UNIVERSITÉ DE LORRAINE

AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10
http://www.cfcopies.com/V2/leg/leg_droi.php
http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm

Archives Henri-Poincaré
Philosophie et Recherches sur les Sciences et les Technologies
(AHP-PReST)

THESE

présentée pour l'obtention du

DOCTORAT DE PHILOSOPHIE

le Jeudi 25 novembre 2021

par

NICOLAS ERDRICH

Les mêmes et les doubles

La tension entre l'imitation et la représentation ou les paradoxes de l'identité fictionnelle

Directeur de thèse : M. Manuel REBUSCHI
Président du jury : M. Jérôme DOKIC

JURY

M. Christophe BOURIAU
Professeur des universités,
université de Lorraine

M. Jérôme DOKIC
Directeur d'études,
EHESS - rapporteur

Mme Ondine BRÉAUD-HOLLAND
ESAP Monaco

M. Filipe DRAPEAU-CONTIM
Maître de conférences,
université de Rennes 1

M. Fabrice CORREIA
Professeur des universités,
université de Genève - rapporteur

M. Manuel REBUSCHI
Maître de conférences HDR,
université de Lorraine

Table des matières

Introduction générale **v**

0.1 L'objet v
0.2 L'enjeu v
0.3 La question vi
0.4 Le projet viii
0.4.1 La méthodologie employée viii
0.4.2 L'état de la question et la démarcation de la thèse viii
0.4.3 Les thèses défendues xi
0.5 Plan analytique du mémoire xiii
0.6 Remerciements xiv

I La possibilité des doubles **1**

Introduction à la première partie **3**

1 Du bon dénombrement des particuliers concrets : identité, unité et persistance **5**

1.1 Identité logique et identité qualitative 6
1.1.1 Identité numérique et identité qualitative 6
1.1.2 Caractériser l'identité comme une relation logique 8
1.1.3 Comparaison entre identité logique et ressemblance 13
1.2 La théorie objectuelle de l'identité 14
1.2.1 La question du rapport au langage 15
1.2.2 Unité et identité 18
1.2.3 La conception frégéenne des liens entre unité et identité 23
1.3 Les paradoxes de l'identité objectuelle 27

1.3.1 Le problème général du changement 28
1.3.2 L'individuation par la constitution matérielle et ses paradoxes 32
1.3.3 Théories endurantiste et perdurantiste 47
1.3.4 L'identité est-elle vague ? 53
1.4 La théorie conventionnaliste de l'identité 61
1.4.1 Objections à la théorie platonicienne de Frege 62
1.4.2 La thèse de l'identité relative de Peter Geach 70
1.4.3 Conventionnalisme de l'unité 75
1.5 Conclusion 85

2 Métaphysique du double **89**
2.1 Des mondes multiples ? 90
2.1.1 L'espace physique et le double 91
2.1.2 L'espace logique et le double 116
2.2 La question des « mêmes » 133
2.2.1 Introduction 133
2.2.2 La question de l'identification transmondaine 134
2.2.3 La question de l'individuation intermondaine 144
2.2.4 Réponses au paradoxe de Chisholm 150
2.2.5 Les haecceités ou essences individuelles 165
2.2.6 Individuation par stipulation et individuation par accointance 173
2.3 La question des indiscernables 176
2.3.1 Pourquoi devrions-nous rejeter la possibilité d'indiscernables ? 177
2.3.2 Le cas des répliques pures 184
2.3.3 Le cas des doubles purs 190
2.3.4 Coïncidents purs 200
2.3.5 Mondes haecceitiques 208
2.4 Conclusion 211

Conclusion de la première partie **215**

II Vers un fictionnalisme intentionnaliste **219**

Introduction à la seconde partie **221**

3 La fiction comme activité de duplication du réel **225**
3.1 Les théories ontologiquement démarquantes . . . 226
3.1.1 La représentation est-elle un double spéculaire du réel ? . . . 226
3.1.2 Théorie mimétique et représentations . . . 234
3.1.3 La valeur didactique des représentations mimétiques . . . 246
3.2 Théories ontologiquement non démarquantes . . . 248
3.2.1 Les théories intentionnelles de la représentation . . . 249
3.2.2 La théorie de l'objet . . . 258
3.2.3 La théorie des objets abstraits . . . 266
3.2.4 Le pan-fictionnalisme . . . 279
3.3 Les mondes possibles comme doubles épistémiques . . . 298
3.3.1 Le perspectivisme nietzschéen . . . 298
3.3.2 Relativiser la vérité aux mondes . . . 302
3.3.3 La vérité en fiction . . . 307
3.3.4 La théorie métaleptique des fictions . . . 318
3.4 Conclusion . . . 321

4 Le texte et ses doubles **323**
4.1 Le texte et ses dupliques . . . 324
4.1.1 Répliques, copies et dupliques . . . 324
4.1.2 Caractériser les reproductions pertinentes d'un texte . . . 327
4.1.3 Équivalence des marques dans un système linguistique . . . 335
4.1.4 Tracer l'authenticité des dupliques . . . 340
4.2 L'œuvre et ses doubles . . . 346
4.2.1 La question de la nature des œuvres littéraires . . . 346
4.2.2 Différentes versions du textualisme . . . 349
4.3 Questions de filiations, de variations et de versions . . . 355
4.3.1 Le textualisme comme essentialisme ontologique . . . 355
4.3.2 Questions philologiques . . . 356
4.3.3 Questions génétiques . . . 357
4.3.4 Questions de traductions, d'adaptations et de reprises . . . 361
4.4 Le problème de Borges . . . 364
4.4.1 Interprétations et identité . . . 364
4.4.2 Écrire une chose pour en dire une autre . . . 366

4.4.3 Le cas Ménard . . . 369
4.4.4 Des textes homographes ? . . . 374
4.4.5 Deux manières d'identifier l'œuvre textuelle . . . 380
4.5 Conclusion . . . 382

5 Le personnage et ses doubles : concevoir le double en fiction 385
5.1 Figures du double en fiction . . . 387
5.1.1 Les analyses du double en fiction . . . 388
5.1.2 Les doubles natifs . . . 393
5.1.3 Les doubles importés . . . 410
5.2 La question de l'individuation des personnages . . . 415
5.2.1 Penser deux fois à un même personnage ? . . . 416
5.2.2 Le réalisme interne . . . 417
5.2.3 Le réalisme possibiliste . . . 426
5.2.4 Le réalisme externe . . . 432
5.3 Pour une théorie dynamique de l'identification des personnages . . . 441
5.3.1 Un réquisit pour une théorie dynamique de l'identité conventionnelle . . . 441
5.3.2 La mêmeté comme relation de ressemblance . . . 443
5.3.3 La théorie des dossiers mentaux . . . 444
5.3.4 La théorie des lignes de monde . . . 448
5.4 Logiques de l'impossible . . . 452
5.4.1 Un exemple littéraire . . . 452
5.4.2 Concevoir des scénarios alternatifs . . . 454
5.4.3 Mondes étranges et mondes merveilleux . . . 459
5.4.4 Contradictions non-adjonctives . . . 461
5.4.5 Hésitation du lecteur et accessibilité à des mondes différents . . . 463
5.5 Conclusion . . . 465

Conclusion de la seconde partie 467

III Des doubles esthétiques aux identités imaginaires 469

Introduction à la troisième partie 471

6 Objets esthétiques et authenticité 473

6.1 Reproduire, imiter, simuler . . . 475
6.1.1 Reproduction et imitation . . . 476
6.1.2 Imitation et simulation . . . 478
6.2 L'artefact esthétique et son double . . . 480
6.2.1 La reproduction des œuvres esthétiques . . . 481
6.2.2 La photographie et son double . . . 484
6.2.3 Le cas des fausses reproductions . . . 488
6.2.4 La question de l'identité spécifique des objets esthétiques . . . 498
6.3 De l'authenticité esthétique à l'esthétique de l'authenticité . . . 501
6.3.1 Connoisseurship et identité personnelle . . . 502
6.3.2 L'esthétisation de l'artiste . . . 504
6.3.3 L'inauthenticité personnelle ou l'homme *kitsch* . . . 506
6.4 L'œuvre d'art *Gesamtkunstwerk* : l'automate parfait . . . 508
6.5 Conclusion : authenticité et reconnaissance . . . 512

7 Fiction de l'identité personnelle et identité fictionnelle de la personne 515
7.1 Les théories de l'identité personnelle ontologiquement engagée . . . 518
7.1.1 Le dualisme cartésien . . . 519
7.1.2 Les théories du Moi absolu . . . 537
7.1.3 Le monisme matérialiste . . . 540
7.2 Pour un désengagement ontologique de la notion de personne . . . 548
7.2.1 Mémoires doubles ou double mémoire . . . 548
7.2.2 Le moi-mêmeté et le moi narratif . . . 562
7.3 La construction psycho-sociologique du moi narratif . . . 570
7.3.1 L'émergence du moi par le jeu de rôle chez Mead . . . 571
7.3.2 Le *doppelgänger* social . . . 574
7.3.3 Pathologies du moi dédoublé . . . 575
7.3.4 Imiter le désir d'un autre . . . 579
7.4 Se voir comme un autre . . . 582
7.4.1 De l'auto-identification spéculaire à l'auto-identification imaginaire . . . 583
7.4.2 Se voir voir et voir à travers . . . 588
7.4.3 L'illusion spéculaire . . . 589
7.5 S'imaginer comme un autre et se croire autre . . . 599
7.5.1 Imagination et pensées personnelles . . . 600

7.5.2 La duplication de l'imaginant en imaginé dans l'imagination *de se* 610
7.5.3 La question du double point de vue 620
7.6 Conclusion 626

Conclusion générale **629**
8.1 Rappel de la problématique 629
8.2 Rétrospective des grandes lignes 630
8.3 Réponse à la problématique 635
8.4 Quelques pistes de réflexion 636

À Alain

Introduction générale

LE PROFESSEUR : Bon. Arithmétisons donc un peu.
L'ÉLÈVE : Oui, très volontiers, Monsieur.
LE PROFESSEUR : Combien font un et un ?
L'ÉLÈVE : Un et un font deux.
LE PROFESSEUR (émerveillé par le savoir de l'Élève) : Oh, mais c'est très bien. Vous me paraissez très avancée dans vos études. Vous aurez facilement votre doctorat total, Mademoiselle.

Ionesco, *La leçon.*

One of the hardest things in philosophy sometimes is to know the difference between a profound problem and a silly question.

Graham Priest, *One*

0.1 L'objet

Ce mémoire est consacré à l'étude de la notion de double en philosophie. Le propos qui va suivre peut sans doute être résumé en une phrase : je souhaite défendre l'idée que les doubles n'ont aucun poids ontologique. Toute chose est singulière, quoique toute chose puisse s'avérer, épistémiquement parlant, le double d'une autre.

0.2 L'enjeu

Lorsque nous parlons de « double » dans un contexte pratique, dans une plaisanterie ou dans le résumé d'un film fantastique, notre interlocuteur sait généralement de quoi il retourne : soit qu'il s'agisse d'une quantité deux fois plus grande, de la reproduction d'un objet, d'une personne extrêmement ressemblante à une autre ou plus généralement de rapports perturbés à l'identité. Mais lorsque nous souhaitons analyser plus précisément ce que nous entendons par « concept de double », l'idée paraît néanmoins plus abstruse. La célèbre remarque d'Augustin, avouant, au livre XI des *Confessions*, une incompétence à définir ce qu'est le temps en dépit de sa compréhension

immédiate du terme pourrait certainement aussi être formulée à propos du double : nous voyons bien de quoi il retourne, mais sitôt qu'on entreprend d'en expliquer les faits, les mots nous fuient et nous nous trouvons dans l'embarras d'en préciser la nature exacte. La question qui oriente le présent travail est précisément d'analyser cette notion.

A quelques exceptions près, les philosophes utilisent le terme de double sans en circonvenir la terminologie, sans en préciser les acceptions retenues. L'expression apparaît dans leurs travaux comme il apparaît aussi dans la vie courante, dans le langage ordinaire. Or les usages de la notion de double indiquent une variété importante, la thématique déployant un ample champ lexical aux analogies éparses. Le mot « double » connote certes des expressions dont l'emploi est précis et délimité, comme « jumeaux » et « clones » mais également des expressions comme « copies », « répliques », « dupliques », « imitations », « simulations », « représentations », « reproductions », « répétitions » dont l'application est non seulement d'une étendue large, mais aussi ambiguë. En conséquence, l'usage de la notion de double génère bien souvent du vague et débouche fréquemment sur des confusions. Décrire ou juger de telle chose, de tel objet, voire de telle personne qu'il est un double de tel autre implique ainsi diverses difficultés. Les explications de la nature de l'expression de double, de son fonctionnement, mais aussi les pratiques qui l'entourent sont imbues de problèmes liés à ces confusions.

Parvenir à déterminer ce qui tombe sous le régime du double présuppose alors de connaître les critères appropriés entrant dans la définition du concept afin de reconnaître les objets auxquels il peut s'appliquer. Or quel genre de propriétés pourrait nous autoriser à attribuer cette notion à un individu ? Serait-ce par des qualités propres et objectives ? Est-il dans l'essence de certaines choses d'être des doubles d'autres choses ? Ou s'agirait-il au contraire de qualités rapportées et seulement subjectives, témoignant de manières proprement humaines de découper le poulet-monde aux articulations ?

0.3 La question

Nous venons d'indiquer l'objet de ce travail, et nous venons d'en décrire l'enjeu global. Précisons maintenant le problème qui se pose, avec la notion de double, en philosophie. La question est la suivante : un double a-t-il un poids ontologique en tant que tel, ou ne s'agit-il que d'un être de langage, une étiquette pour parler de choses, en réalité, uniformément singulières ? Les doubles sont-ils des imitations objectives, des similaires ontologiques, ou s'agit-il de représentations rapprochées par des similarités fonctionnelles ? Peut-on ainsi cliver le double en deux pôles, les doubles par imitation et les doubles par représentation, en dépit des tensions qui tendent à les rapprocher ou à les éloigner ? Enfin, en fonction de nos réponses à ces questions, quelles conclusions devrions-nous alors tirer de la catégorie du double en ce qui concerne nos actions, nos jugements, et plus particulièrement encore, notre rapport à nous-même et à autrui ?

Comme le fait remarquer le professeur dans la pièce de Ionesco citée en épigraphe, si « un et un font deux », c'est que le double, dira-t-on, c'est deux fois l'unité. Le double se rapporte ainsi à une simple opération arithmétique dont la résultante est la constatation qu'une chose et encore une chose, ça fait bien deux choses. Si l'apparente banalité d'un tel truisme peut prêter à sourire, je prétends, à mon tour, hisser son questionnement au niveau d'un doctorat d'université.

Cette perspective de travail est stimulée par le fait que le dénombrement de choses peut parfois prendre une tournure complexe. Il n'y a qu'à penser aux cas des flaques, des nuages, des trous, des particules quantiques pour rendre nos décomptes moins assurés. Nous aurons amplement l'occasion de constater dans ce qui suit que c'est sans doute aussi le cas d'entités tenues généralement pour mieux délimitées, comme les bateaux, les statues, les chats et les personnes. Nous avons donc là un problème de décompte, tenant à la manière de considérer que des choses puissent tenir le rôle d'unité. La question est alors de savoir si la valeur d'unité est ontologique ou seulement épistémique.

À cette fin, commençons par constater que le double se manifeste par une sorte de balancement de la raison entre deux pronostics vraisemblables mais contradictoires. D'une part un pronostic d'identité, d'autre part un pronostic de distinction. Le sentiment du double naît ainsi de l'oscillation entre ces deux alternatives. Avons-nous affaire à *une* chose ou avons-nous affaire à *deux* choses ? Confrontés à un problème de double, il semble que nous ayons autant de raisons d'admettre l'une des alternatives que l'autre. Mais pouvons-nous concilier les deux hypothèses sans céder à l'absurde ? Quelle que soit la réponse apportée à cette question, il est manifeste que travailler sur le double, c'est rôder dans les marges de l'identité logique, s'intéresser à ses cas pathologiques, s'il en est.

Considérons une première illustration. Dans la nouvelle fantastique de Fitz James O'Brien intitulée *The Lost Room*, le personnage principal décrit ainsi sa chambre : « The room was, yet was not mine; and a sickening consciousness of my utter inability to reconcile its identity with its appearance overwhelmed me, and choked my reason. » (O'BRIEN, 1995). L'expression « la pièce était et n'était pas ma chambre » présente une illustration du double, à savoir, un jugement d'apparence paradoxale mêlant identité et différence. Et comme le reconnaît l'auteur, un tel jugement a sans doute pour effet de « choquer la raison ». On pourrait bien sûr n'y voir, d'un strict point de vue logique, qu'une pure absurdité et ne consentir qu'à reconnaître une sorte d'illusion, fut-elle volontaire ou non. Sous cette perspective, l'analyse du double incomberait alors entièrement à la psychologie ou à la psychanalyse, voire à la psychiatrie. À l'ampleur des travaux qui y sont consacrés, il faut d'ailleurs reconnaître que ces disciplines ont fait du double un sujet de prédilection. Mais à côté des caractérisations psychologiques, il est sans doute possible de s'intéresser aussi aux composantes sémantiques, logiques et métaphysiques du double. Car c'est un fait que nous comprenons bon nombre de nouvelles fantastiques qui prétendent confondre ainsi l'identité et la diversité. Le présent travail se propose d'explorer de telles possibilités et se voudrait donc être une contribution, aussi modeste soit-elle, aux enjeux métaphysiques et épistémiques du double.

Attardons-nous encore sur une seconde illustration, celle-ci tirée de la perception plutôt que de la fiction : il est un petit exercice de gymnastique orthoptiste servant de rééducation aux patients qui sont sujets à un strabisme divergent. Il s'agit d'observer deux images, séparées et tendues à bout de bras, en focalisant le regard de manière à ne percevoir qu'une seule image au lieu de deux. Autrement dit, il s'agit de faire converger les deux images de façon à ce qu'elles coïncident entièrement. Ce phénomène illustre sans doute certains usages, motivés par différentes raisons pratiques, qui nous incitent à considérer que ce qu'on tient habituellement pour deux choses ne compte que pour une seule, voire inversement, que ce qu'on tient habituellement pour une chose vaille en fait pour deux. La question est alors de déterminer dans quelle mesure l'unité ne relève que de simples conventions et dans quelle mesure elle témoigne au contraire d'une réalité méta-

physique indépendante de nos manières de compter. De fait, ce qui motive la question du double est certainement lié à nos façons de concevoir le monde.

La question du double peut ainsi être posée selon deux hésitations conflictuelles allant dans des directions opposées, de l'identité à la diversité pour l'une, et de la diversité à l'identité pour l'autre. Le premier mouvement consiste ainsi à présumer d'emblée l'identité, mais à la remettre très vite en question sous le poids d'arguments venant perturber cette hypothèse. Je désignerai sous le nom de question des mêmes cette première catégorie de problèmes, en conservant la marque du pluriel pour souligner le passage de l'unité à la diversité. Le deuxième mouvement consiste inversement à stipuler d'abord une distinction numérique, distinction qui se trouve ensuite remise en question sous la pression de ressemblances frappantes. C'est ce que je désignerai plus proprement sous le terme de problème des doubles. Ces deux tendances inverses indiquent la raison pour laquelle nous parlons, au sujet du double, d'une tension entre l'imitation et la représentation, comme l'indique le sous-titre de la thèse.

0.4 Le projet

0.4.1 La méthodologie employée

Sans surprise, ce travail s'est heurté à différentes difficultés méthodologiques. La plus importante mérite d'être mentionnée. Elle fut bien évidemment de circonscrire le problème du double, et donc son corpus littéraire, qui, pour peu qu'on ne le retienne par de prudentes digues, risquait à tout bout de champ d'inonder sans retenue le contenu par son étendue encyclopédique. Les digues, maintes fois rompues, furent maintes fois étayées, mais, quoi qu'il en soit, il était de toute façon difficile de ne pas laisser certains torrents nous porter, creusant des ravines que nous avons canalisées selon trois principaux défluents du double constituant chacun une des trois parties de ce travail. Le premier, qui relève de la question de l'imitation, est celui des choses qui sont ou qui pourraient être des ressemblants parfaits. Le second canal, celui de la représentation, irrigue les démarcations entre le réel, l'irréel et l'illusoire. La fiction est-elle un double du réel ? Enfin, le troisième canal, est la question de l'identité personnelle, du rapport du double à soi-même.

La question du double sera traitée ici comme une question de philosophie du langage et comme une question de philosophie de l'esprit. La philosophie analytique, pour ses méthodes et pour ses résultats, est accueillie dans les pages qui suivent, sans rejeter toutefois une approche plus « continentale », si l'expression a un sens, de l'enquête philosophique.

0.4.2 L'état de la question et la démarcation de la thèse

Bien que notre approche est essentiellement philosophique, il n'est pas souhaitable d'écarter trop rapidement les autres disciplines intéressées par la notion de double. Je détaillerai plus explicitement, dans le corps du mémoire, l'état de la question selon ces disciplines. Mais relevons déjà quelques jalons.

Le double est un concept très prisé en psychologie et en psychanalyse. Les travaux d'Otto Rank (RANK, 1973), de Sigmund Freud (FREUD, 1985c) considèrent le double comme un révélateur de

complexes latents, et André Green (GREEN, 1980) comme une instance éminente de la subjectivité par une sorte de dialogue de la psyché entre un corps réel et un corps absent.

En anthropologie et en sociologie, le concept est employé par René Girard (GIRARD, 1983b), pour fournir une explication aux rivalités sociales liées aux désirs mimétiques. Chez George Herbert Mead (MEAD, 1934), la notion est utilisée pour clarifier les processus d'émergence du sentiment d'identité personnelle au sein de jeux d'imitation. Chez Helmuth Plessner (PLESSNER, 1960), le double intervient pour expliquer le processus d'incarnation de rôle dans nos activités sociales, nos relations interpersonnelles.

Dans le domaine de la fiction, la critique renonce à donner une définition unanime du double, et elle centre essentiellement la problématique autour de la question de savoir s'il s'agit d'un genre littéraire ou d'une figure présente dans tous les genres.

Dans la littérature philosophique, les doubles surgissent généralement au cours d'expériences de pensées menant à des paradoxes. Ils sont employés pour évaluer la validité d'une thèse, en révéler les contradictions. Les indiscernables apparaissent notamment dans les questions autour de la nature des œuvres esthétiques. Goodman (GOODMAN, 1976) utilise par exemple la résistance des œuvres à la copie pour distinguer deux types de productions artistiques, d'une part les arts allographiques qui regroupent les œuvres dont l'identité résiste à la duplication, et d'autre part les arts autographiques, qui regroupent celles qui ne résistent pas à l'épreuve de la reproduction. Après Goodman, Arthur Danto (A. DANTO, 1981), utilise les indiscernables comme des sortes de témoins, au sens expérimental du terme, permettant de caractériser ce qui fonde la nature esthétique d'une œuvre.

Ces diverses études ont toutes pour point commun de considérer le double comme un outil conceptuel plutôt que comme un objet d'étude. A cette fin, les auteurs visent un aspect plus ou moins particulier du double pour satisfaire à une certaine exigence d'explication. Mais jusqu'à présent, il faut bien reconnaître que les philosophes n'ont pas étudié la question du double pour elle-même. Clément Rosset est toutefois l'un des rares auteurs à prendre le double, oserai-je dire, par les cornes, à lui consacrer une place de taille dans son interrogation philosophique. Mais son propos est essentiellement centré sur la caractérisation du réel. Or, dans l'usage qu'il fait du concept, Rosset n'a pas manifestement pas préalablement caractérisé les divers types de doubles avec suffisamment de netteté, comme il en convient d'ailleurs lui-même dans les premières pages d'*Impressions fugitives* (ROSSET, 2004). Le présent mémoire souhaite s'atteler à combler ce manque en abordant de manière analytique le concept polysémique de double.

Comme je l'ai annoncé plus haut, le but de ce travail est de clarifier le sens du terme « double ». De façon plus précise, l'enquête qui suit vise à cerner les dichotomies qui prétendent distinguer le double du singulier, l'indiscernable de l'identique, la reproduction de l'original, le différent du même. Il n'y a aura pas lieu ici de nier ces oppositions majeures, en souscrivant à une dialectique duale des contraires, à un point de vue dialethéiste cédant aux contradictions vraies. Il s'agira plutôt de saisir les ambiguïtés de la notion de double, afin de les enrégimenter dans une catégorisation plus fine et tributaire des contextes d'usage.

La question est d'abord de déterminer comment nos dénombrements sont articulés sur l'identité logique. La problématique des mêmes regroupe les questions se rapportant à la persistance temporelle, aux relations parties-tout et aux relations modales. Premièrement, des arbres, des chats,

des bateaux sont des choses qui persistent en évoluant dans le temps, autrement dit, qui sont modifiées tout en demeurant identiques à elles-mêmes. Remarquons deuxièmement qu'un arbre, qu'un chat et qu'un bateau sont des unités constituées de parties diverses. En ce sens, notre manière de concevoir de telles choses semble répondre à une sorte de processus dialectique passant de la singularité au pluriel, et inversement, du pluriel à la singularité. Troisièmement, constatons que des choses matérielles disposent d'un certain degré de liberté dans le possible, et qu'elles pourraient, ou qu'elles auraient pu, sous certaines conditions, s'avérer différentes de ce qu'elles sont effectivement. Ce chat aurait pu ne pas avoir sa queue tranchée par le passage d'une voiture. Ce livre aurait pu être moins défraîchi qu'il ne l'est actuellement. Ces situations indiquent que nous pouvons faire référence à une seule et même chose quoique celle-ci puisse s'avérer différente par sa constitution temporelle, par ses parties spatiales et par ses dispositions. Dans de tels cas, il semble que nous présumions d'emblée l'identité numérique, mais que ce constat soit mis en demeure de répondre à des sollicitations contraires engendrées par des altérations temporelles, spatiales et modales.

D'un autre côté, nous pouvons aussi invoquer la figure du double pour parler de choses qui, quoique numériquement différentes, nous apparaissent comme étrangement similaires. Dans de tels cas, c'est plutôt la distinction numérique qui est d'emblée présumée, mais elle est aussitôt remise en question par une familiarité par trop manifeste. Quelle place les conceptions de l'identité laissent-elles à la notion de ressemblance parfaite ? Pourrait-il exister des entités absolument indiscernables ? De quoi s'agirait-il alors ? Pour nombre de philosophes partisans du réalisme des propriétés, la ressemblance parfaite doit être rabattue sur l'identité. Autrement dit, il ne se pourrait pas que x soit parfaitement ressemblant à y sans que x soit identique à y. Deux individus diffèrent toujours par quelque qualité, c'est du moins ce qu'affirme le principe d'identité des indiscernables. Comme l'écrit Leibniz à Arnault : « il n'est pas possible qu'il y ait deux individus entièrement semblables ou différents *solo numero*. » (Gottfried Wilhelm LEIBNIZ, 1989, p. 120) Nous montrerons au cours d'une discussion de ce principe qu'apparaissent différents sous-genres d'indiscernables que je désignerai respectivement par les noms de répliques pures, de doubles purs et de coïncidents purs.

La conclusion de la première partie nous mènera à constater ainsi l'émergence, en philosophie, de deux types de doubles ontologiques, d'une part les doubles-mêmes, qui relèvent des questions d'identité croisée entre mondes possibles, et d'autre part les doubles par indiscernabilité.

Le passage à la fiction, seconde partie de ce mémoire, ouvrira ensuite la porte aux questions d'intentionnalité. Dans quelle mesure y a-t-il des doubles épistémiques ? De fait, la fiction est particulièrement intéressée par la notion de double. Soit que, de manière générale, la fiction reproduit par représentation la réalité, qu'elle en produit une imitation. Soit que, de manière externe, l'authentification de l'œuvre textuelle passe par la question des reproductions correctes des textes qui la supportent. Soit encore que, d'un point de vue interne, la fiction propose des imaginaires du double, qu'elle nous incite à concevoir des personnages dupliqués, comme le William Wilson de Poe (E. POE, 1932) ou comme le Goliadkine de Dostoïevski (DOSTOÏEVSKI, 1969). Il semble important ici de confronter quelques exemples de doubles littéraires aux théories métaphysiques de l'identité objectuelle qui, pour rendre compte de nos usages fictionnels, hypostasient des entités exotiques, des inexistants, des simplement possibles ou des artefacts abstraits afin d'individuer les personnages. Nous défendrons l'idée que de tels doubles ontologiques n'existent pas, que l'identification des personnages de fiction relève de négociations consensuelles enrégimentées par des

pratiques culturelles.

Nous conclurons la seconde partie par la nécessite d'introduire, à côté des doubles ontologiques, un type épistémique de double. Les dupliques fonctionnelles nous serviront par exemple pour qualifier les reproductions d'une œuvre de type allographique.

Nous passerons alors, dans une troisième partie, à la question de l'authenticité, en particulier à son rapport à l'identité spécifique. En effet, la question de l'identification des dupliques d'une œuvre textuelle débouche sur la question plus générale de l'authentification des œuvres esthétiques. La critique de l'authenticité de l'objet esthétique mène ensuite à une critique de l'authenticité de l'artiste, et de là, la question de l'identification du personnage de fiction rejaillit sur la question de l'identité personnelle. Après avoir critiqué les théories substantialistes du moi d'être des théories proclamant l'existence de doubles ontologiques, il s'agira de considérer les rapports de duplications imaginaires que nous mobilisons pour nous doter d'une identité personnelle, dont la nature est narrative.

0.4.3 Les thèses défendues

L'idée générale qui est défendue dans cette thèse consiste à promulguer les doubles comme des instances épistémiques notoires, et à rejeter les doubles ontologiques. Sans la possibilité de tenir une chose pour deux, autrement dit, de voir double, et sans la possibilité inverse de n'en voir qu'une là où s'en tiennent en réalité deux, nous serions tout simplement incapables de produire de nouvelles identifications. Remarquons à cet égard que, même lorsque nous avons mis à jour une identité, comme celle entre Phosphorus et Hesperus, nous sommes encore capables de faire comme s'il s'agissait de deux astres différents, par exemple lorsque nous considérons le système astronomique babylonien. Nous admettrons plus généralement qu'une chose peut toujours, sous un angle donné, s'avérer le double d'une autre. En ce sens, le terme de double par lequel nous qualifions les copies, les reproductions, bref, qui qualifie des objets ou des événements indiscernables apparaît aussi légitime que trompeur lorsque son emploi prétend à une application absolue.

Le point de départ du travail préparatoire à ce mémoire a consisté à admettre la thèse que Clément Rosset soutient à longueur d'ouvrages, affirmant que le Réel est singulier, autrement dit, sans double. La conséquence est donc que, de manière contraposée, ce qui peut être qualifié de « double » caractérise très exactement ce qui n'est pas réel, disons, irréel. Pour ma part, il s'agissait d'examiner jusqu'à quel point pouvait tenir cette thèse, et en particulier, de déterminer les points de divergence entre doubles fictionnels et les doubles illusoires que relève Rosset lui-même, à la fin de sa carrière, dans le court essai *Impressions fugitives* où il reconnaît qu'il existe « certains doubles qui sont au contraire des signatures du réel garantissant son authenticité » ((ROSSET, 2004, p. 9)). La question du double est donc aussi de savoir dans quelle mesure certains doubles relèvent du réel, de la fiction ou de l'illusoire. La distinction entre l'irréalité de la fiction et le caractère illusoire de certains jugements est justifiée en ce que, si ces domaines possèdent très certainement une intersection non vide, ils ne semblent néanmoins pas coïncider complètement. Les fictions comportent en effet des informations sur le réel, ou, pour filer des métaphores, paraissent en constituer des ombres, des reflets ou des échos. A l'inverse, l'illusoire est toujours, en terme de connaissance, stérile, sinon en contradiction complète avec ce qui est réel. Discriminer entre, d'une part, les fictions du double, et d'autre part, les illusions du double, exige donc de clarifier la

distinction entre la fiction et l'irréel. Soulignons d'emblée que le fait qu'une pensée soit sans objet réel ne signifie pas nécessairement qu'elle s'avère également neutre ou sans effet sur la réalité. Les religions et les idéologies politiques témoignent au contraire d'une puissance causale par delà leur caractère irréel.

Résumons-nous : le réel est singulier, sans double. Mais il existe des doubles qui témoignent du réel. Avons-nous affaire à une contradiction ? Est-elle seulement apparente ? Comment la surmonter ?

Les thèses qui sont défendues au fil de ce travail veulent répondre à cette question. Elles sont les suivantes :

Thèse principale Il existe en philosophie, un « double » problème du double. L'un est métaphysique, relevant de l'illusoire, et l'autre est épistémique, relevant d'une sorte de connaissance par la fiction, par l'irréel. Le sous-titre du mémoire est motivé par le fait que les doubles épistémiques relèvent de représentations liées par des similarités fonctionnelles.

Sous-thèse 1 Je défendrai l'hypothèse qu'il n'y a, métaphysiquement parlant, pas de doubles ontologiques. Si une théorie est amenée à en induire, il ne s'agit pas d'authentiques entités, mais de paradoxes causés par une mauvaise interprétation de l'identité. Les principales théories engagées en faveur de l'identité métaphysique induisent l'introduction de doubles métaphysiques, soit qu'il s'agisse d'entités coïncidentes, comme celles qui émanent de la théorie sortaliste de David Wiggins, soit qu'il s'agisse de contreparties alter-mondaines dans la théorie du même nom développée par David Lewis, soit qu'il s'agisse d'entités menant des « double vies » comme dans la théorie essentialiste de Saul Kripke.

Sous-thèse 2 Il y a une certaine liberté de regrouper sous le terme d'« objet » différentes ostensions. Nous pouvons dans ce cas parler de doubles *épistémiques*. De tels doubles sont, par nature, ontologiquement inoffensifs.

Sous-thèse 3 L'identité des personnages de fiction ne relève pas de l'identité d'entités fictionnelles, mais de consensus d'identifications dans des pratiques culturelles.

Sous-thèse 4 Je défendrai l'idée que la notion métaphysique d'identité personnelle, entendue comme substance du moi, relève d'un double métaphysique, et qu'elle s'avère donc, de ce fait, illusoire. Je propose de la remplacer par une notion d'identité fictionnelle qui s'alimente de rapports d'identification *de se* à des doubles épistémiques.

Tout ce qui suit prétend défendre ces thèses.

0.5 Plan analytique du mémoire

Partie 1 La première partie a trait à la tension entre identité qualitative et identité numérique. Deux conceptions de l'identité, objectuelle et conventionnaliste, sont détaillées. Deux catégories de doubles ontologiques, les doubles-mêmes et les doubles par indiscernabilité sont ensuite discutées.

Chapitre 1 Dans ce chapitre, il s'agira d'inscrire le double dans un rapport non contradictoire à l'identité. Au cours de ce premier chapitre, on demandera, avec insistance, « une ou deux choses ? ». Deux types de questions apparaissent. La première, épistémique, engage à déterminer ce en quoi nous pouvons savoir que nous avons affaire à une chose plutôt qu'à deux ou, au contraire, à deux choses plutôt qu'à une seule. La deuxième question, métaphysique, consiste à déterminer ce qui fait que la chose visée est une chose plutôt que deux, ceci indépendamment de toute intention à son égard. A cet effet, on synthétise et on critique les théories de l'identité. 1.1 On commence par distinguer l'identité numérique de l'identité qualitative, en les comparant. 1.2 On introduit ensuite la question de l'identité objectuelle, identité métaphysique que posséderait tout objet singulier. 1.3 On rappelle les paradoxes temporels et méréologiques sur lesquels achoppe l'identité métaphysique. 1.4 On opte finalement pour une théorie conceptualiste de l'identité en déclinant les versions relativiste et conventionnaliste. C'est cette seconde option que j'adopte.

Chapitre 2 A cette étape, nous passerons à l'étude des doubles métaphysiques en analysant ce que pourraient être de telles choses, s'il y en avait. 2.1 Nous commençons par traiter la question des mondes répétés, qui nous mène 2.2 à la question des identités croisées entre mondes possibles. Nous abordons alors le problème des doubles-mêmes. Le paradoxe de Chisholm, central pour la discussion, y est analysé. 2.3 Le principe d'identité des indiscernables est discuté et trois types de doubles ontologiques par indiscernabilité sont mis à jour, les répliques pures, les doubles purs et les coïncidents purs, dont on cherche ensuite le degré de réalité.

Partie 2 La seconde partie traite des fictions. Il s'agit de mettre en relief les articulations entre possibilités logiques et possibilités métaphysiques. Il s'agit d'interroger cette fois la notion de double du point de vue de la production humaine. Nous poserons dans cette partie essentiellement trois questions, la première étant de savoir s'il convient de considérer la fiction comme un double du réel, la seconde étant de déterminer en quoi consiste la production d'un double textuel, la troisième étant d'analyse en quoi consiste le fait de concevoir le double en fiction.

Chapitre 3 La fiction pose, par son activité propre, un problème de double : s'agit-il d'une duplication mimétique du réel ? La question est ancienne, elle est détaillée dans la section 3.1, pour être ensuite abordée selon les diverses réponses apportées. Dans 3.2 nous discutons des théories ontologiquement non démarquantes pan-réalistes et pan-fictionnalistes avant 3.3 de considérer une autre voie, utilisant les mondes possibles comme des doubles épistémiques de la réalité.

Chapitre 4 La fiction est médiatisée par des textes, ces textes sont assujettis à des reproductions matérielles qui manifestent la participation, dans le monde, de l'œuvre textuelle. Pour décrire ce qui constitue la nature d'une œuvre textuelle, nous introduirons la notion de duplique, double fonctionnel d'un texte. Dans 4.1 nous abordons le rapport du texte à ses reproductions. Nous comparons les dupliques aux répliques et aux copies. Nous passons ensuite, dans 4.2 à la question de la nature des œuvres textuelles. Il s'agit de situer l'œuvre textuelle par rapport au texte. Nous sou-

tenons que les œuvres textuelles n'ont pas de poids ontologique, qu'elles ne sont pas des doubles ontologiques des textes. L'enjeu de la section 4.3 est de traiter la question des versions, des traductions et des interprétations, phénomènes qui semblent diffracter l'unité du texte en doubles. Dans 4.4 nous passons à l'examen du problème des textes indiscernables.

Chapitre 5 La fiction nous confronte à concevoir le double. La question est de déterminer quels sont les paradoxes du double qui apparaissent dans ce cadre. Dans la section 5.1, on considère les figures du double en fiction. Les problèmes afférents à l'identification des personnages sont exposés dans la section 5.2. La section 5.3 propose des éléments pour une théorie contextuelle de l'identification des personnages. La section 5.4 traite de la manière de gérer, à l'aide des mondes possibles, certaines identités impossibles.

Partie 3 La troisième partie s'intéresse aux rapports de l'authenticité et de l'identité personnelle aux doubles épistémiques. L'enjeu est de rejeter les théories substantialistes de l'identité au profit d'une théorie constuctiviste de l'identité, basée sur la notion de double épistémique.

Chapitre 6 Ce chapitre offre une transition vers la question de l'identité personnelle via la question de l'authenticité des objets dits esthétiques. Après avoir mis en évidence, dans la section 6.1, la distinction entre imitation et reproduction, nous examinons, dans la section 6.2, la notion de double appliquée aux objets matériels, sous l'angle des diverses pratiques que constituent les reproductions, les dupliques, les contrefaçons, les faux. Dans la section 6.3, nous nous concentrerons sur le cas des objets dits esthétiques, qui intéressent particulièrement la notion d'authenticité. Dans la section 6.4, nous considérerons le rêve de l'œuvre d'art idéale, l'automate parfait, dans son rapport à l'imitation et à la simulation.

Chapitre 7 Il s'agira dans ce chapitre de souligner l'importance de la notion de personnage pour la constitution de l'identité personnelle. Dans la section 7.1, je présente et critique les théories substantialistes de l'identité personnelle. Dans la section 7.2, j'adopte une thèse fictionnaliste de l'identité personnelle : l'identité est narrative. La sociologie et la psychologie nous aident ensuite à mettre à jour les figures du double dans leur rapport à la constitution d'un moi singulier. Dans la section 7.3, on suit une étape importante de la construction du moi par les jeux de rôle et les duplications imaginaires. On analyse ensuite plus finement les phénomènes liés à l'imaginaire du double en y décrivant deux instances de doubles épistémiques, d'une part la figure-écho, symptomatique des comportements par imitation qui renvoie au sujet de manière réflexive sa position dans le monde. D'autre part, la figure-filtre relative aux manières de faire-comme si l'on était un autre en adoptant sa position. La section 7.4 est consacrée à la reconnaissance autoscopique, et la section 7.5 aux identifications imaginaires.

0.6 Remerciements

Ce mémoire est l'aboutissement d'un travail mené durant plusieurs années parallèlement à l'exercice de professeur de mathématiques dans le secondaire. Il a été enrichi par de nombreuses contributions, directes ou indirectes.

Je dois d'abord remercier mon directeur de thèse, Manuel Rebuschi pour son soutien continuel et la générosité avec laquelle il a suivi, critiqué et enrichi par ses conseils le présent travail.

Je remercie également les membres des Archives Poincaré pour leur accueil et leur accompagnement, d'abord au sein du master MADELHIS, et ensuite dans l'encadrement de la thèse, en particulier Christophe Bouriau, Baptiste Mélès, Philippe Nabonnand, et Roger Pouivet.

Je remercie aussi le professeur Claudine Tiercelin et Alexandre Declos pour leur accueil au sein du séminaire du Groupe d'Études en Métaphysique au Collège de France. Merci aussi aux participants du séminaire pour leurs critiques et conseils, Guillaume Bucchioni, Muriel Cahen, Filipe Drapeau Vieira Contim, Vincent Grandjean, Jean-Baptiste Guillon, Jean-Maurice Monnoyer, Frédéric Nef et Pierre Saint-Germier.

Je remercie les participants du séminaire *Logiques et philosophie, entre paraconsistance et paradoxes* pour leurs discussions passionnantes : Stefan Jokulsson, Marion Renauld, Philippe Ruisseau, Florent Schoumacher et Guillaume Schuppert.

Je remercie également les différents collègues de la cité scolaire André Maurois qui ont encouragé, voire enrichi, d'une façon ou d'une autre ce travail : Brigitte Brocard, Estelle Burlett, Guillaume Funfrock, Aline Werlé.

Je remercie la communauté Emmaüs pour l'approvisionnement en livres, une partie importante de la bibliographie de ce mémoire ayant été dénichée sur leurs rayonnages.

Merci enfin à ma « double » famille, François et Dominique d'un côté, Marie-Anne et Alain de l'autre, qui ont soutenu de toutes les manières possibles ce travail.

Enfin, merci à mon double, Angélique, qui, elle, n'avait pas besoin de tant d'années de recherches pour savoir ce que cela signifie. Ce travail n'aurait pas pu être réalisé sans son aide quotidienne.

Première partie

La possibilité des doubles

Introduction à la première partie

Une pierre
deux maisons
trois ruines
quatre fossoyeurs
un jardin
des fleurs
un raton laveur

Jacques Prévert, *Inventaire*

La question du double est sans doute, et avant tout, une question de juste dénombrement. Les ethnologues Edward Burnett Tylor (TYLOR, 1920), Alf Sommerfelt (SOMMEFELT, 1938) et Archibald E. Hunt (HUNT, 1899) ont remarqué que les systèmes de numération de certaines peuplades, voire de certaines civilisations, se limitent à n'employer que deux chiffres différents représentant l'unité et la paire [1]. La base binaire a d'ailleurs retrouvé un usage fondamental en informatique. Mathématiquement parlant, les deux chiffres « 1 » et « 2 » suffisent pour représenter tous les autres multiples de l'unité puisque toute autre valeur entière peut être obtenue par des opérations successives ou par des relations biadiques [2]. En outre, le théorème de décomposition en base 2 certifie que n'importe quel entier n supérieur ou égal à 0 peut s'écrire sous la forme $a_p \times 2^p + a_{p-1} \times 2^{p-1} + ... + a_0$ où $\forall i \in [\![0, p]\!]\ a_i \in \{0, 1\}$. Il conviendra donc de garder en vue, dans ce qui suit, que le passage de « un » à « deux » sert toujours aussi de passage de l'unité à n'importe quel de ses multiples. Le double indique non seulement la répétition de l'unité ajoutée une fois, mais aussi la répétition de cette même unité à un multiple quelconque. Attribuer le double, c'est donc, aussi, attribuer le triple, le quadruple, etc.

Si le double désigne une quantité de choses deux fois plus grande, il reste à savoir ce qui détermine, pour un objet visé, l'unité (ce qui lui permet d'être l'étalon d'un dénombrement) et l'unicité (le fait d'être unique). L'identité est-elle un fait objectif, une propriété réelle qui constitue la singularité d'une chose ? Ou doit-on seulement parler de *jugements d'identité* pour désigner des manières proprement humaines d'individuer des choses ? La question relève de nos activités ordinaires de dénombrements, et de la place de l'identité dans ces rapports.

Cette première partie est donc consacrée aux relations entre l'identité et le double lorsqu'on a

[1] On pourra consulter à ce sujet l'*Histoire Universelles des Chiffres* de Georges Ifrah (IFRAH, 1981).

[2] Par exemple, si le nombre 2 est représenté par l'écriture (1 ; 1), le nombre 3 peut alors être représenté par l'écriture (1 ; (1 ; 1)), 4 par ((1 ; 1) ; (1 ; 1)), et ainsi de suite.

affaire à des objets concrets.

Chapitre 1

Du bon dénombrement des particuliers concrets : identité, unité et persistance

> What's true of the many is not exactly what is true of the one. After all they are many while it is one.
>
> David (LEWIS, 1991, p. 87)

L'objet du chapitre est de distinguer deux sortes de théories de l'identité, l'une métaphysique, l'autre, épistémique. J'appelle identité objectuelle la thèse que les choses possèdent une identité métaphysique absolue, indépendante du contexte d'application.

L'enjeu du chapitre est d'expliquer notre capacité à adapter nos dénombrements, notre application de l'identité numérique aux objets qui persistent dans le temps.

La question est alors de savoir s'il y a quelque chose comme une identité métaphysique des objets, une identité qui ne dépende pas des intentions. En souscrivant à cette idée, on s'enlise dans de nombreux paradoxes que je rappellerai. Aussi, je favorise une théorie conventionnaliste de l'identité. Ma conclusion est que l'identité logique n'est ni vague, ni relative, mais son application est contextuelle, ce qui permet d'expliquer qu'une chose puisse compter pour une ou pour deux selon le contexte.

Dans ce qui suit, j'appellerai « paradoxes du double » les paradoxes qui se ramènent à la forme générale $a = b \wedge a \neq b$. On le voit, le problème du double relève ainsi d'une tension occasionnée entre le singulier et le pluriel, entre l'unicité et la multiplicité, entre l'identité et la différence. Comme l'écrit Gérard Genette, le double, c'est « à la fois un *autre* et un *même* » (GENETTE, 1966, p. 21). Ainsi, le problème du double relève autant de la signification de « *deux* fois un » que de la signification de « deux fois *un* ». Car compter jusqu'à deux demande préalablement de parvenir à saisir l'unité. Autrement dit, déterminer où commence *deux* nécessite d'abord de savoir ce qui fait *un*. L'objet principal de ce chapitre est de circonscrire le problème du double en le désolidarisant d'un supposé rapport inévitable à la contradiction logique. En particulier, il s'agit d'éclaircir les liens qui articulent la notion d'identité à la notion d'unité. Il s'agira essentiellement

de considérer en quoi nous pouvons parler de l'identité et de l'unité d'une chose matérielle. Nous examinerons les questions de l'identité temporelle et des totalités composées en observant leur rapport problématique à l'identité numérique.

Au cours de ce premier chapitre, on demandera, avec insistance, « une ou deux choses ? » Or, précisons d'emblée qu'il faut distinguer ici deux types de questions. La première, épistémique, engage à déterminer ce en quoi nous pouvons savoir que nous avons affaire à une chose plutôt qu'à deux ou, au contraire, à deux choses plutôt qu'à une seule. La deuxième question, métaphysique, consiste à déterminer ce qui fait que la chose visée est *une* chose plutôt que *deux*, ceci indépendamment de toute intention à son égard.

Dans une première section, il s'agira de comparer de manière formelle l'identité logique à l'identité qualitative en faisant apparaître cette dernière comme l'identité numérique appliquée à des propriétés. Dans une seconde section, on abordera la question de l'individuation comme la recherche du critère d'identité d'une chose qui persiste dans le temps. Il s'agira de distinguer plusieurs questions liées respectivement à la notion d'unité et d'identité en ciblant des critères et des cribles respectivement métaphysiques et épistémiques. Les théories métaphysiques de l'identité des objets concrets, que j'appelle théories de l'identité objectuelle, achoppent, comme nous le verrons dans la section 3, sur différents paradoxes bien connus. Dans la section 4, nous passerons aux conceptions conventionnalistes l'identité.

1.1 Identité logique et identité qualitative

On commencera dans cette section par caractériser l'identité logique. L'enjeu est de la comparer à l'identité qualitative.

1.1.1 Identité numérique et identité qualitative

La tension entre l'identité et la différence, entre le singulier et le pluriel apparaît d'emblée dans les relations ambiguës qu'entretiennent la mêmeté et la ressemblance. Lorsque nous disons par exemple que Valérie et Christian conduisent la *même* voiture, de fait, nous pouvons affirmer deux choses différentes. Nous pouvons dire qu'ils partagent le même véhicule qu'ils conduisent de manière alternée. Mais nous pouvons aussi dire qu'ils conduisent chacun une voiture de la même marque et du même modèle. Dans le premier cas, il n'y a qu'une seule voiture (unicité), dans le second, deux unités, et les deux situations ont en commun l'usage du mot « même ». Aussi, les philosophes distinguent-ils deux sens de la « mêmeté ». D'une part *l'identité numérique* d'une chose, son unicité, ou identité logique et d'autre part *l'identité qualitative* entre différentes choses, relation fondée sur la ressemblance. L'ambiguïté qui réside dans l'assertion « la même voiture » peut donc être remédiée en précisant s'il s'agit de l'une ou l'autre identité. Ainsi, l'un des usages de « même » est référentiel, puisque parler de la même voiture signifie parler d'un certain individu à l'exception de tous les autres. Le deuxième usage est quant à lui descriptif : dire que les voitures sont les mêmes, c'est dire qu'elles entretiennent certaines ressemblances. Cela pourrait inciter à interpréter ainsi la signification de notre illustration introductive : dire d'une chambre qu'« elle est et n'est pas la même », pourrait donc signifier que cette chambre est numériquement identique, quoique devenue qualitativement différente.

Le problème du double est certainement à la jonction des relations d'identité et de ressemblance. Il devient alors nécessaire d'expliquer ce qui permet de différencier ces deux genres de relations. On trouve dans la littérature philosophique deux conceptions opposées qui articulent l'identité numérique à l'identité qualitative. La première conception est fondée sur un réalisme des universaux, la deuxième, sur un point de vue nominaliste.

Le courant réaliste se diversifie selon deux versions. La première version prétend que l'identité qualitative n'est finalement qu'un cas particulier d'identité numérique. L'identité de qualités n'est rien d'autre que la stricte identité numérique, mais portant sur des propriétés conçues comme des universaux. L'identité qualitative présumée entre deux pommes tiendrait de ce qu'elles instancieraient par exemple le même universel de rondeur. Cette conception, qui est bien représentée par la théorie des faisceaux de Bertrand Russell, consiste donc à rabattre la ressemblance sur l'identité numérique, moyennant la stipulation d'entités ayant la particularité d'être exemplifiées ou instanciées de manière synchronique à des emplacements différents. De telles entités seraient donc des doubles métaphysiques, des choses capables d'être présentes au même instant ici et là, et toute la difficulté consisterait alors à motiver la tolérance à l'égard de ce genre plutôt fantasque d'entités. Car comme le constate aussi John Heil :

> Si les universaux sont des entités platoniciennes, la relation d'instanciation est profondément mystérieuse. Si les universaux sont *in rebus*, il s'ensuit que l'universel est entièrement présent dans chacun de ses instances. Mais on a du mal à dire ce que cela pourrait signifier. (HEIL, 2011, p. 30)

Il existe encore une seconde théorie réaliste. Celle-ci préconise inversement de considérer que l'identité numérique est un cas particulier d'identité qualitative. La théorie de l'identité relative de Peter Geach admet qu'une chose x puisse s'avérer numériquement identique à y sous l'aspect de la sorte F, tout en étant numériquement différente de y sous l'aspect de la sorte G. Nous aurons l'occasion de revenir sur cette théorie, qui a le mérite de proposer une solution aux paradoxes de transitivité de l'identité, mais qui est rejetée par la plupart des philosophes parce qu'elle enfreint les caractéristiques fondamentales de la version standard de l'identité logique, qui est généralement considérée comme une relation absolue. En outre, si l'on considère que les sortes ne sont rien d'autre que des universaux, cette théorie est alors engagée envers les mêmes doubles métaphysiques que la version précédente.

L'on peut évidemment ne pas accepter l'existence des universaux. Autrement dit, l'on peut aussi considérer que tout ce qui existe est singulier, que tout ce qu'il y a est particulier. De ce point de vue, dire que la rondeur de la pomme a est identique à la rondeur de la pomme b ne serait qu'une manière impropre de parler, une façon lâche d'utiliser le langage de l'identité. Si l'on reconnaît des ressemblances entre des choses numériquement différentes, cela conduit alors à considérer que la relation de ressemblance est une relation qui n'est pas réductible à l'identité. Ainsi, la seconde conception, nominaliste, de l'articulation entre l'identité numérique et l'identité professe un anti-réalisme des universaux fondé sur l'idée que la ressemblance est une relation primitive [1].

[1] Soulignons qu'existent divers types de nominalismes, selon leur tolérance à l'égard des particuliers abstraits (tropes).

1.1.2 Caractériser l'identité comme une relation logique

Une définition problématique

L'identité numérique et l'identité qualitative partagent le fait d'être des relations d'ordre total, autrement dit, elles sont applicables à tout couple de choses. Ceci étant, elles se différencient sur différents points. Notons d'abord que, contrairement à l'identité qualitative, l'identité numérique est une relation absolue : soit une chose x est identique à y, soit elle ne l'est pas. Soit la voiture qui a été mal garée est la même que celle qui est retrouvée à la fourrière, soit c'est une voiture différente. Il n'y a pas de possibilité intermédiaire : la voiture ne saurait être « plus ou moins la même ». En outre, la relation d'identité numérique est isolante : tout ce qui existe est numériquement identique à lui-même et numériquement différent de toute autre chose. L'identité qualitative est, quant à elle, non-isolante (une chose peut être qualitativement identique à une chose numériquement différente) et relative (x peut plus ou moins ressembler à y, lui ressembler sur certains aspects et non sur d'autres).

Pour mieux saisir les liens entre l'identité numérique et la ressemblance, commençons par préciser au juste en quoi consiste l'identité logique dans sa version orthodoxe, autrement dit, absolue. La caractériser comme « la relation qu'une chose entretient avec elle-même à l'exclusion de toute autre » ne constitue pas une définition de l'identité, parce qu'« à l'exclusion de toute autre » est une périphrase de l'expression « avec toute chose qui ne lui est pas identique ». Nous pourrions alors tenter de la définir de manière numérique, par le dénombrement. Dire que x est identique à y, c'est dire que x et y ne sont qu'*une* chose. Néanmoins, il s'avère tout aussi vain de chercher à caractériser l'identité par l'unité, parce qu'affirmer qu'il y a une chose et deux ou plus, c'est justement considérer l'identité numérique de cette chose. Plus généralement, on remarquera qu'il semble de toute façon exclu de parvenir à une définition non circulaire de l'identité, parce que définir une chose, c'est justement indiquer en quoi il s'agit *de la même chose*. Plutôt que d'en rester au langage ordinaire, qui conduit systématiquement à des définitions circulaires, nous pouvons aussi envisager de circonscrire les frontières de l'identité en l'approchant d'une manière moins directe, en délimitant les principes forts qui la caractérisent dans nos pratiques langagières. Le formalisme sert d'appui considérable pour une telle caractérisation. En effet, non seulement il procure un abri contre les circonlocutions circulaires charriées par l'usage du langage naturel, tout en contribuant à mettre en avant les propriétés formelles de la relation que nous cherchons.

Dans « Fonction et Concept », article de 1891 qui fait suite à « Sens et dénotation » [2], Frege fait remarquer que l'égalité numérique n'est rien d'autre que l'identité logique. Distinguer entre identité et égalité reviendrait à vouloir « distinguer la violette odorante de la *violeta odorata* parce que ces noms sonnent différemment » (FREGE, 1971b, p. 82). Cela incite à formaliser l'identité comme une relation binaire. De manière générale, une relation nécessite la donnée de deux ensembles $\mathcal{E}$ et $\mathcal{F}$. Pour tout $x \in \mathcal{E}$ et tout $y \in \mathcal{F}$, on définit ensuite la notion de couple ordonné (x, y) comme l'ensemble $\{\{x\}, \{x, y\}\}$, qui est une partie du produit cartésien $\mathcal{E} \times \mathcal{F}$. Le fait de former de tels couples ordonnés permet alors de définir la notion de relation de manière extensionnelle : on dit que R est une relation entre $\mathcal{E}$ et $\mathcal{F}$ si et seulement si R est une partie du produit cartésien $\mathcal{E} \times \mathcal{F}$. Pour tout couple (x, y) appartenant à R, on dit que x est en relation R avec y et on

[2] Précisons que si la date de publication de « Fonction et Concept » est antérieure à celle « Sens et Dénotation », le deuxième article fut néanmoins écrit par Frege avant le premier.

note Rxy (notation préfixe) ou xRy (notation infixe). A partir de cette notion générale de relation, nous pouvons alors caractériser celle que nous appelons habituellement « identité », généralement notée « = ».

Tout autant que le langage ordinaire, le formalisme semble cependant échouer à parvenir directement à une définition logique de la notion standard d'identité, suivant laquelle tout objet est identique à lui-même et différent de tout autre. Plus précisément, considérons une structure $\mathfrak{U}$ pour un langage de premier ordre $\mathcal{L}$ contenant le signe « = ». Nous dirons que ce signe s'interprète de manière standard si la relation qu'il dénote entre x et y est vraie si et seulement si x est le même élément que y. Bien sûr, ce signe peut aussi s'interpréter de manière très différente. Mais existe-t-il, dans un langage $\mathcal{L}$ muni de l'égalité, un principe logique répondant à des modèles qui répondraient exclusivement à l'identité standard ? Comme le souligne Wilfrid Hodges, il n'en est rien :

> Let L be a langage containing the symbol '='. It would be pleasant if we could find a theory Δ in L whose models are exactly the L-structure with standard identity. Alas, there is no such theory. For every L-structure $\mathfrak{U}$ with standard identity there is an L-structure $\mathfrak{B}$ which is a modele of the same sentences of L as $\mathfrak{U}$ but doesn't have standard identity. (HODGES, 2004, p. 64)

Pour tout langage contenant le symbole d'identité, il y en a toujours un autre, contenant les mêmes expressions, mais dans lequel on peut déterminer des relations « = » qui ne s'interprètent pas en termes d'identité standard. Ainsi, l'égalité ne peut pas être directement définie de manière formelle par un principe logique exclusif.

Il nous reste cependant une solution de secours. Nous pourrions aussi demander s'il existe une théorie qui serait vraie dans toute L-structure contenant l'identité standard, et qui impliquerait logiquement toute affirmation vraie dans n'importe quelle de ces L-structures. Une telle théorie existe, elle est donnée par deux principes, la réflexivité et l'indiscernabilité des identiques.

Réflexivité et principe d'indiscernabilité des identiques

Une caractérisation immédiate de la relation d'identité est le principe de réflexivité :

$$\textbf{[Réflexivité de l'identité]} \quad \forall x \, (x = x)$$

Cependant, ce principe ne suffit pas pour isoler la relation qui nous intéresse. Il faut encore y ajouter le principe d'*indiscernabilité des identiques* : si deux choses sont identiques, alors toute propriété prédicable à l'une est également prédicable à l'autre et réciproquement. Utiliser ce principe nécessite ainsi une quantification sur les objets et sur les propriétés. Afin de l'exprimer formellement, nous devons donc passer d'une logique du premier ordre à une logique du second ordre :

$$\textbf{[Indiscernabilité des identiques (PIId)]} \quad \forall x \, \forall y \left((x = y) \rightarrow \forall \mathrm{P} \, (\mathrm{P}x \leftrightarrow \mathrm{P}y) \right)$$

Ce principe ne doit pas être confondu avec sa converse, dénommée principe d'identité des indiscernables, qui peut être, quant à elle, formalisée ainsi :

$$\textbf{[identité des indiscernables (PIdI)]}\ \forall x\,\forall y\left(\forall \mathrm{P}\,(\mathrm{P}x \leftrightarrow \mathrm{P}y) \rightarrow (x = y)\right)$$

Celui-ci peut s'interpréter ainsi : si toute propriété prédicable à un objet est prédicable à un autre et réciproquement, alors ces deux objets sont identiques. Ce principe peut aussi s'exprimer sous sa forme contraposée, appelée principe de dissemblance du divers : si deux objets sont distincts, alors il existe au moins une propriété qu'ils ne partagent pas.

Nous aurons amplement l'occasion de discuter de la validité, bien plus polémique, du principe d'identité des indiscernables, mais pour l'instant, restons-en à l'indiscernabilité des identiques, qui est une généralisation d'un principe appelé « principe de substituabilité des identiques *salva veritate* », principe utilisé par Leibniz pour le calcul propositionnel et qui explique la fonction de l'égalité. Il énonce que si x et y sont deux termes qui co-réfèrent au même objet, alors, dans tout énoncé, l'un des deux termes peut être substitué à l'autre sans modifier la signification de la proposition :

$$\textbf{[SI]}\ \forall x\,\forall y\,\forall \mathrm{P}\left((x = y) \wedge \mathrm{P}x \rightarrow \mathrm{P}y\right)$$

A ce principe logique correspond une règle d'inférence particulièrement utile.

Selon W.V.O. Quine, l'indiscernabilité des identiques est complètement équivalente au principe de substitution des identiques *salva veritate*. Il n'est donc qu'un principe linguistique, ne s'appliquant qu'aux seuls signes et non aux choses. De fait, Quine reproche à Leibniz et à ses successeurs de confondre signe et objet dans la formulation du principe. C'est cette raison qui amène Quine à réinterpréter le principe de substitution des identiques *salva veritate* ainsi : « étant donné un énoncé d'identité vrai, l'un de ses deux termes peut être substitué à l'autre dans tout autre énoncé vrai et le résultat sera vrai » (W. v. O. QUINE, 1980, chapitre VIII).

Soulignons encore que la contraposée de ce principe d'indiscernabilité des identiques est d'une grande utilité pour discriminer des choses. En effet, elle affirme que si un prédicat peut se dire d'une chose mais pas d'une autre, alors ces choses sont numériquement différentes :

$$\textbf{[Distinction des discernables]}\ \forall x\,\forall y\left(\exists \mathrm{P}\,(\mathrm{P}x \wedge \neg \mathrm{P}y) \vee (\neg \mathrm{P}x \wedge \mathrm{P}y) \rightarrow (x \neq y)\right)$$

Principes dérivés : symétrie et transitivité

Selon la conception classique, ces deux principes suffisent à circonscrire l'identité logique. En effet, ils permettent de dériver toutes les autres caractéristiques particulièrement représentatives de la relation d'identité, notamment la symétrie et la transitivité. La symétrie signifie que si x est identique à y, alors y est également identique à x :

$$\textbf{[Symétrie]}\ \forall x\,\forall y\Big((x = y) \rightarrow (y = x)\Big)$$

Prouvons ce principe. Supposons que x et y soient tels que $x = y$. Par le principe d'indiscernabilité des identiques, on a alors :

$$x = y \rightarrow \forall \mathrm{P}\ (\mathrm{P}x \leftrightarrow \mathrm{P}y)$$

En prenant pour Pa l'expression $a = x$, il vient que :

$$(x = y) \rightarrow \Big((x = x) \leftrightarrow (y = x)\Big)$$

Par hypothèse que $x = y$ et par *modus ponens*, on peut en déduire que :

$$(x = x) \leftrightarrow (y = x)$$

Par réflexivité de l'identité et par *modus ponens*, il vient alors que :

$$y = x$$

ce qui confirme bien que la relation est symétrique.

On peut également dériver de la réflexivité et de l'indiscernabilité des identiques le principe de transitivité qui affirme que si x est identique à y et que si y est identique à z, alors x est également identique à z.

[Transitivité] $\forall x\ \forall y\ \forall z\ \Big((x = y) \wedge (y = z) \rightarrow (x = z)\Big)$

Prouvons ce nouveau principe. Supposons que $(x = y) \wedge (y = z)$. Par le principe d'indiscernabilité des identiques, on a :

$$(x = y) \rightarrow \forall \mathrm{P}\ (\mathrm{P}x \leftrightarrow \mathrm{P}y)$$

En prenant pour Pa l'expression $(x = a) \wedge (a = z)$, il vient que :

$$(x = y) \rightarrow \Big((x = x) \wedge (x = z) \leftrightarrow (x = y) \wedge (y = z)\Big)$$

Par hypothèse que $(x = y)$ et par *modus ponens*, on peut en déduire que :

$$(x = x) \wedge (x = z) \leftrightarrow (x = y) \wedge (y = z)$$

Par hypothèse que $(x = y) \wedge (y = z)$ et par *modus ponens*, on peut en déduire que :

$$(x = x) \wedge (x = z)$$

Donc finalement que :

$$x = z$$

Ce qui prouve bien la propriété de transitivité.

Unicité extensionnelle de la relation d'identité

Ainsi, puisque l'identité satisfait aux axiomes de réflexivité, de symétrie et de transitivité, il s'agit d'une relation d'équivalence. Il reste à savoir si cette relation est unique, autrement dit à connaître si une telle axiomatique ne définit pas tout un ensemble de relations qui ne seraient pas équivalentes d'un point de vue extensionnel. Supposons par exemple que la réflexivité et l'indiscernabilité des identiques satisfassent non seulement la relation d'identité classique mais également une deuxième relation, mettons la relation d'« udentité » que l'on notera ici « $\equiv$ ». Nous allons montrer dans un premier temps que $x = y$ implique $x \equiv y$, puis réciproquement que $x \equiv y$ implique que $x = y$.

Supposons que x et y soient tels que $x = y$. Par le principe d'indiscernabilité des identiques appliqué à la relation d'identité, on a donc :

$$(x = y) \rightarrow \forall \mathrm{P}\ (\mathrm{P}x \leftrightarrow \mathrm{P}y)$$

Considérons alors, pour tout objet s, la propriété d'être udentique à x, notée $\mathrm{U}_x := (s \equiv x)$. En substituant U_x à P dans l'expression précédente, on obtient que :

$$(x = y) \rightarrow (x \equiv x \leftrightarrow y \equiv x)$$

Par l'hypothèse que $(x = y)$ et par *modus ponens*, il vient alors :

$$(x \equiv x \leftrightarrow y \equiv x)$$

Puisque, par hypothèse, la relation d'udentité est réflexive, on a bien sûr :

$$x \equiv x$$

De ceci, de l'équivalence précédente et de *modus ponens*, il vient que :

$$y \equiv x$$

Or, par symétrie d'udentité, on peut en dériver que :

$$x \equiv y$$

On peut donc conclure que :

$$(x = y) \rightarrow (x \equiv y)$$

Passons maintenant à la réciproque. Supposons que x et y soient tels que $x \equiv y$. Par le principe de l'indiscernabilité des identiques appliqué à la relation d'udentité, on a que :

$$(x \equiv y) \rightarrow \forall \mathrm{P}\ (\mathrm{P}x \leftrightarrow \mathrm{P}y)$$

En substituant à P la propriété qui à tout s associe la propriété d'être identique à x, autrement dit $\mathrm{I}_x := (s = x)$, on obtient que :

$$(x \equiv y) \rightarrow ((x = x) \leftrightarrow (y = x))$$

Par l'hypothèse que $x \equiv y$ et par *modus ponens*, il vient alors :

$$(x = x) \leftrightarrow (y = x)$$

La relation d'identité étant réflexive, on a bien :

$$x = x$$

Par ceci, par l'équivalence précédente et par *modus ponens*, il vient alors que :

$$y = x$$

Par symétrie de l'identité, on peut en déduire que :

$$x = y$$

On en conclut ainsi que :

$$(x \equiv y) \rightarrow (x = y)$$

Finalement on a bien :

$$(x = y) \leftrightarrow (x \equiv y)$$

Ainsi, la relation que nous avons dénommée « udentité » est extensionnellement équivalente à celle d'identité. Les classes d'objets que départage l'udentité seront une-à-une correspondantes aux classes d'objets que départage l'identité. Aussi, les deux axiomes de réflexivité et d'indiscernabilité des identiques définissent bien, de manière extensionnelle, une unique relation.

Pour achever ce tour des propriétés formelles de l'identité, nous pouvons ajouter une dernière caractéristique : le principe de nécessité de l'identité affirme que si deux choses sont identiques, alors elles le sont nécessairement. Nous aurons l'occasion de revenir dans la suite sur les conséquences de cette dernière propriété de l'identité.

1.1.3 Comparaison entre identité logique et ressemblance

Nous pouvons maintenant mieux percevoir les différences entre l'identité qualitative et l'identité numérique en comparant leurs principales caractéristiques formelles. Notons $\mathrm{R_F}xy$ pour la relation « x ressemble à y sous l'aspect F ». Il semble immédiat que nous ayons affaire à une relation, et

que celle-ci soit relative à l'aspect F sous lequel nous considérons les choses comparées. Il est peu douteux que cette relation soit réflexive : $\forall x$ $\mathrm{R_F}xx$. Il est également assez peu contestable qu'il s'agisse d'une relation symétrique. Si x ressemble à y alors certainement, y ressemble à x : $\forall x$ $\forall y$ $\mathrm{R_F}xy \rightarrow \mathrm{R_F}yx$. En revanche, la transitivité semble moins assurée : il se pourrait qu'un x ressemble à y sous un certain aspect F, que y ressemble à z sous le même aspect sans que x ressemble à z sous cet aspect. Il ne s'agit donc pas d'une relation d'équivalence. En outre, elle n'obéit pas au principe d'indiscernabilité, puisque deux choses qui se ressemblent sous un aspect F ne sont pas nécessairement ressemblantes sous d'autres aspects. Par ailleurs, les relations de ressemblance sont diverses. Il y a autant de relations de ressemblance que d'aspects comparatifs Fs. Enfin, il est certain que les relations de ressemblance sont, en général, contingentes plutôt que nécessaires : le vin d'une bouteille peut ressembler à du vinaigre par l'acidité, mais si les conditions de stockage de la bouteille avaient été plus appropriées à sa conservation, cela n'aurait peut-être pas été le cas.

Résumons les différences entre l'identité numérique et la ressemblance par le tableau suivant :

	Relation	Ordre total	Réflexivité	Symétrie	Transitive	Équivalence	Indiscernabilité	Unicité	Absolue	Isolante	Nécessaire
Identité numérique	O	O	O	O	O	O	O	O	O	O	O
Identité qualitative (ressemblance)	O	O	O	O	N	N	N	N	N	N	N

Les différences mentionnées dans ce tableau entre l'identité logique et la ressemblance incitent à considérer qu'il s'agit bien de deux relations distinctes. Dans le chapitre suivant, nous allons considérer comment s'articulent ces deux relations quand on considère le cas particulier d'une ressemblance qui ne serait pas seulement partielle, mais exacte, autrement dit, dans le cas de choses indiscernables. Mais avant cela, il nous faut encore poursuivre notre enquête générale sur la relation d'identité numérique. En particulier, nous devons maintenant considérer si cette relation est ou n'est pas purement formelle.

1.2 La théorie objectuelle de l'identité

L'objet de cette section est de distinguer une conception métaphysique de l'identité (identité objectuelle) d'une caractérisation épistémique. Nous retraçons de manière historique cette question à partir de l'analyse de Frege. L'enjeu est de faire saillir les différentes notions qui sous-tendent les processus d'identification selon deux catégories, métaphysique et épistémique, subdivisées en quatre questions (de l'unité, de l'unicité, de la persistance, de l'identité spécifique).

La caractérisation standard de l'identité logique confirme certaines de nos intuitions concernant la mêmeté. Cette relation obéit aux mêmes principes que ceux qui guident nos intuitions d'identification : indiscernabilité des identiques, réflexivité, symétrie et transitivité. En revanche, elle ne fournit pas de règle pour déterminer la façon dont elle peut s'appliquer ou non aux choses réelles. La question est donc de déterminer quel genre de pont relie le langage de l'identité au monde.

1.2.1 La question du rapport au langage

La sémantique de Frege

Gottlob Frege a souligné qu'il est difficile de déterminer la nature exacte de l'identité :

> La notion d'égalité propose à la réflexion quelques questions irrémissibles, auxquelles il n'est pas aisé de répondre. Est-ce une relation ? Une relation entre des objets, ou entre des noms ou signes d'objets ? (FREGE, 1971b, p. 102)

On peut se demander si la relation d'identité qu'entretient *Phosphorus* à elle-même est comparable à d'autres sortes de relations. Est-ce une relation dans le même sens qu'on parle de relation pour dire que *Phosphorus* est à une certaine distance du Soleil ? L'identité numérique n'est-elle qu'une forme logique reliant des termes, ou s'agit-il d'un principe métaphysique connectant les choses à elles-mêmes ? Énonce-t-elle des faits à propos du monde ou s'agit-il seulement de descriptions concernant nos manières spécifiques de concevoir les choses ?

Appelons d'une part « relation objectuelle » une relation ayant cours entre des objets, et appelons d'autre part « relation conceptuelle » une propriété ou une relation articulant seulement des concepts. La question qui nous intéresse ici est de savoir si l'identité est une relation objectuelle ou conceptuelle.

Dans les *Begriffsschrifte* (1879) Frege avait d'abord émis l'hypothèse que l'identité ne s'applique qu'aux seuls signes, autrement dit, qu'elle est entièrement réductible au principe de substitution des identiques *salva veritate*, principe qui suppose qu'une expression référentielle contribue seulement par l'objet qu'elle désigne aux conditions de vérité d'un énoncé qui la contient. Cependant, dans *Sens et dénotation* (1892), Frege a révisé cette hypothèse en raison de la différence épistémique qui sépare les expressions suivantes :

(H=H) Hesperus est identique à Hesperus.

(H=P) Hesperus est identique à Phosphorus.

Notons que si la même relation d'identité apparaît dans les deux expressions, leur contenu informatif varie pourtant de manière très sensible. Car la première concerne une vérité purement analytique là où la seconde s'avère être une découverte astronomique notable.

Frege constate que si l'identité ne portait que sur les objets, la différence épistémique relevée serait tout simplement inexplicable, parce que l'égalité ne signifierait rien d'autre que le fait, pour une chose, d'être identique à elle-même. Nous ne pourrions donc pas *apprendre* que Phosphorus est la même planète que Hesperus. D'autre part, si l'identité ne portait au contraire que sur les signes, il s'en suivrait que la seconde égalité ne concernerait que notre manière de désigner, mais pas la planète Vénus elle-même. Les signes nouant avec les choses des liens purement arbitraires, là encore, il n'y aurait aucun moyen d'expliquer l'apport épistémique constaté. Frege en tire la conséquence que l'identité porte à la fois sur les choses et sur les signes qui expriment ces choses.

Mais cela n'explique pas encore complètement la différence de signification qui ressort des deux usages précédents de l'identité. Selon Frege, sa bivalence découle d'une double composante

sémantique : la signification d'une expression simple ou complexe consiste à la fois en sa référence (*Bedeutung*), autrement à l'objet réel qu'elle vise, et à son mode de présentation, que Frege désigne sous le nom de « sens » (*Sinn*). À une expression donnée, comme par exemple « Hesperus », est associé un sens déterminé, ici, « la première étoile visible à la tombée de la nuit », sens qui correspond lui-même à une référence unique, à savoir, la planète Vénus. En revanche, à une référence donnée, comme Vénus, peuvent correspondre plusieurs sens, comme « l'étoile du soir » ou comme « l'étoile du matin ».

Du point de vue de cette distinction sémantique, Frege souligne que deux expressions ne sont synonymes qu'à la condition qu'elles partagent à la fois leur dénotation et leur sens. En revanche, des expressions comme « Hesperus » et « Phosphorus » sont certes dotées de la même référence, mais puisqu'elles n'ont pas le même sens, elles ne sont pas synonymes. Quant au cas où deux expressions partageraient le même sens mais pas la même dénotation, Frege en rejette complètement la possibilité. Chez Frege, la référence semble déjà plus ou moins indiquée par le mode de présentation [3]. En effet, le sens étant défini comme le mode d'accès à la référence, l'identité de sens implique nécessairement la co-référence.

La double composante sémantique des expressions, dénotation et sens, explique ainsi les aspects contradictoires des contenus épistémiques des assertions (H=H) et (H=P) :

> Si $a = b$ la dénotation de b est bien la même que celle de a, et la valeur de vérité de $a = b$ est aussi la même que celle de $a = a$. Toutefois, le sens de b peut être différent du sens de a et par là la pensée exprimée dans $a = b$ peut être différente de celle exprimée dans $a = a$. Dans ce cas, les deux propositions n'ont plus la même valeur pour la connaissance. Si, comme précédemment, nous entendons par « jugement » le progrès qui va d'une pensée à sa valeur de vérité, nous dirons encore que les jugements sont différents. (FREGE, 1971b, p. 127)

La partition entre dénotation et sens répond aussi au problème des contextes dans lesquels le principe de substituabilité des identiques *salva veritate* ne peut être appliqué. Frege explique ceci en introduisant la notion de contexte indirect, contexte dans lequel la signification du terme employé n'a pas pour dénotation sa référence habituelle, mais son sens. L'échec de la substituabilité tient donc de cette différence dans la manière de dénoter selon qu'on a affaire à un contexte direct ou non. Il n'y a donc pas véritablement d'entorse au principe de substituabilité des identiques.

Reste à savoir comment concevoir une telle relation entre les signes et les choses. La stratégie de Frege consiste à dupliquer le monde factuel par un autre monde, tout aussi objectif, le monde des sens. Est donc stipulé un « troisième domaine » entre le monde objectif matériel et le domaine purement subjectif des représentations :

> Les pensées ne sont ni des choses du monde extérieur, ni des représentations. Il faut admettre un troisième domaine. Ce qu'il enferme s'accorde avec les représentations en ce qu'il ne peut pas être perçu par les sens, mais aussi avec les choses en ce qu'il n'a pas besoin d'un porteur dont il serait le contenu de conscience. Telle est par exemple la pensée que nous exprimons

[3] C'est d'ailleurs ce qui pousse Manuel Rebuschi à affirmer que la théorie de la signification de Frege est « une vraie-fausse sémantique dualiste : le niveau du Sens détermine, i.e. contient implicitement celui de la Dénotation » (REBUSCHI, 2008).

> dans le théorème de Pythagore, vraie intemporellement, vraie indépendamment du fait que quelqu'un la tienne pour vraie ou non. (FREGE, 1969b, p. 184)

Dans ce qui suit, nous voudrions souscrire en partie à l'esprit de la sémantique frégéenne, en conservant l'idée qu'une dénotation est toujours dédoublée d'un mode de présentation, tout en évitant l'inflation ontologique du *Sinn* et tout en évitant également de rendre la dénotation tributaire du mode d'accès. Car de fait, il semble bien possible de concevoir sous un même mode de présentation des choses qui seraient néanmoins numériquement différentes. Nous pouvons avoir des pensées indiscernables concernant des individus néanmoins différents.

La notion d'objet comme particulier

Si nous voulons savoir comment départager le fait d'être en présence d'un objet ou de deux objets, il dévient nécessaire de préciser plus clairement ce que nous entendons par le terme « objet ». Comment caractériser la notion recouverte par ce terme ? Remarquons que si le mot « objet » peut s'employer dans la position d'un prédicat, comme dans la phrase « cette table est un objet », il ne semble pourtant rien dire de la sorte de chose qu'est une table. « Objet » n'est pas un prédicat sortal, mais catégorial, comme le sont aussi les termes « chose » et « particulier » que je prendrai ici pour synonymes.

Frege a donné une caractérisation, que nous suivrons ici, en distinguant les notions d'objet et de propriété par une opposition formelle entre terme singulier et terme général (FREGE, 1971a). Un terme désigne un objet s'il est saturé, autrement dit, s'il n'a besoin de rien d'autre pour signifier. A la différence des premiers, les termes prédicatifs sont insaturés, c'est-à-dire qu'ils nécessitent une connexion à un ou plusieurs objets pour parvenir à signifier quelque chose. Il y a donc une frontière linguistique étanche entre objets et propriétés. Il arrive toutefois, précise Frege, que nous fassions référence à un concept, par exemple lorsque nous préfixons le prédicat considéré par l'expression « le concept de... », comme dans l'expression « le concept d'homme ». Dans cette situation, la propriété est considérée elle-même comme un objet, autrement dit, on réifie cette propriété. De cette manière, Frege établit une correspondance entre le couple objet/propriété et le couple référence/prédication.

Bertrand Russell note que la distinction de Frege est sans doute la plus pertinente pour démarquer de manière exclusivement logique les particuliers des universaux :

> If predication is an ultimate relation, the best definition of particulars is that they are entities which can only be subjects of predicates or terms of relation *i.e.*, that they are (in the logical sense) substances. This definition is preferable to one introducing space or time, because space and time are accidental characteristics of the world with which we happen to be acquainted and therefore are destitute of the necessary universality belonging to purely logical categories. (RUSSELL, 1912, p. 23)

Russell distingue encore deux types de prédicats : les relations, qui sont dyadiques, triadiques, etc. et les qualités. Les entités peuvent donc être rangées dans deux classes. La première contient les particuliers vus comme sujets de prédicats ou comme termes de relations, particuliers qui, s'ils sont de notre monde, existent dans le temps en occupant, à un moment donné, une unique position.

La seconde contient les universaux, qui peuvent remplir le rôle de prédicats ou de relations, qui n'existent pas dans le temps et qui n'ont aucune relation à un quelconque emplacement :

> The ground for regarding such a division as unavoidable is the self-evident fact that certain spatial relations imply diversity of their terms, together with the self-evident fact that it is logically possible for entities having such spatial relations to be wholly indistinguishable as to predicates. (RUSSELL, 1912, p. 24)

Ainsi, lorsque nous parlons d'objet, nous véhiculons par ce terme l'idée de particulier, qui semble mêlée de manière inextricable à la notion d'unité. Il nous faut donc voir maintenant comment s'articule l'identité à l'unité.

1.2.2 Unité et identité

Les relations entre l'identité et l'unité

Dans les conditions habituelles, nos yeux nous procurent deux images de la réalité qui se superposent. L'ivrogne a cependant la réputation de voir double, son cerveau perdant provisoirement sa faculté de superposer ces deux images. Mais Clément Rosset fait remarquer que c'est sans doute plutôt l'homme sobre qui voit double, quand l'ivrogne qui observe une fleur est :

> [...] hébété par la présence sous ses yeux d'une chose singulière et unique [...] Une chose toute simple, c'est-à-dire saisie comme singularité stupéfiante, comme émergence insolite dans le champ de l'existence. (ROSSET, 1978, p. 41)

La « singularité stupéfiante » de toute chose concrète relève de son *unicité* ontologique, autrement dit du fait de n'être identique à rien d'autre, de son identité métaphysique.

Mais quelles relations entretiennent l'unicité et l'unité ? Dans les *Fondements de l'Arithmétique*, Frege remarque une certaine ambiguïté entre des notions distinctes recouvertes sous celle, plus générale, d'unité. Une certaine contradiction semble en effet s'immiscer dans les rapports entre l'idée d'unités, dont le pluriel marque la négation de l'identité, et le nombre un, qui est par nature, un terme singulier. Comme l'indique Frege, ce rapport discordant entre unités et unicité apparaît déjà dans les septièmes *Éléments* d'Euclide, à travers les différentes définitions que l'auteur donne du nombre. La première définition d'Euclide stipule en effet que l'unité est ce suivant quoi toute chose existante est *une*, mais la seconde définition identifie le nombre à une multitude composée d'unités.

Des unités peuvent-elles s'avérer différentes, demande Frege, ou y a-t-il au contraire unicité des unités ? Si des objets, numériquement différents, peuvent constituer les unités du décompte « $1 + 1 = 2$ », faut-il en inférer que de ces unités sont, entre elles, différentes ? Dans ce cas, nous devrions sans doute écrire plus rigoureusement « $1' + 1''$ » plutôt que « $1 + 1$ ». Mais comme le fait remarquer Frege, nous perdrions alors toutes les égalités numériques (FREGE, 1969a, p. 166-167). Il faut donc, prévient Frege, dissocier l'unité numérique de l'identité d'individuation. Il ne faut donc pas confondre la sommation arithmétique « $1 + 1 = 2$ » avec la conjonction « 1 et 1 » qui fait encore 1 :

> Si l'on emploie « et » comme dans « Bunsen et Kirchhof », 1 et 1 et 1 ne fait pas 3, mais 1, de même que l'or et de l'or et de l'or ne fait rien d'autre que de l'or. (FREGE, 1969a, p. 167)

On dégagera ainsi deux notions connexes : celle d'« unité », qui est corrélative à un dénombrement, et celle d'unicité, qui relève d'une identité métaphysique. En insistant sur cette différence, Frege introduit une désolidarisation entre la conception numérale de l'unité et une notion d'unicité qui ressortit à l'individuation métaphysique. Cette deuxième notion est désignée par Frege du nom de « critère d'identité » [4].

Sous cette distinction, observons que pour dénombrer, il faut saisir *l'unité* d'une chose tout en abandonnant ce qui en fait l'*unicité*, de sorte que n'importe quel objet puisse servir d'étalon possible pour la mesure de quantité de choses du même type. Partant, l'appréhension de l'unité et l'appréhension de l'unicité semblent convoquer deux attitudes contradictoires. En effet, donner à un objet le rôle d'*unité* d'un décompte consiste en une sorte de dénégation de son unicité. Cette manœuvre est pourtant nécessaire afin de promouvoir la chose considérée au rôle de représentant unitaire de toute une classe. En gommant sa singularité, c'est-à-dire en négligeant les traits saillants qui la différencient des autres choses, on ramène ainsi la chose visée à la seule condition d'être *une* chose parmi *plusieurs* choses de la même sorte.

Un principe d'individuation doit alors combiner à la fois un critère d'unité et un critère d'identité, comme le fait notamment remarquer Filipe Drapeau Contim écrit dans l'entrée « identité » de *L'Encyclopédie Philosophique* que :

> Le critère d'identité d'une sorte F de choses nous dit à quelles conditions un F est le même ou pas qu'un F existant au même moment (identité synchronique) ou à un moment différent (identité diachronique), mais il ne nous dit pas comment discriminer ou individuer en pensée un F en premier lieu ; on présuppose qu'on a déjà effectué la tâche de segmentation par laquelle on sépare un F des autres objets. Le rôle d'un critère d'unité est précisément de procéder à cette délimitation, en décrivant le genre d'unité que des régions matérielles doivent entretenir entre elles pour faire partie d'un même F, l'absence d'un tel lien marquant la limite entre ce F et les objets environnants. (DRAPEAU VIEIRA CONTIM, 2016)

De son côté, Jonathan Lowe constate lui aussi que les notions d'unité et d'unicité sont souvent embrouillées dans nos conceptions des individus, mais il considère toutefois que dans certains cas, elles peuvent être entièrement désintriquées : « [. . .] it is questionable to assume that identity and countability are mutually inseparable - that items can constitute a countable plurality if and only if they are determinatly distinct from one another » (E. J. LOWE, 2011, p. 77). Ainsi, à côté des individus qui répondent au double critère d'individuation, il y a des choses qui ne répondent à aucun critère d'unité, mais qui possèdent cependant un critère d'identité, comme les termes de masse, et il y a aussi des choses (que Lowe appelle « dividus ») qui, inversement, répondent à un critère d'unité mais pas à un critère d'identité, comme par exemple les particules quantiques [5]. Certains philosophes considèrent ainsi que, tout comme l'unicité, l'appréhension de l'unité dépend elle aussi d'un principe métaphysique.

[4] Soulignons que le terme « critère » est ici trompeur en ce qu'il pourrait faire croire à une notion purement épistémique alors qu'il relève en fait pour Frege d'une individuation métaphysique.

[5] Nous aurons l'occasion de revenir de manière plus approfondie sur ces particuliers dans le chapitre suivant.

La question est de savoir si l'unité relève ou non de la métaphysique. De fait, on peut concevoir la nature ontologique de l'unité selon deux points de vue différents. Le premier affirme que l'unité existe indépendamment de l'esprit humain, soit de manière transcendante par rapport aux choses sensibles (réalisme transcendantal), soit de manière immanente (réalisme immanentiste). La seconde option, conventionnaliste, affirme au contraire que l'unité résulte de l'activité de l'esprit humain, qu'elle n'existe nulle part ailleurs que dans nos têtes, voire qu'elle n'est qu'une simple convention de langage. Nous allons considérer dans ce qui suit ces deux perspectives.

La conception immanentiste de l'unité

Notre question est de savoir, pour une occasion donnée, ce qui nous incite à compter *une* chose plutôt que plusieurs. L'unité est-elle un fait métaphysique inscrit dans la nature des choses ? Autrement dit, s'agit-il d'une réalité du monde dont la vérité s'étend au delà de nos concepts ? Ou est-ce au contraire une certaine disposition de notre entendement, voire de notre biologie ? À moins qu'il ne s'agisse plutôt d'une simple manière de voir le monde, d'une psychologie sociale dont la convention aurait pu s'avérer bien différente ? La question est de savoir si nous sommes libres de considérer qu'on a affaire tantôt à une chose ou tantôt à deux ou si le monde est au contraire véritablement découpé en unités ontologiques. Et si tel est le cas, sommes-nous capables de saisir de telles unités métaphysiques ? Correspondent-elles à nos propres segmentations conceptuelles ?

Comme nous l'avons précisé plus haut, le réalisme immanentiste soutient la thèse que l'unité est une propriété tirée des choses elles-mêmes ou tirée d'un agrégat de choses. Si je peux dire qu'il y a un livre et non deux sur la table, c'est du fait que je tire du livre en question la propriété d'être *un* (livre) [6].

Une telle explication est bien sûr favorisée par les théories empiristes, qui, en faisant découler la connaissance de l'expérience, ont une évidente inclinaison à affirmer que l'unité provient de l'expérience que nous faisons des choses. Si nous parvenons à connaître l'idée d'unité, c'est que celle-ci doit provenir de certaines expériences et donc que l'unité est bien une propriété des choses. C'est notamment ce que soutient John Locke qui considère de manière générale que les nombres sont des qualités primaires des choses, à l'instar de la taille, de la forme, de la position et du mouvement :

> Les qualités qui sont dans les corps correctement envisagés sont de trois sortes : d'abord, la taille, la forme, le nombre, la situation et le mouvement ou le repos de leurs éléments solides [...] Je les appelle *qualités primaires*. (LOCKE, 1998c, Livre II, chapitre VIII, §23)

Comme on le voit, pour Locke, les nombres sont mis sur le même plan que les propriétés physiques. Mais tous les nombres ne posent pas avec une égale insistance la question de leur présence dans les choses. L'unité semble avoir une certaine préséance parce qu'elle donne naissance à une idée plus simple - car non composée, et surtout, universelle. Locke souligne cet aspect de l'unité, idée qui résulte de toute chose, qui accompagne chacune des idées de l'entendement et chacune des pensées de l'esprit. Du fait de son universalité, l'unité occupe donc une place particulière parmi les idées. Dans l'*Essai*, Locke écrit ainsi que :

[6] Quoique nous centrions pour l'instant notre attention sur la seule question de l'unité, nous verrons plus loin qu'il faut la relier à la sorte de chose considérée, à savoir, ici, la propriété d'être (un) *livre*.

> Parmi toutes nos idées, aucune n'est suggérée par plus de voies à l'esprit et en même temps aucune n'est plus simple, que celle *d'unité* ou d'un. Il n'y a pas l'ombre d'une différence ou d'une composition en elle ; tout objet dont s'occupent les sens, toute idée de l'entendement, toute pensée de l'esprit l'impliquent. C'est donc l'idée la plus intime à nos pensées, et c'est aussi, par sa concordance avec toutes les autres choses la plus universelle que nous ayons [...]. (LOCKE, 2001, Livre II, chapitre XVI, §1)

L'unité possède ainsi une certaine primauté dans l'ordre des idées, et donc dans l'ordre des nombres, ces derniers résultant de la répétition de l'unité dans l'esprit. En cela, et à la différence de l'idée d'unité, qui est simple, les nombres sont quant à eux des idées complexes : « ainsi, en additionnant un et un, on obtient l'idée complexe d'un couple, en mettant en ensemble douze unités, on obtient l'idée complexe d'une douzaine ; et de même pour une vingtaine, un million ou tout autre nombre » (LOCKE, 2001, Livre II, chapitre XVI, §2).

Locke souligne également le caractère bien démarqué des idées de nombre. Par comparaison avec elles, d'autres idées, quoique distinctes, peuvent s'avérer indiscriminables entre elles : « qui tentera de trouver en effet une différence de blancheur entre cette feuille et la nuance de blanc qui lui est immédiatement voisine ? Ou qui pourra former des idées distinctes de chacune des plus petites différences d'étendue ? » (LOCKE, 2001, l.II, XVI, §3). En comparaison, remarquons que la l'idée de différence entre des nombres est toujours limpide, que ces nombres soient proches ou éloignés. La distinction entre 2 et 3 est tout aussi claire que celle entre 2 et 10 millions. Si l'idée de nombre provient des choses sensibles, il s'agit d'une idée particulièrement claire, qui exclut le vague. Nous pouvons bien sûr nous tromper dans un décompte de choses, voire ne pas parvenir à les compter, mais une fois que nous sommes parvenus au compte, autrement dit aux idées adéquates afférentes aux nombres qui sont exemplifiées par ces choses, nous reconnaissons leur différenciation de manière absolument claire.

Il faut conclure que, avec la conception de l'unité de Locke, le double ne serait certainement qu'une sorte d'erreur, d'égarement, ou de confusion intellectuelle. Car l'unité résulte de tout objet et accompagne toutes nos pensées, elle est tout aussi absolue que limpide. Mais cette soit-disant limpidité peut sans doute être remise en question.

Les objections contre l'immanentisme

Certaines objections peuvent être adressées à la conception immanentiste de l'unité. Pour commencer, prenons pour illustration un passage de la bande dessinée d'Hergé intitulée *On a marché sur la Lune*, dans lequel les Dupondt entreprennent une excursion à pied sur la Lune. Leur promenade prenant la forme d'un cercle, les deux inspecteurs finissent par revenir sur leurs propres pas, qu'ils prennent pour ceux d'autrui, en déduisant ainsi une présence étrangère. Communiquant la cause de leur stupeur à l'équipage de la fusée, les Dupondt se voient rapidement gourmandés par le professeur Tournesol. Car ce dernier a bien compris de quoi il retourne réellement et les avertit de leur méprise : « - Je parie que vous êtes revenus sur vos pas, et que ce sont donc vos propres traces ». Après un petit moment d'hésitation, durant lequel les Dupondt se remémorent avoir déjà été dupés d'une telle manière à l'occasion d'une autre excursion, ils réfutent ensuite catégoriquement l'hypothèse de Tournesol : « - Absolument exclu ! s'exclament-ils, car il y a *deux* traces et nous sommes *seuls* ! » (HERGÉ, 1954).

Il paraît évidemment inepte de concéder aux inspecteurs d'Hergé *qu'à deux, ils sont seuls*. Mais on aurait tort de mépriser la cause du caractère spécieux de l'argument. L'allégation des Dupondt pourrait en effet procéder de l'argumentation suivante : « puisque chacun d'entre nous, pris isolément, possède la propriété d'être seul, autrement dit, d'être *un*, notre réunion possède également cette même propriété. En effet, le fait d'être plusieurs ne modifie pas les propriétés que porte chacun d'entre nous. Donc nous demeurons, chacun, un. Aussi, ensemble, nous sommes un « à deux ». En revanche, pour ce qui concerne les traces laissées sur le sol lunaire, les Dupondt usent cette fois d'un argument différent, relevant d'une simple addition arithmétique : une trace plus une trace, ça fait deux traces. Il n'y a plus qu'à admettre alors que les traces ne peuvent conséquemment pas être les leurs.

C'est le genre d'argument que mobilise Berkeley contre la thèse immanentiste de Locke, objectant qu'il est toujours possible de dénombrer différemment une même chose. Autrement dit, le nombre ne peut être abstrait directement par une chose qui l'instancierait, puisqu'une autre manière de compter lui attribuera un tout autre nombre. On peut par exemple varier le nombre exprimant une distance en changeant ses unités, par exemple en comptant « 1000 m » ou « 1 km ». De plus, les choses étant des agrégats, elles peuvent toujours être comptées de différentes manières, selon la façon dont on les décompose mentalement en parties distinctes. L'*Odyssée* est non seulement une œuvre, mais aussi vingt-quatre livres et bien plus encore de vers. Or de tels dénombrements sont tout aussi légitimes les uns que les autres.

> [...] the same thing bears a different denomination of number, as the mind views it with different respects. Thus, the same extension is one or three or thirty six, according as the mind considers it with reference to a yard, a foot, or an inch. Number is so visibly relative and dependent on men's understanding, that it is strange to think how any one should give it an absolute existence without the mind. We say one book, one page, one line; all these are equally units, though some contain several of the others. And in each instance it is plain, the unit relates to some particular combination of ideas arbitrarily put together by the mind. (BERKELEY, 2009a, p. 87)

Dans les *Fondements de l'Arithmétique*, Frege souligne à son tour les difficultés qui surgissent quand l'on tente de définir le nombre par l'agglomération de choses dont chacune est considérée comme une unité [7]. Être un, ce n'est pas une propriété qui, vraie de deux choses isolées, le resterait encore une fois ces choses regroupées : « Si on peut réunir 'Solon était sage' et 'Thalès était sage' en 'Solon et Thalès étaient sages', on ne peut pas dire 'Solon et Thalès étaient uns'. Ce qui serait possible si 'un' était une propriété possédée par Solon et Thalès. » (FREGE, 1969a, pp. 40-41).

Frege reprend aussi les arguments de Berkeley contre Locke, constatant qu'un même objet peut recevoir des attributions cardinales contradictoires, insistant sur l'idée qu'un objet étant un tout, il peut être différemment dénombré par ses parties :

> On peut partager un agrégat de manières bien différentes, et on ne peut pas dire que l'une seulement d'entre elles soit caractéristique. Par exemple, on peut partager une botte de paille

[7] La critique du réalisme immanentiste occupe une place important des *Fondements de l'Arithmétique* puisqu'elle s'étend du paragraphe 5 au paragraphe 28 sur les 109 paragraphes que contient l'œuvre.

> en fendant chaque brin, ou en les prenant un à un, ou en faisant deux bottes. (FREGE, 1969a, p. 150)

En particulier, Frege fait aussi remarquer que la dimension du contenu d'un concept diminue quand son extension s'accroît. Par exemple, si l'on compare les concepts « homme » et « mammifère », on peut constater que l'extension du premier est moins grande que l'extension du second car bien sûr, il y a moins d'hommes que de mammifères, la classe des premiers étant incluse dans la classe des seconds [8]. Comparons maintenant les contenus des deux concepts. Puisque le concept de mammifère est contenu dans celui d'homme, il appert que le contenu en compréhension du concept de mammifère est analytiquement plus restreint que celui du concept d'homme : un homme n'est pas réduit à sa seule spécificité de mammifère, il est également un bipède, et bien d'autres choses encore. Frege pousse alors le raisonnement à sa limite : si toutes les choses tombaient universellement sous un même concept, comme celui d'unité, alors le langage serait incapable d'en produire une compréhension. Le contenu d'un concept F ne peut avoir une épaisseur que lorsque certaines choses sont F là où d'autres ne le sont pas. Donc, si l'idée d'unité accompagnait toute chose, comme le prétend Locke, autrement dit, si elle était universelle, alors sa compréhension serait réduite à peau de chagrin.

> Il serait étrange que chaque chose possède la propriété en question (d'être un) [...]. L'assertion que Solon est sage a un sens parce qu'il est possible qu'un être soit dépourvu de sagesse. Le contenu d'un concept diminue quand son extension s'accroît ; et si celle-ci englobe toute chose, il faut que le contenu du concept s'amenuise à l'extrême. On ne voit pas bien comment le langage aurait pu produire un qualificatif impuissant à déterminer plus précisément un objet. (FREGE, 1969a, p. 158)

Finalement, Frege souligne que la meilleure façon d'infirmer la conception immanentiste de l'unité est encore de constater que de ce point de vue, le nombre 0 serait librement attribuable à tout inexistant. En affirmant que les nombres sont des idées issues des choses, on devrait alors déduire que Vénus n'ayant aucune lune, 0 serait une idée provenant de l'inexistante lune de Venus [9]. Mais Frege souligne que cela n'aurait aucun sens. On ne peut comprendre l'idée de 0 que si l'on prédique ce nombre, non à l'inexistante lune de Vénus, mais à son concept, qui est vide.

1.2.3 La conception frégéenne des liens entre unité et identité

La conception de l'unité comme cardinal d'un concept

La conception platonicienne de Frege concernant la notion d'unité l'oppose à la fois à celle de Locke, immanentiste, stipulant que l'unité est directement tirée de nos expériences sensibles, et à

[8] Évidemment, le fait que la partie soit plus petite que le tout n'est vrai que dans le cas d'ensembles finis. En revanche, l'argument est faux dans le cas d'ensembles infinis. Cela se voit en considérant par exemple l'existence d'une bijection des entiers naturels sur les entiers positifs pairs. Cette bijection permet de vérifier qu'à chaque élément du premier ensemble correspond un et un seul élément du deuxième, et réciproquement qu'à chaque élément du deuxième correspond un et un seul élément du premier. Il y a donc autant de nombres naturels que d'entiers positifs pairs.

[9] Cette conception a eu ses défenseurs, par exemple Meinong qui considère les inexistants comme des objets définis par compositionnalité de propriétés, comme la Montagne d'Or ou le cercle carré. Nous y reviendrons plus loin.

celle, conceptuelle, de Kant, qui en fait une donnée innée de notre appareil cognitif, c'est-à-dire une forme *a priori* de l'intuition. Ce que montrent les arguments de Frege, c'est qu'il n'est pas possible de considérer que l'unité, et plus généralement le nombre, soient la propriété d'une chose : « un objet auquel je peux légitimement attribuer plusieurs nombres n'est pas le suppôt propre d'un nombre » (FREGE, 1969a, p. 149). Le nombre n'est pas un prédicat décrivant une propriété comme celle d'être d'être rond ou d'être mou. Les particuliers concrets ne sont pas les véritables sujets de telles attributions cardinales. Ce sont seulement certains types de concepts à qui l'on doit attribuer ces propriétés. Le nombre est en effet pour Frege une propriété de l'extension du concept qui subsume les choses que l'on considère. Il ne s'applique donc pas aux objets, mais aux concepts qui subsument ces objets :

> Donner un nombre, c'est énoncer quelque chose d'un concept. [...] Si je dis : « le carrosse de l'empereur est tiré par quatre chevaux », c'est au concept « cheval qui tire le carrosse de l'empereur » que j'attribue le nombre quatre. (FREGE, 1969a, p. 176)

Frege définit l'unité à partir de l'identité. Il traduit l'énoncé « il y a exactement un F » par l'énoncé « il a au moins un F et il y a au plus un F ». Formellement, cela donne :

$$\exists x \, (\mathrm{F}x \wedge \forall y \, (\mathrm{F}y \rightarrow x = y))$$

De même, l'affirmation « il y a deux Fs » est interprétée par « il y a au moins deux Fs et il y a au plus deux Fs », soit, formellement :

$$\exists x \, \exists y \Big(x \neq y \wedge \mathrm{F}x \wedge \mathrm{F}y \wedge \forall z (\mathrm{F}z \rightarrow z = x \vee z = y) \Big)$$

L'activité consistant à dénombrer des Fs est basée sur une relation d'équivalence entre, d'une part, des ensembles de choses et d'autre part, des classes générales indiquant directement les cardinaux des premières. La conception cardinale du nombre repose sur l'équipotence des ensembles, autrement dit, sur le fait de corréler différents concepts par le même nombre si et seulement s'ils sont reliables par une relation un-à-un, c'est-à-dire par une bijection. Cette méthode repose elle-même sur le principe de Hume qu'on peut formuler ainsi :

Principe de Hume Le nombre de choses qui appartiennent à un concept F est identique au nombre de choses qui appartiennent au concept G si et seulement si il est possible de faire correspondre un à un les objets qui tombent sous le concept F avec ceux qui tombent sous le concept G.

Principe d'extensionnalité En outre, deux ensembles de choses sont identiques s'ils ont la même extension, autrement dit, s'ils ont exactement les mêmes membres.

Néanmoins, Frege conteste que le principe de Hume suffise à déterminer le nombre d'un concept. Car il faut encore pour cela qu'on dispose d'un *critère d'identité* qui permette de dire en quoi un particulier donné est bien différent d'un autre. Il s'agit d'une relation d'équivalence qui puisse servir de condition suffisante et nécessaire à l'identité sans être elle-même l'identité. Autrement dit, un critère d'identité est la donnée d'une relation d'équivalence $\mathcal{R}$ qui, par clause de non-circularité, ne doit pas être elle-même l'identité ni en dépendre. La relation $\mathcal{R}$ doit ainsi satisfaire les deux réquisits suivants :

Condition nécessaire Pour un individu x et pour un individu y, $x = y$ seulement si $\mathcal{R}xy$.

Condition suffisante Pour un individu x et pour un individu y, $x = y$ si $\mathcal{R}xy$.

Ces deux conditions peuvent finalement se résumer par la condition suivante :

Critère d'identité Pour un individu x et pour un individu y, alors $x = y$ si et seulement si $\mathcal{R}xy$.

Frege donne deux exemples de telles relations. Le premier relève de l'identité des ensembles. Un ensemble E est identique à un ensemble F si et seulement s'ils sont co-extensifs, c'est-à-dire, si et seulement si leurs éléments peuvent être mis en correspondance par une relation d'appariement « un-à-un ». Plus formellement :

$$\forall x\, \forall y \left(x = y \leftrightarrow \forall z(z \in x \leftrightarrow z \in y)\right)$$

Le second exemple de Frege traite de l'identité de direction de droites, dont le critère d'identité est la relation de parallélisme :

$$\forall x \forall y\ (d(x) = d(y) \iff \mathrm{P}xy)$$

Remarquons que les critères d'identité dans ces deux exemples sont à deux niveaux. En effet, pour le premier exemple, la relation d'identité entre deux nombres est corrélée à une relation bijective prenant pour sujet des ensembles. Dans le second exemple, l'identité des directions est interprétée en terme de parallélisme de droites. De fait, la relation d'équivalence ne doit pas concerner directement les objets considérés dans l'égalité afin d'éviter une circularité vicieuse. Elle doit au contraire s'appliquer à des entités d'une autre nature que celles dont on cherche à caractériser l'identité.

Différents critères d'identité ?

Il est manifeste que nous avons tendance à considérer qu'un arbre, qu'un chat et qu'un bateau constituent chacun un individu. Mais le sont-ils seulement pour nous ou le sont-ils aussi, « en soi » ? Les démarcations ontologiques qu'ils semblent imposer relèvent-elles de faits à propos du monde, ou s'agit-il seulement de conventions sociales véhiculées par le langage ordinaire ? En quoi consistent au juste nos jugements d'unité ? Doivent-ils être distingués de nos critères d'identité ? Qu'est-ce qui nous pousse à utiliser certains critères plutôt que d'autres ? Qu'est-ce qui nous incite à fonder nos jugements d'unité sur certaines conditions plutôt que d'autres ?

La première question est de déterminer la portée ontologique de la notion d'individu :

1. Quelle est la portée ontologique de la notion d'individu. Est-ce seulement une manière de parler des particuliers ou s'agit-il d'une réalité métaphysique ?

La question de l'individuation doit ainsi être distinguée selon un mode épistémique et selon un mode métaphysique. À cette fin, ma terminologie différencie *critère* et *crible*. Traditionnellement, les termes de « critère d'identité » et de « critère d'individuation » servent tous les deux à

indiquer la relation d'ordre métaphysique qui constitue l'identité de la chose considérée. J'utilise en revanche le terme de « crible » exclusivement dans un sens épistémique.

La question de l'individuation doit ensuite être déclinée selon trois questions différentes, quoique extrêmement corrélées.

2. Quels sont les critères de particularité qui permettent de dire qu'un particulier d'une sorte F est différent de tout autre particulier de la même sorte ?

3. Quels sont les critères de persistance qui permettent de dire d'un particulier x donné à un instant t_1 qu'il est le même individu qu'un particulier à un instant t_2 ?

4. Quels sont les critères d'unité qui permettent de dire de x qu'il compte pour un ?

Le tableau qui suit résume ces différentes interrogations selon le mode épistémique ou métaphysique d'interrogation qui porte sur elles :

Niveau	Crible épistémique	Critère métaphysique
Unité	Dans quel contexte x compte-t-il pour un ?	Par quelle propriété x fait-il unité ?
Unicité	Comment reconnaître qu'un certain x qui est F est différent de tout autre y qui est F ?	En quoi un certain x qui est F est-il différent de tout autre y qui est F ?
Persistance	Comment reconnaître qu'un certain x à t_1 est le même que y à t_2 ?	En quoi un certain x à t_1 est-il le même que y à t_2 ?
Identité sortale	De quels x peut-on dire qu'ils sont F ?	Qu'est-ce qui fait d'un x un F ?

Historiquement parlant, les deux premières questions se rapportent au principe d'individuation. De fait, l'identité pose des problèmes inverses aux nominalistes et aux réalistes des universaux. Pour les premiers, la particularité est primitive, il s'agit d'un fait brut ne nécessitant pas d'explication. Les individus sont d'emblée individués par leur existence, parce qu'ils sont des singularités. Aussi, pour le nominaliste, la question du dénombrement consiste à rendre compte de la manière dont des choses essentiellement différentes peuvent néanmoins être qualifiées avec le même prédicat. Il s'agit de reconnaître, au travers des expériences que nous faisons, la mêmeté d'une chose qui est *de toute façon d'emblée assurée d'être singulière*. Pour ce qui concerne les réalistes des universaux, le problème est inverse, puisqu'il consiste principalement à déterminer un principe qui assure qu'un particulier donné est bien différent de tout autre particulier de la même espèce. Les questions de l'identité et de la persistance doivent être différenciées de celle de l'unité qui s'affaire de parvenir à dénombrer des choses parce qu'elles sont segmentées en objets sur le fond du monde. Un « crible d'unité » permet ainsi d'affirmer que telle chose constitue *une* chose plutôt que la réunion de deux ou trois choses.

La question est alors de savoir si ces notions ont une quelconque portée métaphysique ou si elles ne relèvent que des seuls cribles épistémiques.

1.3 Les paradoxes de l'identité objectuelle

L'objet de cette section est de critiquer les théories de l'identité objectuelle en les recontextualisant par rapport aux paradoxes de l'identité des objets matériels.

Qu'est-ce qui nous permet de dire que nous avons affaire à une et même chose plutôt qu'à une autre ? L'identité logique s'applique-t-elle aux objets du monde ?

Dans *L'objet singulier* ((ROSSET, 1979)), Clément Rosset constate que l'étymologie du mot « identique » distingue deux acceptions opposées, quoique conjointes, de la mêmeté. Le premier sens, métaphysique, porté par l'expression « *is dem* » relève de l'identité ontologique, de la pleine coïncidence d'une chose avec elle-même. En revanche, la seconde acception, qui est portée par la signification « *idem* » qui indique une égalité, un processus d'identification, est épistémique. Or, commente Rosset, la première signification témoigne de la singularité essentielle de toute chose, là où la seconde témoigne de la présence du double. Toute équivalence, toute manière de parler d'égalité entre deux occurrences est en effet une tentative de *re*-produire la singularité, donc de la dédoubler en autre chose. De ce point de vue, la « mêmeté » revendiquée à l'égard de deux instances temporelles ne serait finalement qu'une manière parmi d'autres de dupliquer la singularité d'une chose en autre chose :

> Le privilège d'être celui-ci est très différente de celui de demeurer le même, [...] il y a presque autant de différence entre le ceci et son même dans le temps qu'entre le ceci et tout autre objet dont il puisse lui advenir d'être tenu pour le même [...]. Si je suis le même que celui que je serai demain, c'est à peu près au sens où je suis le même que mon voisin : un homme parmi d'autres. [...] Ce qui fait l'identité d'un ceci demeure ainsi étranger à la somme de ses égalités possibles, étranger même à une égalité avec son propre soi considéré à un autre instant du temps, l'écart au présent d'un des deux ceci étant la marque d'une différence qui suffit à faire de cette identification une affaire d'égalité et non d'identité. Ce qui fait l'identité d'un ceci est d'être ceci, et non d'être identique ou assimilable à quelque chose que ce soit. (ROSSET, 1979, p. 19)

On voit que Rosset entremêle ici les deux conceptions ontologiques et épistémiques de l'identité, et qu'il va plus loin que d'affirmer simplement que ce n'est pas parce qu'on identifie deux choses qu'elles sont nécessairement identiques. C'est une thèse plus forte qui apparaît ici, à savoir, que l'identification est justement une sorte de preuve de la distinction ontologique. Par exemple, l'identité temporelle tenue entre deux occurrences n'est qu'une forme de duplication, les *mêmes* (les occurrences identifiées entre elles) n'étant rien d'autre que des sortes de doubles particulièrement insidieux.

> La pensée de l'identité rassemble et confond les deux idées contraires du même et de l'autre : désignant à la fois et contradictoirement ce qui est *sans égal* et ce qui est *égal* à autre chose. Impossible, en bref, de penser le même sans penser du même coup son propre contraire, d'imaginer une identification qui ne passerait pas par le biais obligé d'un « ne pas être identique », en quoi se résume paradoxalement sa dernière et meilleure preuve : comme en témoigne la banale carte d'identité qui n'assure de la personne que pour autant qu'elle la présente comme différente d'aucune autre, non identique à quelque autre personne que ce soit, garantissant

> ainsi son « identité » particulière par l'exhibition d'un certificat général de « non-identité ». (ROSSET, 1979, p. 20)

Ce qu'on peut retenir des remarques précédentes de Rosset, c'est que penser l'identité métaphysique nous confronte d'emblée au problème du double, parce que nous ne pouvons concevoir l'identité métaphysique sans identification : pas de possibilité d'affirmer d'une chose qu'elle est la même sans charrier, dans le même temps, l'idée d'une identification à autre chose, c'est-à-dire la pensée d'une distinction numérique, comme le montre l'exemple de la carte d'identité, qui, en retour, prétend rendre compte de l'unicité ontologique de son détenteur.

1.3.1 Le problème général du changement

Une première objection qu'on pourrait retenir à l'égard de l'application de l'identité aux objets matériels est que son usage est résolument contradictoire avec leurs conditions d'existence, puisque ceux-ci sont toujours soumis au changement.

Si la question du double est à l'articulation de l'identité et de la ressemblance, elle peut certainement se poser quant aux rapports qu'entretient un individu à ses diverses occurrences temporelles. D'un côté, rien ne ressemble plus au chat qui a quitté la pièce que celui qui était présent il y a une minute encore. Nous disons d'ailleurs qu'il s'agit du même chat. Mais d'un autre côté, il est possible d'attribuer des prédicats différents aux deux occurrences : le chat était assis sur le tapis, maintenant, il court dehors. D'après le principe de distinction des discernables, nous devrions déduire la distinction numérique des deux occurrences temporelles. On aboutit alors à la conclusion paradoxale qu'il s'agit du même chat et qu'il ne s'agit pas du même chat. Des paradoxes de double surgissent selon la façon d'accommoder l'identité logique à l'identité diachronique d'une chose.

Dans ce qui suit, nous utiliserons le terme neutre de persistance pour qualifier la capacité d'une chose à demeurer la même à travers le temps. Une chose persiste si elle existe à différents moments du temps. Or, lorsqu'on pose la question de l'identité d'une chose à travers ses manifestations physiques, nos jugements oscillent sans doute entre deux possibilités contradictoires nous plaçant devant une alternative embarrassante :

(MU) soit les objets changent et ne demeurent pas les mêmes ;

(IU) soit ils restent identiques et ne changent pas.

Ces deux hypothèses représentent en fait deux thèses radicales et contradictoires : la première, le mobilisme universel (MU), affirme que l'identité n'est, en dernière analyse, qu'une illusion et que fondamentalement toute chose matérielle tombe nécessairement sous le régime de la différence, puisque toute chose est, par ses changements successifs, toujours actuellement différente de ce qu'elle fut. Mais ceci n'est-il pas contraire à nos usages ordinaires, par exemple lorsque nous disons que le chat que nous nourrissons aujourd'hui est le *même* que celui que nous avons nourri hier ? Nous pourrions alors céder à la seconde possibilité, l'immobilisme universel (IU), en affirmant que seule l'identité existe et que le changement n'est, *in fine*, qu'une illusion [10]. L'illusion

[10] Cette thèse a été défendue par l'école éléate sous la préséance de Parménide. Rappelons que Zénon d'Élée, disciple

ne devrait donc pas être attribuée à l'identité mais bien aux choses qui changent, les seules réalités étant les substances inaltérables, intemporelles, éternelles. Mais là encore, cela semble contraire à nos attitudes ordinaires : nous disons que le matou qui nous observe devant sa gamelle vide est bien différent du chaton qu'il fut.

La théorie du mobilisme universel

La paternité philosophique du mobilisme universel est traditionnellement attribuée à Héraclite d'Éphèse dont les écrits stipulent que rien ne demeure le même et que tout devient. La thèse rejette la persistance temporelle des choses. Le soleil est nouveau chaque jour, et l'on ne peut entrer deux fois dans le *même* fleuve. De sorte que l'identité n'est finalement prédicable que du devenir lui-même : la seule chose qui demeure, c'est le changement lui-même. Tout le reste *devient*. Les choses s'écoulent et passent dans le flux du temps. Cette thèse a connu divers partisans, puisqu'on la retrouve chez Descartes à propos de la substance étendue. La thèse resurgit également, d'une façon plus radicale, chez David Hume, pour qui la croyance en la persistance des choses résulte d'une habitude psychologique contractée par la répétition de faits ressemblants. Le monde de Hume ne contient que des nouveautés évanescentes, des existences fulgurantes qui apparaissent et disparaissent instantanément.

David Sedley ((David SEDLEY & BRUNSCHWIG, 1989)) rappelle aussi que le poète grec Épicharme fait usage de l'argument du changement comme ressort comique dans certaines de ses comédies. En particulier, l'argument qu'une personne qui change n'est plus la même sert de prétexte à diverses situations cocasses [11]. En voici un exemple caractéristique : un créancier fait constater à son débiteur l'échéance de son emprunt. Mais ce dernier en conteste la dette : « vois comme je suis gras et souviens-toi de la maigreur de celui à qui tu confias ton argent. N'étant manifestement pas le même que celui qui fit l'emprunt, je ne te suis donc pas redevable. » La manigance du débiteur consiste ainsi à faire passer le changement qualitatif d'une même personne pour un changement numérique de personne. Cependant, l'argument est retourné contre le débiteur, qui, rossé par son créancier, pousse de vaines suppliques sollicitant l'indulgence. Mais tel est pris qui croyait prendre, car le créancier renvoie au débiteur son propre argument : si l'écoulement du temps change inexorablement les individus, alors la personne qui est suppliée n'est pas la même que celle qui inflige le supplice. Aussi la première pourrait-elle bien se résigner à épargner le débiteur, cela demeurerait cependant sans effet sur la seconde qui continuera néanmoins à le rosser. Selon un tel point de vue, créanciers et débiteurs seraient des individus impossibles à atteindre car, à peine apparus, ils seraient déjà évanouis. Toute chose temporelle serait ainsi soumise au régime du double par différences successives : ne demeurant jamais la même personne, devenant toujours

---

de Parménide, a échafaudé de célèbres paradoxes visant à réfuter l'existence du mouvement. Puisque le changement est illusion, les choses ne changent pas même de place. Toute flèche tirée demeurera immobile. Toute tentative de rattraper à la course une tortue est vaine, tant celle-ci restera à jamais inaccessible. L'erreur de Zénon est toutefois de concevoir le temps et l'espace de manière discontinue. Il considère les lieux et les instants comme des parties fragmentées. Les périodes de temps et les distances spatiales sont, entre elles, cloisonnées, discontinues. D'où l'immobilité de la flèche, sans cesse présente en chacune des tranches spatio-temporelles où elle réside. Et d'où l'immobilisme d'Achille, condamné, à la manière de Sisyphe, à accomplir une tâche sans cesse réajournée, incapable de franchir une distance qui le sépare perpétuellement de son but.

[11] Dans l'analyse de Sedley, il s'agit plus précisément de l'argument de croissance que l'on retrouve dans les paradoxes sorites. Mon illustration en simplifie ici le propos.

une autre. On croyait avoir affaire à x, mais le voilà déjà devenu y.

L'argument qui alimente la pièce d'Épicharme semble généralisable à tout ce qui existe : toute chose étant soumise au changement, elle devient constamment autre chose. En cela, le double pourrait se dire de tout ce qui existe, toute chose étant toujours différente d'elle-même parce que constamment changeante. Dès lors, l'identité serait absolument inapplicable aux choses matérielles et la temporalité serait tout simplement le domaine des doubles. Le monde ne serait qu'un vaste miroir aux doubles, où toute chose laisserait immédiatement présence à une autre, en dépit de nos inclinaisons à voir, entre ces « mêmes » choses, des liens serrés, mais illusoires.

On peut effectivement se demander si l'identité numérique est véritablement applicable à des choses constamment changeantes. Distinguer l'identité de la ressemblance ne semble pas ici suffire à régler le paradoxe. Peut-on saisir l'identité numérique à travers le changement qualitatif, ou devrions-nous plutôt affirmer la duplicité de toute chose en disant que nos décomptes ne sont que des façons lâches de parler ? Les changements qualitatifs qui affectent un *même* individu semblent en effet constituer une entorse logique au principe d'indiscernabilité des identiques. Car si un individu à t_0 est discernable du « même » individu à t_1, autrement dit, s'il diffère par une propriété, ne serait-ce que parce qu'il est assis à t_0 et debout à t_1, ne devrions-nous pas en conclure qu'il s'agit, à proprement parler, du « même » individu qu'en un sens extrêmement « relâché » ? Plus précisément, l'argument général qui motive le mobilisme universel peut être formulé de la manière suivante :

(1)	Toutes les choses existent dans le temps.	hypothèse 1
(2)	Une chose x à t_0 est discernable de la chose x à t_1.	hypothèse 2
(3)	x à t_0 est numériquement différente de x à t_1.	(1) et (2) application du principe de distinction des discernables

Principe d'indiscernabilité des identiques relativisé

Ni le mobilisme universel ni l'immobilisme universel ne paraissent répondre à nos attentes concernant l'existence des choses dans le temps. Car l'une et l'autre théorie contredisent bien trop nos intuitions. Comme le souligne Épicharme, si nous ne donnions pas quelque crédit à l'identité temporelle des personnes, nous serions quittes de ne jamais parvenir à nous engager contractuellement. La présumée incompatibilité entre changement et identité repose en fait sur un usage abusivement diachronique du principe d'indiscernabilité des identiques, qui pourrait être formulé ainsi : si un individu x existe à deux moments quelconques t_1 et t_2, alors toute propriété que x possède à t_1 est aussi une propriété que x possède à t_2 et réciproquement. Pourtant, quand il s'agit de la même fleur, il est tout aussi vrai de celle d'hier et de celle d'aujourd'hui que, hier, elle était pleine de fraîcheur et qu'aujourd'hui, elle est fanée. Le principe d'indiscernabilité des identiques et sa contraposée, le principe de différence des discernables, doivent être relativisés par rapport au temps :

(IIdS) Si x est le même individu que y, alors toute propriété que x possède à un quelconque moment t est aussi une propriété que y possède à t et réciproquement.

(DDiS) S'il existe une propriété qu'un individu x possède à un moment t sans que l'individu y ne

la possède à t, alors x est distinct de y.

Ainsi, le fait qu'un même objet exemplifie au moment t_1 une propriété contradictoire avec la propriété qu'il exemplifie au moment t_2 n'implique pas une distinction numérique, car le principe d'indiscernabilité des identiques n'est applicable que de manière synchronique.

De ce point de vue, le changement ne paraît pas absolument incompatible avec l'identité. Plus encore, il en est même la condition. C'est justement parce qu'une chose matérielle change qu'on peut dire d'elle qu'il s'agit de la même chose. Nous considérons les individus comme des choses qui persistent tout en changeant. La persistance ne relève pas de l'inaltérabilité d'une substance qui serait comme préservée des aléas de la temporalité. Un objet, un individu persistent parce qu'ils évoluent à travers les modifications occasionnées durant leurs carrières. Comme le souligne Quine, « l'identité continue d'une personne à travers le temps lui est attribuée non pas en fonction de la conservation de sa substance, mais sur la base du renouvellement continu de celle-ci et des modifications continues qui affectent la forme, la masse, et les habitudes de cette personne » ((W. V. O. QUINE, 1992, p. 111)). Les créanciers peuvent donc continuer à dormir sur leurs deux oreilles : leurs débiteurs restent redevables des dettes qu'ils ont contractées : car si les choses sont des devenants, elles sont également des continuants.

Parmi les candidats qui semblent propices à jouer le rôle de critères d'individuation pour un objet matériel, deux cribles d'identification retiennent immédiatement l'attention : d'une part l'emplacement qu'occupe une chose et d'autre part sa constitution matérielle. Il s'agit bien sûr de cribles que nous utilisons couramment pour *reconnaître* une certaine chose. Un stylo retrouvé à l'emplacement de sa perte est identifié avec plus d'assurance que s'il était retrouvé ailleurs. Ainsi, les relations spatio-temporelles et matérielles sont certainement propices à assumer le genre de relations recherchées. Mais peut-on en faire des critères d'individuation ? Notons d'emblée que la relation de coïncidence spatiale n'est pas une condition nécessaire à l'identité temporelle, puisqu'une même chose peut occuper, de façon diachronique, des emplacements différents. Nous pouvons certes amender notre critère en recourant à la notion de déplacement. Nous pouvons tenir compte que les objets, au moins à l'échelle macroscopique ne font ni sauts dans l'espace ni sauts dans le temps. Les déplacements d'un individu sont continus. La continuité signifiant ici que pour tout écart de distance δ aussi petit soit-il, il existe un intervalle temporel dans lequel les différences de distances séparant une même entité considérée à des moments différents de cet intervalle n'excèdent pas δ. Ainsi, la continuité spatio-temporelle et la continuité de constitution matérielle (physico-chimique) semblent fournir des critères pertinents pour l'individuation des individus :

(CSM) x et y sont deux individus de la même sorte occupant la même position spatiale ou il existe une relation de continuité spatio-temporelle menant de la position de x à t_1 à la position de y à t_2 et x et y partagent une relation de continuité constitutive.

Un individu peut changer continûment de position spatio-temporelle, mais il peut aussi voir sa constitution modifiée. Un homme qui perd un cheveux demeure le même. En règle générale, les changements de constitution se font de manière continue. Un homme ne passe pas instantanément, mais progressivement, de l'état d'enfant de 40 kg à celui d'adulte de 80 kg. Cela pourrait inciter à affirmer que pour un individu donné et pour tout écart de masse δ très petit, il existe un

intervalle temporel dans lequel les différences de masses considérées à des moments différents de cet intervalle n'excèdent pas δ.

1.3.2 L'individuation par la constitution matérielle et ses paradoxes

Il faut toutefois reconnaître que si la continuité de constitution fournit un crible d'identification généralement assez efficace, elle n'est néanmoins pas une condition nécessaire pour l'individuation métaphysique. Car un individu peut perdre un bras en réduisant brusquement sa masse de 70 kg à 63 kg, cela sans jamais passer par une masse de 69 kg. Certes, il est sans doute difficile de déterminer le moment exact où l'homme passera de 70 kg à 63 kg, mais il est clair que ce changement ne se fera pas par gradations continues. Les changements de constitution matérielle tolèrent des sauts, ils ne sont pas toujours progressifs. La discontinuité de constitution des artefacts est même un phénomène assez fréquent : un bateau peut être démonté puis remonté. D'un moment à l'autre, c'est bien le même bateau, nonobstant qu'entre les deux, il n'existe tout simplement plus.

Les principes de tolérance et de restriction constitutive

Si l'on admet que les choses peuvent changer en persistant, il faut toutefois reconnaître que certains changements s'avèrent néfastes à leur survie. Une statue peut perdre un bras et poursuivre sa carrière, mais si elle est brisée ou refondue, sa carrière temporelle se trouvera interrompue. Notons que la discrimination entre les changements qui n'affectent pas la carrière d'une chose et ceux qui l'affectent n'est pas seulement de nature. En effet, l'accumulation de changements bénins mais successifs peut occasionner la disparition de l'objet auquel elle s'applique. Comme le fait remarquer Eubulide, en retirant un à un les grain d'un tas de sable, on finit par n'avoir plus de tas. Ainsi, si les objets tolèrent certains changements, leur persistance est également tributaire d'une forme de restriction sur l'accumulation des modifications qu'ils subissent. Nous pouvons formuler ceci en énonçant deux principes. Le premier principe, dit de « tolérance constitutive » affirme que :

(TC) Pour tout individu d'une sorte S, il existe un seuil en deçà duquel les changements occasionnés ne rompent pas la persistance temporelle de cet individu.

Le deuxième principe, dit « principe de restriction constitutive » affirme que :

(RC) Pour tout individu d'une sorte S, il existe un seuil au delà duquel les changements occasionnés rompent la persistance de l'objet.

S'il existe un seuil d'altération qui, dépassé, engendre la cessation de l'existence de la chose considérée, il n'est cependant pas évident de situer avec exactitude ce seuil. Nous ne pouvons généralement pas positionner le curseur à l'endroit exact qui délimite la persistance temporelle d'une chose de la cessation de son existence. On se contente habituellement d'un degré qui fait consensus. L'articulation des deux principes implique donc un certain vague. En outre, il faut encore remarquer que le degré de tolérance et de restriction au changement n'est pas absolu, qu'il dépend évidement de la sorte d'individu considérée. Comme nous l'avons déjà remarqué, une statue peut perdre un bras tout en demeurant la même statue. Mais le volume de marbre qui la constitue ne peut perdre cette quantité de matière sans cesser d'être le même volume de marbre.

Paradoxes sorites de diminution et d'accroissement

L'articulation des principes de tolérance et de restriction constitutive avec la transitivité de l'identité conduit à des difficultés bien connues. De petits changements successifs occasionnent de grandes altérations. C'est ce que constatent les paradoxes sorites dont l'énoncé général peut être formulé ainsi :

(Paradoxe sorite) Un artefact duquel on retire (on ajoute) progressivement, une à une, d'infimes parties matérielles constitutives est à la fois identique et différent de l'artefact original.

Ces paradoxes concernent donc la constitution matérielle d'une chose qui se voit continuellement amenuisée jusqu'à disparaître. Leur schéma général obéit au mécanisme suivant :

- On considère un individu x à un moment t_0 de son existence.
- On applique ensuite, à un moment t_1, un certain type de changement dont le degré est inférieur au seuil de tolérance constitutive de cet individu.
- Puis à t_2, on pratique une nouvelle fois une altération bénigne de la constitution de cet individu. Par TC, on peut inférer de x-à-t_1 qu'il demeure identique à x-à-t_0.
- On réitère n fois de tels changements sur chacune des nouvelles occurrences de x. On obtient ainsi une succession d'individus x-à-t_0, ..., x-à-t_n telle que pour tout $i \in \{0, ..., n-1\}$ on ait x-à-$t_i = x$-à-t_{i+1}.
- De la transitivité de l'identité, on peut déduire alors que x-à-$t_0 = x$-à-t_n.
- Si le nombre n de changements est suffisamment grand, le seuil de tolérance et dépassé, et il advient que, par principe de restriction constitutive, x-à-t_n n'est plus identique à x-à-t_0.
- On conclut alors qu'on a à la fois x-à-$t_n = x$-à-t_0 et x-à-$t_n \neq x$-à-t_0.

Paradoxes de substitution (bateau de Thésée)

Dans les paradoxes sorites, l'altération compatible avec le principe de tolérance constitutive consiste en un accroissement de matière (paradoxe du créancier) ou à une diminution de matière (paradoxe du tas de sable). Dans d'autres paradoxes, les altérations de constitution subies sont d'une autre nature. Elles peuvent par exemple consister en substitutions successives de parties. C'est notamment le cas du bateau de Thésée, paradoxe dont Plutarque rapporte ainsi la trame :

> Le Vaisseau sur lequel Thésée alla et retourna était une galiote à trente rames, que les Athéniens gardèrent jusqu'au temps de Démétrius le Phalérien, en ôtant toujours les vieilles pièces de bois, à mesure qu'elles se pourrissaient, en y remettant des neuves en leurs places : tellement que depuis, dans les disputes des philosophes, touchant les choses qui s'augmentent, à savoir si elles demeurent unes, ou si elles se font autres, cette galiote était toujours alléguée pour l'exemple de doute, parce que les uns maintenaient que c'était un même vaisseau, les autres, au contraire, soutenaient que non (PLUTARQUE, 1937, p. 21)

Le bateau rénové est-il le même bateau que celui dont toutes les parties ont été changées ? En un sens, c'est le même bateau dont on parle, en un autre, il s'agit d'un bateau neuf.

Pour simplifier le mécanisme du paradoxe, supposons qu'un bateau soit la réunion de trois parties $A \cup B \cup C$. Admettons en outre que ce que nous appelons « bateau » obéisse aux principes de tolérance et de restriction constitutives suivants : le bateau survit au changement d'une ou deux de ses parties constitutives, mais pas au changement de plus de deux parties. Appliquons maintenant ces deux principes au schéma sorite comme illustré sur la figure 1.1. Au moment t_1, la partie A du bateau originel est remplacée par la partie A'. Par (TC), le bateau-à-t_0 est identique au bateau-à-t_1. Au moment t_2, la partie B est remplacée par la partie B'. A nouveau, par (TC), le bateau-à-t_1 est identique au bateau-à-t_2. Enfin, au moment t_3, la partie C est remplacée par la partie C'. Toujours par (TC), le bateau-à-t_2 est identique au bateau-à-t_3. Dès lors, la transitivité de l'identité implique que le bateau-à-t_0 est identique au bateau-à-t_3. Mais d'un autre côté, par principe de restriction constitutive, le bateau n'a pas survécu au changement de plus de deux pièces.

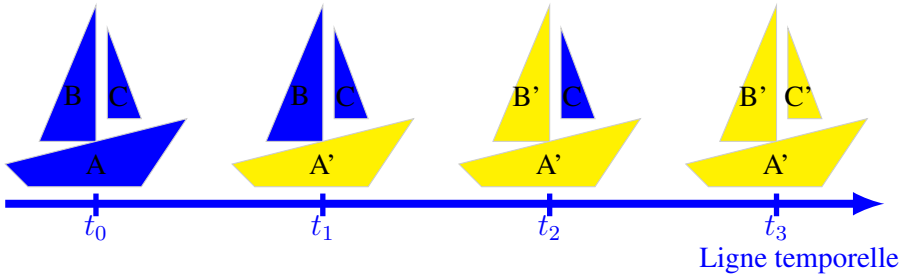


FIGURE 1.1 : La version de Plutarque du paradoxe du Bateau de Thésée

Quelle réponse offrir à ce paradoxe ? Nous pouvons tout simplement refuser de dire que les bateaux sont distincts. On peut soutenir par exemple que ce problème d'identité n'est finalement qu'un problème purement conceptuel. Car il semble que la nature intentionnelle de l'artefact serve d'issue au paradoxe. En effet, nous utilisons le terme « bateau » pour désigner une certaine production matérielle reproductible dans des conditions similaires et remplissant une certaine fonction. Lorsque l'on change des pièces, on n'altère pas l'identité du bateau parce qu'elle dépend de notre intention de le considérer comme un bateau et parce que nous pouvons reproduire les conditions de sa création. Ainsi, la matière du bateau dont on a changé toutes les parties n'est certes pas la même, mais en tant que bateau, c'est-à-dire en tant qu'artefact, il est encore pensé comme *le même bateau* que celui d'origine. C'est donc le principe de restriction qu'il faut rejeter : un bateau complètement rénové peut demeurer le même, parce que cela ne dépend pas de conditions métaphysiques, mais de nos intentions de le considérer ainsi.

Le paradoxe du bateau de Thésée peut cependant prendre une tournure plus complexe. Dans la version qu'il rapporte dans *De Corpore* ((HOBBES, 1655, p. 83, ch. XI, § 7)), Thomas Hobbes ajoute à la version de Plutarque une hypothèse supplémentaire. Il demande de considérer que les différentes parties détachées sur le bateau d'origine servent progressivement à constituer un second bateau comme illustré sur la figure 1.2. En plus d'un problème de double temporel, on est maintenant également confronté à un double synchronique du bateau original.

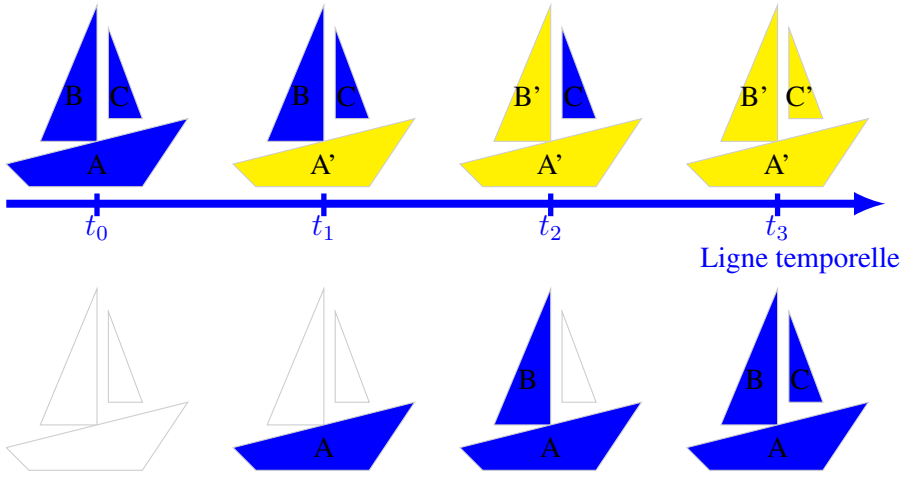


FIGURE 1.2 : La version d'Hobbes du paradoxe du Bateau de Thésée

En effet, on n'a donc plus affaire à un seul, mais à deux bateaux : du bateau original (désormais noté O pour « original ») émanent maintenant deux objets : d'une part un bateau constitué de parties neuves (désormais noté R pour « rénové ») et d'autre part un second bateau (désormais noté D pour « démonté ») constitué des parties usées. Si nous posons la question de savoir quels sont les liens qui relient ces trois choses, nous pourrions demander lequel des deux est identique à l'original. Or il ne semble y avoir ici que quatre options comme illustré sur la figure 1.3 :

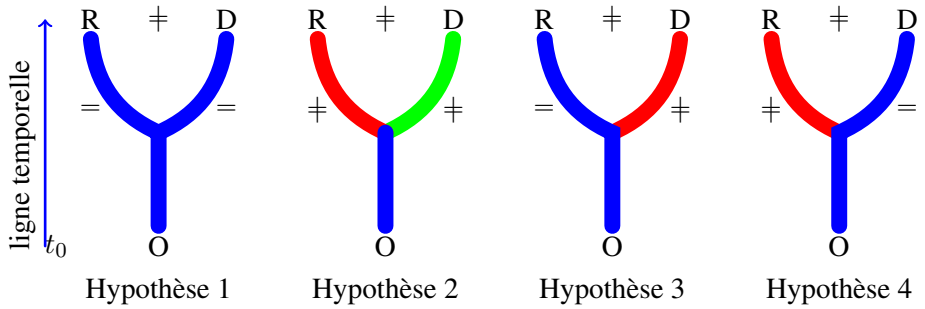


FIGURE 1.3 : Quatre hypothèses pour la fission d'un artefact

1. les deux bateaux D et R sont identiques à O ;
2. aucun des deux bateaux n'est identique à O ;
3. R mais pas D est identique à O ;
4. D mais pas R est identique à O.

Considérons successivement chacune des possibilités. La première pourrait paraître incongrue. En effet, si R et D étaient tous les deux identiques à O, la transitivité de l'identité appliquée sur R = O et O = D impliquerait inévitablement l'identité R = D, ce que n'accepterait pas un partisan de l'identité métaphysique. La deuxième hypothèse n'est pas plus pertinente, car il paraît peu probable qu'aucun des deux bateaux ne soit identique à l'original. En effet, cela contredirait ce que nous avons affirmé dans la version précédente du paradoxe et à l'intuition qu'un bateau puisse survivre à son démontage, à son remontage ou au changement graduel de ses parties constitutives. Il reste donc les deux possibilités qu'un seul parmi les deux bateaux est identique à l'orignal. Mais lequel ? Deux expériences de pensée modales semblent tirailler l'identification à O d'un bateau à l'autre. D'un côté, on pourrait avoir de bonnes raisons de croire que c'est D qui est identique à O. En effet, puisque les pièces de O auraient pu être détachées sans être remplacées, il aurait pu n'y avoir qu'un seul bateau, à savoir, D. En fait, il ne s'agirait ni plus ni moins que la situation assez banale du démontage et remontage d'un même bateau, comme le pratiquaient d'ailleurs les romains. Mais d'un autre côté, on pourrait également avoir de bonnes raisons de croire que c'est R qui est identique à O. En effet, les planches de O auraient pu ne pas être ré-utilisées pour construire D. Or, puisqu'un bateau peut survivre au remplacement graduel de ses parties, R devrait être identique à O. Le bilan qu'en tirera le métaphysicien de l'identité est donc qu'aucune des possibilités n'est enviable car il n'est pas satisfaisant de faire dépendre l'identité du bateau d'événements aussi contingents que la décision de remplacer ou non ses parties constitutives.

Que répondre à cette nouvelle version du paradoxe ? Commençons par constater avec Stéphane Ferret ((FERRET, 1998)) que le paradoxe perd de sa force si on considère que le bateau est démonté en une seule fois et non plus progressivement. Dans ce cas, nous dirions probablement que le bateau rénové est différent de l'original et que celui qui est ré-assemblé est bien identique au bateau original. Cette hypothèse semble faire disparaître le paradoxe parce que nous obtenons deux bateaux dont l'un est nouveau (R) et l'autre est un bateau simplement démonté-remonté (O = D). D'autre part, si l'on considère que les parties détachées étaient ré-assemblées en une cabane, l'identité serait inévitablement attribuée au bateau rénové pour la raison qu'une cabane et qu'un bateau sont des sortes différentes de choses. La cabane et le bateau initial sont bien constitués des mêmes parties, mais en revanche, ils n'ont pas le même usage.

Ce qui précède indique que nos pratiques d'identification changent en fonction du contexte. Précisons toutefois qu'une théorie du meilleur candidat, qui permettrait de choisir le candidat le plus apte à remplir la condition d'identité en fonction du contexte et de l'usage, n'implique pas nécessairement une identité vague. La contingence du choix n'est pas à situer au niveau de la relation d'identité, mais au niveau de la fonction que nous attribuons aux artefacts, et au contexte dans lequel nous employons cette relation.

Paradoxe des coïncidents temporaires (Lumpl et Goliath)

Nous n'en avons pas encore terminé avec les paradoxes de constitution matérielle. Considérons maintenant le paradoxe des coïncidents temporaires, introduit en 1975 par Allan Gibbard ((GIBBARD, 1975)). Il met en scène un morceau d'argile, désigné sous le nom de Lumpl, qui existe à l'instant t_0. A un autre instant, t_1, Lumpl est façonné en statue de Goliath puis, à t_2, cette statue est ensuite aplatie. On constate donc que, dans l'intervalle temporel allant de t_1 à t_2, le morceau

d'argile Lumpl coïncide complètement avec la statue Goliath. En revanche, puisque Lumpl existait avant Goliath et subsiste à l'aplatissement de la statue, le bloc d'argile possède des propriétés historiques différentes de celles de Goliath. La contraposée du principe d'indiscernabilité des identiques implique donc que Lumpl est bien numériquement différent de Goliath. Il y a donc deux choses qui se tiennent, simultanément, au même emplacement : une statue et un bloc d'argile.

(Paradoxe des coïncidents temporaires) Dans le cas de carrières qui ne sont pas coextensive, une statue n'a pas les mêmes propriétés historiques que le bloc de matière qui la constitue. De ce fait, la statue est et n'est pas identique au bloc qui la constitue.

Devons-nous stipuler que la statue est numériquement différente du bloc d'argile ? Doit-on en inférer qu'il y a deux choses présentes au même emplacement, d'une part une statue, et d'autre part un bloc de marbre ? Les doubles ressurgissent à nouveau.

Une réponse donnée par les métaphysiciens au paradoxe de Lumpl et Goliath passe par l'admission d'objets coïncidents : le bloc, prétendent-ils, n'est pas identique à la statue, bien que tous les deux occupent le même emplacement spatio-temporel. C'est en ce sens que David Wiggins soutient par exemple que :

> [...] the absoluteness thereby discovered in identity will carry with it certain unexpected consequences, the possibility for instance that a portion of space may be co-occupied at a time by distinct things (provided, at least, that the things in question belong to different basic or individuative kinds). (WIGGINS, 1980, p. 22)

Une conséquence qu'en tire Wiggins, c'est que l'identité est *sortalement* dépendante. Il n'y a aucun sens à demander si x est le même que y si l'on ne précise pas dans le même temps le genre de choses que sont x et y. Une affirmation d'identité n'a donc de sens que relativement à une sorte de chose :

(Thèse de dépendance sortale de l'identité) La question exprimée par la phrase « x est-il le même que y » n'admet pas, *simpliciter*, de réponse déterminée. Pour lui conférer une signification, il est nécessaire de préciser un terme sortal « F » qui détermine à quelles conditions x est le même F que y.

La statue et le bloc comptent-ils pour une ou deux choses ? Si l'on admet que nos dénombrements reposent sur l'identification, et si l'on admet que celle-ci est sortalement dépendante, alors il faut aussi admettre que nos dénombrements sont sortalement dépendants. Frege constatait par exemple qu'il est vain de demander de compter des choses sans autre précision. En effet, comment compter les objets présents sur une table ? Devrions-nous compter non seulement telle assiette et telle baguette de pain, mais aussi les miettes qui en sont tombées, les tâches sur la nappes, les cheveux, les flaques de sauce sur les assiettes, voire les ombres projetées par les ustensiles ? Le concept « chose sur la table » n'est pas suffisamment individuant pour indiquer ce qui doit compter. Comme le fait remarquer aussi Frege,

> Si je donne une pierre à quelqu'un, en lui disant d'en déterminer le poids, j'ai parfaitement indiqué ce qu'il doit chercher. Mais si je lui mets dans la main un paquet de cartes à jouer en

> lui disant d'en déterminer le nombre, il ne sait pas si je veux connaître le nombre des cartes, ou celui des jeux de cartes complets, ou encore la valeurs de ces cartes à l'écarté. En lui remettant le paquet de cartes, je n'ai pas encore indiqué parfaitement ce qu'il devait chercher ; je dois ajouter un mot : carte, jeu ou valeur. (FREGE, 1969a, pp. 148-149)

Pour compter, nous devons d'abord préciser la catégorie des choses qui importent pour notre décompte, autrement dit, la catégorie de choses qui comptent comme des choses à compter.

(Thèse de dépendance sortale du dénombrement) La question exprimée par la phrase « Combien ? » n'admet pas, *simpliciter*, de réponse déterminée. Pour lui conférer une signification, il est nécessaire de préciser un terme sortal « F » qui détermine le nombre de choses qui sont à compter.

Pour Wiggins, l'identification des objets est relative aux concepts sortaux, mais ce n'est pas le cas de leur identité métaphysique, qui demeure, quant à elle, absolue. La solution que Wiggins apporte aux paradoxes de constitution matérielle consiste en une thèse essentialiste néo-aristotélicienne. L'identité métaphysique est corrélée à l'essence de la chose considérée. Mais, demandera-t-on alors, qu'est-ce qui confère sa nature sortale à une chose ? Quel lien y a t-il entre une chose et le prédicat qui s'y applique ? Pour ceux comme Wiggins qui prennent les propriétés au sérieux, il s'agit de déterminer une relation $\mathrm{R}_{\mathcal{E}}$ dépendant d'une espèce $\mathcal{E}$ obéissant au principe suivant :

[Identité spécifique] Un x appartient à la même espèce $\mathcal{E}$ que y si et seulement si $\mathrm{R}_{\mathcal{E}}xy$.

La théorie essentialiste est une théorie de la substance qu'Aristote a principalement exposé dans les *Catégories* et dans la *Métaphysique* [12]. La conception aristotélicienne de la substance repose sur l'idée que pour être *un* bateau, *un* tigre ou *une* personne, il faut être *bateau*, *tigre* ou *personne*.

Une telle conception implique une démarcation entre des entités de différentes natures, les particuliers et les universaux, qui peuvent être conçus comme transcendantaux (platonisme) ou *in rebus* (Aristote), autrement dit, n'ayant pas d'existence en dehors du particulier dans lequel elles s'instancient. Contrairement aux particuliers qui sont les substances premières, les universaux, ou substances secondes, ont la capacité d'être instanciés de manière synchronique à différents emplacements. Par ailleurs, les particuliers peuvent être prédiqués par des universaux mais la réciproque est fausse :

> La substance dont on parle principalement, d'abord et avant tout, c'est celle qui ne se dit pas d'un certain sujet et n'est pas inhérente à un certain sujet. Ainsi, un certain homme ou un certain cheval. Sont dites, en revanche, substances secondes, les espèces où prennent place les substances dites au sens premier, ces espèces tout comme leurs genres. Ainsi un certain homme se trouve dans une espèce, et le genre de l'espèce est l'animal. Ce sont donc elles qu'on dit secondes, substances comme l'homme et l'animal (ARISTOTE, 2001, p. 7)

[12] Notons au passage que le terme substance vient du latin *substantia*, il est généralement employé afin de traduire le terme grec οὐσια (*ousia*) qui signifie propriété, appartenance, ce qui est essentiel à une chose. Le mot grec a aussi été traduit par *essentia*, par exemple par Boèce dans *Contre Eutychès*.

La relation $\mathrm{R}_{\mathcal{E}}$ recherchée est ainsi définie par l'instanciation, capacité à refléter une propriété universelle portée par un particulier. Ainsi, $\mathrm{R}_{\mathcal{E}}xy$ signifie que x et y instancient le même universel $\mathcal{E}$.

La théorie de Wiggins admet donc à la fois les doubles synchroniques que constituent les universaux, mais en outre, elle admet l'existence d'entités coïncidentes comme les statues et les volumes de matière qui les constituent. C'est d'ailleurs l'infléchissement essentialiste de sa théorie de l'identité qui produit l'effet de redondance ontologique. Il semble inévitable qu'en acceptant d'attribuer une teneur ontologique à nos prédicats, nous soyons contraints de considérer qu'une même chose vue sous différents aspects (statue, matière) relève en réalité de plusieurs choses coïncidentes. De ce point de vue, les doubles métaphysiques émanent directement des théories de l'identité métaphysique reposant sur la théorie des universaux.

Selon la position essentialiste représentée par Wiggins, les questions de l'unité et de la persistance sont finalement ramenées à la seule question de l'appartenance sortale. C'est parce qu'une chose appartient à une sorte F que cette chose dispose de certaines conditions de persistance et qu'elle est découpée comme unité sur le fond du monde.

Paradoxe de Dion et Théon

Considérons encore un dernier paradoxe de constitution, celui-ci appliqué aux organismes. Il s'agit du paradoxe de Dion et Théon, qui fut probablement élaboré par le philosophe stoïcien Chrysippe. Voici la trame de la version antique. Soit, à un instant t_0, un individu nommé « Dion ». On considère la partie propre de Dion constituée de l'intégralité de son corps moins sa jambe gauche. On donne pour nom « Théon » à cette partie propre de Dion. Admettons maintenant qu'à un instant ultérieur t_1, Dion subisse l'amputation de sa jambe gauche. Constatons alors qu'après t_1, Dion et Théon possèdent exactement la même constitution matérielle. En outre, Dion et Théon partagent également exactement les mêmes comportements, les mêmes relations. Quand Dion parle, Théon parle aussi, quand Dion s'appuie sur sa jambe droite, Théon le fait également, etc. De fait, les deux individus sont, du point de vue matériel, parfaitement coïncidents [13].

Nous serions sans doute inclinés à penser que Dion et Théon sont simplement le même individu. En effet, comment expliquer autrement leur coïncidence parfaite ? Mais ce n'est peut-être pas si simple que cela. Car notons que si Dion et Théon sont parfaitement indiscernables du point

[13] Le paradoxe de Dion et Théon a été accommodé de la façon suivante par Peter Geach ((P. T. GEACH, 1962)) : un chat, Tibbles, se repose sur un tapis. Appelons Tib la partie de tissu félin composée de Tibbles exception faite d'un de ses poils. Ainsi, sont présentes sur le tapis deux entités partageant quasiment, mais pas tout à fait (nous pourrions dire, « à un poil près »), l'intégralité de leur constitution matérielle : d'une part un chat, Tibbles, et d'autre part un tissu félin, Tib. Admettons alors que Tibbles perde le poil dont il a été question. Avons-nous encore deux entités distinctes ? Puisque les propriétés historiques de Tibbles et de Tib diffèrent (l'un a possédé le poil, l'autre pas), le principe de diversité des discernables implique la différence numérique entre Tibbles et Tib. Autrement dit, sur le tapis, nous avons affaire à deux choses qui ne diffèrent ni temporellement, ni spatialement, ni matériellement et qui pourtant ne sont pas identiques. En outre, au moment de la perte du poil, Tibbles partage avec Tib toutes ses qualités, dont ses qualités essentielles. Cela induit qu'après t_1, Tib possède la même identité spécifique que Tibbles. Autrement dit, Tib est un chat. Mais si Tib est un chat après t_1, on n'a alors aucune raison de nier qu'il le fut aussi avant t_1. Ainsi, à t_0, il n'y a pas un mais deux chats sur le tapis. À moins qu'il n'y en ait encore plus, car en répétant le même argument à chacun des poils de Tibbles, on peut invoquer d'innombrables chats simultanément présents sur le tapis (qui sont d'ailleurs, notons-le, eux-mêmes multiples pour des raisons similaires).

de vue de leur constitution matérielle, il existe néanmoins quelque chose qui les distingue, à savoir, des propriétés historiques différentes. En effet, à t_0 Dion instancie la propriété d'avoir deux jambes, ce qui n'est pas le cas de Théon. Cette différence historique heurte l'identité présumée entre Dion et Théon. En effet, par le principe de distinction des discernables, nous devrions alors admettre que Dion n'est pas identique à Théon. Dans ce cas, nous n'aurions pas affaire à un individu compté deux fois, mais bien à deux individus comptés chacun une fois. Étant donné le caractère arbitraire de la partie propre choisie sur Dion, nous devrions même admettre plus généralement que tout individu coïncide avec une quantité infinie d'autres individus de la même sorte. Ce qui paraît, pour le moins dire, être une conséquence sinon inacceptable, du moins bien étrange de nos explications de la persistance temporelle des choses.

De fait, si l'on souhaite conserver le principe d'identité des coïncidents matériels, il semble nécessaire de faire disparaître l'un des deux protagonistes de l'histoire comme indiqué sur la figure 1.4. Il semble pourtant *a priori* aussi arbitraire d'éliminer Dion, qui survit à l'amputation, que d'éliminer Théon qui n'est affecté par aucun changement.

L'élimination d'un des protagonistes est néanmoins une solution défendue par certains auteurs. Celle apparemment la solution proposée par Chrysippe, du moins telle qu'elle est rapportée par Philon d'Alexandrie dans *De l'indestructibilité du monde*. Après l'opération, Dion prendrait la place de Théon, le faisant disparaître. Théon ne survivrait pas en plus de Dion, mais serait remplacé par ce dernier. Constatons cependant que l'argument paraît assez déroutant, étant donné que Théon n'a rien perdu durant l'opération et l'on ne voit donc pas pourquoi il ne persisterait pas.

Michael Burke ((BURKE, 1994)) a repris l'idée de la solution de Chrysippe, en accréditant l'argument de la seule persistance de Dion après l'opération. En effet, si Théon était, avant l'opération, seulement la partie propre d'un individu humain, après l'opération il devient un individu humain à part entière. Il semble donc que l'opération sur Dion agisse bien sur Théon. Or Burke rejette l'opinion que Théon puisse devenir un être humain. Il stipule donc les deux principes suivants qui tiennent pour les personnes :

(Principe de maximalité) Le concept de personne est maximum, autrement dit, des parties propres d'une personne ne sont pas elles-mêmes des personnes.

(Principe d'essentialité) Les personnes sont essentiellement des personnes et les non-personnes essentiellement des non-personnes.

La première affirmation implique que Théon, avant l'opération, n'est pas une personne. La deuxième affirmation implique qu'il est une non-personne de manière essentielle, et donc qu'il n'aurait pu souffrir un changement faisant de lui une personne. C'est donc que Théon n'a pas survécu au changement occasionné par l'opération.

On peut aussi contester l'existence d'objets définis à partir de parties arbitrairement détachées d'objet matériels. Selon Peter Van Inwagen, le concept d'amas organique, c'est-à-dire de partie matérielle propre d'un organisme n'a tout simplement pas d'occurrence réelle. Il n'existe rien qui soit un Dion-sans-jambe-gauche. Autrement dit, la solution d'Inwagen consiste à réfuter l'existence de Théon - mais pas celle de Dion - en rejetant le principe des parties arbitrairement détachées qui peut s'énoncer ainsi :

(Principe de parties arbitrairement détachée) Pour tout objet matériel x et pour R la région matérielle occupée par x à l'instant t, si R^- est une sous-région quelconque de R alors il existe un objet matériel qui occupe cette sous-région à l'instant t.

Mais une telle manière de procéder est une entrave aux principes méréologiques qui accordent la liberté de considérer toute partie d'un objet comme un objet. Elle semble, de ce point de vue, complètement arbitraire.

Quelle autre solution reste-t-il pour nos décomptes ? On peut se demander s'il est possible que des choses numériquement différentes avant une certaine date deviennent une seule chose après cette date. Autrement dit, peut-on admettre l'idée de fusions entre deux individus, qui passeraient ainsi, après une certaine date, de deux entités à une unique entité ? Constatons que l'hypothèse d'une fusion heurte la conception classique de l'identité logique. En effet, si Dion et Théon sont identiques après t_1, il faut aussi qu'ils aient été identiques avant t_0, du moins, s'ils sont bien la même sorte de chose [14]. Mais comment ces deux individus pourraient-ils relever de sortes différentes s'ils sont constitués exactement des mêmes constituants après t_1 ? Si Théon est un homme après t_1, il n'y a aucune raison de le croire autre chose avant cet instant. Ainsi, l'identité de Dion et Théon devrait s'étendre sur tout le segment temporel durant lesquels ils persistent. En effet, la transitivité de l'identité implique que si Théon-à-t_1 est identique à Dion-à-t_1, alors, puisque Théon-à-t_1 est identique à Théon-à-t_0, et que Dion-à-t_1 est identique à Dion-à-t_0, par transitivité de l'identité, nous aurions aussi que Théon-à-t_0 est identique à Dion-à-t_0. L'identité présumée entre Théon-à-t_0 et Dion-à-t_0 ne s'accorde pourtant pas avec la différence historique que nous avons relevée. Une conséquence du principe relativisé de diversité des discernables est donc qu'une discernabilité provisoire implique la différence perpétuelle des individus qui en sont l'objet. De ce point de vue, il n'est donc pas possible que deux entités de la même sorte puissent devenir une seule entité de cette même sorte. Notons aussi qu'un raisonnement symétrique, appliqué aux cas des fissions, montre qu'il est tout aussi impossible pour une chose de devenir deux choses différentes. Le phénomène de la division cellulaire et la capacité de régénération des ombilics coupés en deux pourraient néanmoins paraître heurter cette constatation. La duplication de deux cellules devrait impliquer qu'il n'y avait pas une, mais deux cellules avant la méiose ou avant la mitose.

Une autre solution, déjà évoquée plus haut, nihiliste et développée par Peter Unger ((UNGER, 1979a)) consiste à faire disparaître les deux individus. L'identité ne s'appliquerait qu'aux simples, autrement dit, aux entités n'ayant pas de parties propres. Dion et Théon étant des composés, ils ne relèveraient pas d'authentiques entités, ce qui lèverait le paradoxe par un dénombrement nul. Nous n'aurions alors affaire ni à un, ni à deux individus mais à un compte nul : il n'y aurait ici aucun individu. Nous serions conduits à éliminer les organismes de la même manière que les artefacts hors de l'ontologie.

On peut enfin, à l'instar de Roderick Chisholm ((CHISHOLM, 1976)), admettre la thèse dite de l'essentialisme méréologique, qui affirme que chaque partie d'un individu est essentielle à son identité. Sous cette perspective, Dion n'a pas survécu à son amputation : il ne subsiste, après l'opération, qu'un seul individu, à savoir Théon. Cette manière de concevoir les individus comme ne

[14] La raison de la précaution de reconnaître des choses de la même sorte sera développée un plus loin dans la section de ce chapitre consacrée à la dépendance sortale de l'identité.

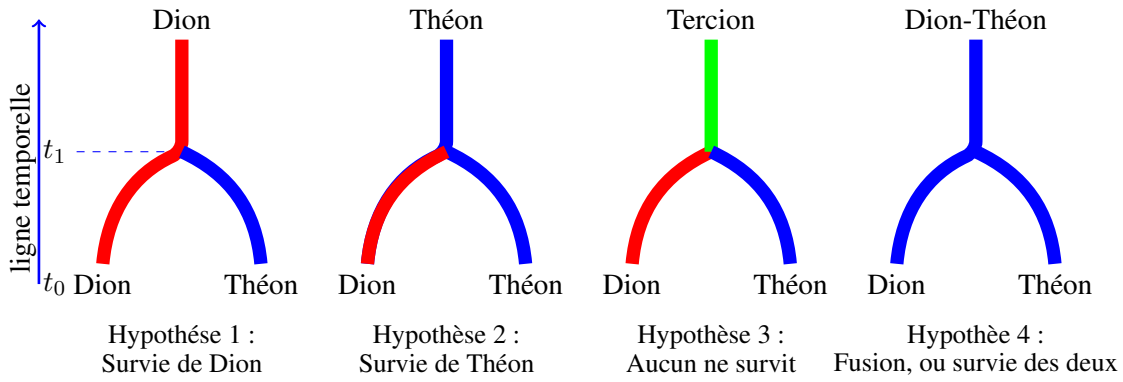


FIGURE 1.4 : Quatre hypothèses au paradoxe de Chrysippe

résistant pas à la moindre perte de matière fait d'eux des entités évanescentes. Puisque la moindre particule de leur constitution leur est essentielle, ils deviennent sans cesse d'autres individus. L'idée d'identité temporelle, c'est-à-dire la persistance d'une et même chose à travers le temps semble dès lors complètement vidée de sa substance.

Paradoxes de métamorphoses

Acceptant la coïncidence d'entités stipulée par David Wiggins, Mark Johnston ((JOHNSTON, 1992)), Judith Thomson ((THOMSON, 1983)) et Lynne Rudder Baker ((BAKER, 2000)) soutiennent que la résolution du paradoxe des statues passe par la distinction entre l'identité et la constitution matérielle. Si deux choses de la même sorte ne peuvent pas occuper simultanément le même emplacement, en revanche, ces philosophes concèdent que rien n'empêche la coïncidence spatio-temporelle de deux choses *de sortes différentes* comme le sont, d'après eux, Goliath et Lumpl. C'est même valable pour des organismes, affirment-ils. Dion et de Théon ne sont pas des individus de la même sorte, Dion est un homme, alors que Théon est quelque chose comme un amas organique, en tout cas, il n'est pas un homme. Il n'y a donc pas de véritable paradoxe à ce qu'ils coïncident. Dans ce cas, il est faux d'affirmer que Goliath est Lumpl. Ce n'est en réalité qu'une façon relâchée de dire que « Goliath possède la même constitution que Lumpl ». L'identité est donc désolidarisée de la constitution matérielle, et on doit alors distinguer deux emplois de l'auxiliaire « être », dont l'un ressortit à l'identité (comme dans « Hesperus est Phosphorus ») et l'autre à la constitution (comme dans « la statue est en argile »).

Le sortaliste (j'appellerai dorénavant ainsi le partisan de la thèse de dépendance sortale de l'identité) considère que sans un critère d'identité, rien ne permettrait de reconnaître un même objet entre deux instants. En outre, le sortaliste souligne que sans critère d'identité, rien n'empêcherait de légitimer les métamorphoses les plus farfelues, car plus rien ne saurait brider les changements que peuvent subir les choses. En assemblant des parties temporelles de manière judicieuse, nous pourrions avoir un objet qui serait un chat jusqu'au 21 septembre 2020 et qui serait un grille pain après cette date.

Comme nous l'avons dit, selon la thèse de dépendance sortale de l'identité défendue par Wiggins, une identité numérique est toujours associée à une identité sortale, entendue comme identité de substance, au sens aristotélicien du terme. Or Wiggins distingue entre les deux thèse de dépendance sortale suivantes :

- (DS1) $(\forall x)(\forall t)\Big((x \text{ existe à } t) \rightarrow (\exists g)(g(x) \text{ à } t)\Big)$
- (DS2) $(\forall x)(\exists g)(\forall t)\Big((x \text{ existe à } t) \rightarrow (g(x) \text{ à } t)\Big)$

La première thèse affirme que pour tout ce qui existe, il y a, à tout instant de l'existence de cette chose, un sortal sous lequel tombe cette chose. Mais cette thèse ne garantit pas que la chose ne puisse relever d'un seul et même concept privilégié tout au long de son existence. Aussi, Wiggins prend-t-il parti pour la thèse (DS2), qui stipule qu'il existe un concept sous lequel cette chose tombe de manière concomitante à son entière carrière temporelle. Une chose x appartenant à une catégorie spécifique F demeure numériquement la même si et seulement si elle demeure de la même sorte F, où F réfère à une certaine catégorie spécifique comme « homme » ou « cancrelat ». Un changement de catégorie implique alors inévitablement une rupture d'identité. Il y a peut-être, au petit matin, un insecte là où il y avait, au coucher, un homme, mais on ne doit pas inférer d'identité entre les deux individus. Un vase demeure le même vase à condition qu'il ne devienne pas un tas de débris, sans quoi il ne serait plus un vase. Un homme n'est pas identique à son cadavre. Aussi, l'identité numérique d'une chose est-elle dépendante d'un concept F de telle manière que l'entière carrière de cette chose soit recouverte par ce concept spécifique. De cette façon, Wiggins distingue d'une part les concepts de substance (comme homme) qui sont concomitants à l'entière carrière des entités qu'elle subsument, et d'autre part les concepts de phases (comme enfant) qui ne le sont pas :

> All phase-sortals are of their very nature qualifications or, [...], restrictions of underlying more general sortals. Boy is definable as human being that is male and biologically immature, and so on. A distinction between substance-sortals and restricted or phase-sortals might be based on the test whether 'x is no longer f' entails 'x is no longer' (or, 'for all f, x is no longer f'). (WIGGINS, 1967, p. 30)

De ce point de vue, découvrir les conditions d'identité d'une chose, c'est enquêter sur la nature de concepts, et déterminer ceux, plus fondamentaux, de substance.

Il reste donc à savoir ce qui est entendu par *sorte* pour indiquer la persistance d'une chose. Comparons par exemple les termes : « noir » et « stylo ». Le premier, adjectival, peut être utilisé afin d'attribuer une qualité à une chose. Il est essentiellement descriptif. Son application ressortit de conditions qui stipulent ce qu'une chose doit satisfaire pour être prédiquée ainsi. Mais ces conditions épargnent complètement la relation d'identité, et n'exigent pas leur persistance pour assurer celle de l'objet : « noir » ne prescrit pas les conditions que doit remplir l'objet qui s'y applique pour continuer à être. Par exemple, un stylo pourrait perdre sa couleur noire par une exposition trop longue à l'ensoleillement, il serait toujours le même stylo. Autrement dit, ce genre de terme est neutre quand aux conditions de persistance. En revanche, ce n'est pas le cas du deuxième terme : appartenir à la classe des « stylos » revêt des conditions de persistance qui sont

données dans cette qualification. Un stylo ne subsiste pas à l'opération consistant à déloger sa réserve d'encre et sa mine. Au mieux, il sera devenu une sarbacane improvisée, autrement dit, un tout autre objet. Mais il aura perdu son rôle de stylo. Les termes sortaux sont donc moins utilisés pour décrire des objets, mais plutôt pour en décrire les conditions de persistance. Par conséquent, la différence entre adjectivaux et sortaux se tient entre la question « comment est-ce ? » (noir, avec une mine, etc.) et la question « qu'est-ce que c'est ? » (un artefact, un instrument d'écriture, un stylo, etc.) Bien sûr, en affirmant « ceci est un stylo », l'on indique inévitablement des descriptions comme avoir une mine et une réserve d'encre, servir à écrire, etc. De surcroît, ces termes prescrivent également les critères de persistance des objets auxquels ils sont attribués, à la fois entre objets différents appartenant à la même classe de choses, mais aussi entre leurs différentes occurrences temporelles. Bien sûr, toute la question est de déterminer si cette partition est légitime.

La spécificité implique la conservation de certaines propriétés. Pour rester le *même* stylo, il faut rester un *stylo*. Si la comptabilité entre le changement et l'identité implique que ce qui change voit sa constitution altérée, cette altération ne doit pas porter préjudice aux conditions de persistance de l'individu considéré sous la sorte de chose qu'il est. Contrairement aux réprobations exprimées par l'élève à qui on l'a confisqué, le stylo évidé n'est plus un stylo, mais bien une sarbacane. Si une tasse ébréchée reste la même tasse, en revanche, que dire d'une tasse dont l'anse est cassée ? S'agit-il encore d'une *tasse* ? Posséder une anse est essentiel à la tasse, autrement, il s'agit d'un bol, d'un gobelet, voire d'un verre. La tasse démunie de son anse n'est plus la *même* tasse puisqu'il ne s'agit tout simplement plus d'une *tasse*.

Il est sans doute difficile d'admettre cependant qu'une statue est d'une nature différente de l'amas d'argile qui la constitue, tout comme il semble difficile d'admettre que Dion soit sortalement différent de Théon. Car de fait, chacun des termes de ces deux couples manifeste exactement les mêmes propriétés que l'autre. La réponse de Wiggins ne répond pas entièrement au paradoxe de Tibbles et Tib, car l'entorse au principe d'exclusion sortale qui affirme que deux choses de la même sorte ne peuvent coïncider n'est pas véritablement levée. Si l'on distingue constitution, coïncidence et identité, il faut alors interroger les principes suivants :

(Coïncidence des identiques) Si une chose x est identique à une chose y, alors, à tout instant t de leur carrière, x et y coïncident à t.

Ce premier principe n'est qu'une conséquence du principe d'indiscernabilité des identiques : les propriétés spatio-temporelles attribuables à x sont également attribuables à y, et respectivement. Il serait certainement néfaste à nos réflexions de remettre en question ce principe. En revanche, la condition nécessaire semble plus discutable. Nous pourrions la décliner dans deux versions, faible et forte :

(Identité des coïncidents - version faible) Si une chose x et une chose y coïncident à tout moment t de leur carrière, alors x est identique à y.

(Identité des coïncidents - version forte) Si une chose x et une chose y coïncident à un moment t de leur carrière, alors x est identique à y.

Comme le résume Stéphane Ferret dans *Le bateau de Thésée* (FERRET, 1996), la position de Wiggins consiste ainsi à stipuler deux types de changements. Là où les premiers n'affectent pas l'existence d'une chose, qui peut ainsi continuer à demeurer la même, les changements du second type rompent l'identité en affectant directement leur existence. La sorte à laquelle appartient une chose constitue la limite infranchissable d'altération. Tout changement que subit un particulier est circonscrit par la catégorie sortale qui le subsume. Mais que signifie au juste continuer à exister ? Perdre l'existence, ce n'est pas perdre des propriétés. Un cadavre est encore un corps, même s'il n'est plus une personne. Selon Wiggins, exister, c'est, pour un particulier donné, continuer à être le même F tout au long de sa carrière, où F désigne un concept sortal qui subsume ce particulier. L'existence est donc inféodée au fait de tomber sous un certain concept, mais vu comme une entité réelle. De ce point de vue essentialiste, qu'on peut qualifier de continuiste, c'est l'essence d'une chose qui détermine son existence. Exister, c'est continuer à être la même substance. La sorte à laquelle appartient une chose détermine donc l'existence de cette chose. Cette perspective est donc contradictoire à celle de Locke, qui écrit au contraire que « c'est l'existence elle-même qui détermine un être, quelle qu'en soit la sorte, à un temps et à une place particuliers, incommunicable à deux êtres du même genre » ((LOCKE, 1998b, p. II, xxvii, 3)). Sous la perspective continuiste, la persistance d'un individu durant un certain laps de temps est conditionnée par la nécessité qu'un certain terme général soit nécessairement vrai de cette chose durant toute sa carrière, et que le terme en question soit plus spécifique que les notions d'entité et de chose.

La question est donc de savoir si un individu tombant sous un concept de substance est *nécessairement* tributaire du concept en question tout au long de son existence. Considérons l'assertion suivante : « un matin, au sortir d'un rêve agité, Grégoire Samsa s'éveilla transformé dans son lit en une véritable vermine. ». Le lecteur aura reconnu le célèbre récit de Kafka décrivant la métamorphose d'un commis de commerce en insecte. Certains critiques littéraires rangent la *Métamorphose* dans les récits fantastiques de doubles [15]. Pour quelle raison ? En quoi la métamorphose a-t-elle trait au double ?

Constatons d'abord qu'elle relève d'une contradiction logique. Dans le récit, le personnage, transformé, conserve néanmoins certaines prérogatives humaines tout en manifestant d'autres propriétés synchroniquement incompatibles avec celles-là. Le personnage est à la fois humain et non humain, autrement dit, l'auteur nous demande d'admettre à la fois que « a est un F à t_0 » et que « a n'est pas un F à t_0 ».

Comparons cette sollicitation de notre imagination à celles qui servent d'autres récits, où l'identité n'est problématique que de manière diachronique. Dans le cas de la métamorphose du Docteur Jeckyll en Mister Hyde, l'auteur nous demande d'imaginer que « a est un F à t_1 » et que « a n'est pas un F à t_2 ». En ce cas, en quoi au juste s'agit-il encore d'un paradoxe ? Dans le cas de métamorphoses, nous disons qu'un certain individu appartenant à l'espèce F est le même qu'un individu appartenant à une espèce non-F. En quoi cela pourrait poser un problème de double ? Si les choses peuvent changer, nous pourrions accepter qu'une et même chose soit un F à un instant t_1 et un non-F à un instant t_2. Le problème se rapporte en fait, pour le partisan de l'essentialisme, aux relations entretenues entre l'identité temporelle d'une chose et son « identité sortale », c'est-à-dire,

[15] Voir par exemple Kurt Fickert ((FICKERT, 1979, p. 34)) qui écrit que dans la *La Métamorphose*, « Kafka se représente lui-même et son Doppelgänger ».

à l'espèce de chose qu'elle est [16].

Du point de vue de la biologie, on pourrait soulever la question de la métamorphose des chenilles en chrysalides et des chrysalides en papillons adultes. La réponse apportée par Wiggins est que chenilles, chrysalides et papillons adultes ne sont que des concepts de phase, qui tombent eux-mêmes sous le concept substantiel de papillon. En dépit des transformations qu'ils subissent, les papillons ne changent pas de concept substantiels. Ils sont des papillons d'un bout à l'autre de leur carrière temporelle. En revanche, ils passent par diverses phases.

C'est un fait que les métamorphoses touchent certains insectes, poissons et amphibiens, mais qu'elles demeurent étrangères aux mammifères. Or, si le principe de dépendance sortale est métaphysiquement vrai, il doit alors répondre plus généralement à toute situation logiquement possible. Ainsi, celui qui doute de la validité de (DS2) peut être amené à considérer des cas plus extraordinaires de métamorphoses. Dans *Sameness and Substance* (WIGGINS, 1980), David Wiggins mentionne l'histoire de la femme de Loth métamorphosée en statue de sel, dont le récit est exposé dans la *Genèse*. De son côté, dans *Identity Through Time* ((PRICE, 1977)), Marjorie Price envisage le cas d'un chien (Rover) qui, transporté sur Mars, est affecté d'une modification d'ADN d'où résulte une transformation en une masse informe (Clover). Concevoir que la statue est numériquement identique à la femme de Loth ou que Clover est identique à Rover implique de concevoir une identité en dépit qu'entre les deux occurrences, nous n'ayons pas de concept spécifique « naturel » recouvrant la carrière complète de l'individu en question. Un partisan réaliste de l'identité sortale est alors amené à choisir parmi les trois options suivantes :

1. Soit nier l'identité entre les deux occurrences avant et après transformation.
2. Soit globalement rejeter de telles histoires au motif qu'elles sont absurdes.
3. Soit constituer un concept sur mesure afin de recouvrir la carrière des individus qui sont identifiés.

Il appert que nous comprenons de nombreuses histoires de métamorphoses, et qu'il est ainsi difficilement acceptable de rejeter l'identification entre deux occurrences. Aussi, il peut sembler raisonnable de rejeter la première option. En outre, de telles histoires ne paraissent pas être logiquement contradictoires, tout au plus physiquement ou biologiquement contradictoires. On devrait donc aussi rejeter la deuxième option. Tout ceci pourrait donc inciter le réaliste à conserver la troisième option. Mais puisque l'on peut forger de telles histoires d'identification avec tout type de choses - un cendrier transformé en vallée, une perdrix transformée en nombre premier - le réaliste serait amené à accepter l'idée que les prédicats qu'il emploie sont artificiels et arbitraires. Artificiels parce que relatifs à des catégories spécifiques constituées sur mesure, et arbitraires parce que l'on ne verrait pas bien au juste ce qui pourrait encore circonscrire les catégories ontologiques dont le réaliste des universaux prétend rendre compte, puisque le concept de « perdrix-nombre premier »

[16] Comme dans le cas de l'identité qualitative, l'identité sortale peut être entendue de deux façons différentes. Soit littéralement comme une identité d'espèce entre des particuliers donnés, ce qui présuppose l'existence de propriétés ou d'universaux qui subsument de tels particuliers (réalisme des universaux), soit au sens de classes de particuliers regroupés soit de manière conventionnelle (nominalisme de classe) soit en raison d'une ressemblance (nominalisme de la ressemblance).

aurait finalement autant de légitimité métaphysique que celui d'insecte ou d'homme. Autrement dit, une telle perspective serait certainement une forme d'abdication du réalisme en faveur du nominalisme linguistique : nos concepts ne seraient que le fruit d'élaborations linguistiques, mais non des réalités métaphysiques.

Comme le reconnaît Wiggins lui-même ((WIGGINS, 1980, p. 65)), si nous étions libres de produire des concepts substantiels pour expliquer la persistance temporelle des choses, nous serions également libre d'en produire pour représenter le fait qu'une chose n'ait pas persisté après une certaine transformation. Mais du point vue de Wiggins, ce n'est justement pas une bonne manière de voir les choses, parce que nous devrions, selon lui, conserver une distinction importante entre le fait, pour une chose, de *changer*, et le fait d'être *remplacée*. Aussi, prise au sérieux, l'histoire de la femme Loth ou de Rover nous conduirait à la recherche quasi-scientifique d'une authentique substance liant l'identité de la femme à celle de la statue :

> If we took the Lot story seriously, if we treated it as if it were a historical narrative, then, enamoured as we are of the change/replacement and true/false distinctions, we should need to interrogate the narrative very closely, and press all sorts of questions about what sort of a being Lot's wife really was. We should need to think of that question as one to which an answer could be discovered 'in spite of ourselves'. (WIGGINS, 1980, p. 67)

L'histoire de la femme de Loth impliquerait un monde où des substances naturelles auraient la propriété d'être des femmes pendant une certaine durée, et d'être ensuite des statues de sel. L'histoire de Rover et Clover inciterait à considérer une catégorie sortale à laquelle appartiendrait et le chien et la masse informe, qui ne seraient que des concepts de phases. On peut se demander si le concept d'organisme ferait ici l'affaire : Rover et Clover seraient le même organisme, mais pas le même chien. Cela signifie alors que la catégorie de chien n'est pas réellement substantielle, mais qu'en réalité il s'agit qu'un concept de phase. Dans ce cas, on peut se demander s'il n'en est pas de même pour tout autre catégorie « spécifique », puisqu'il est toujours possible d'inventer une histoire logiquement valide dans laquelle un concept de substance s'avérerait n'être au fond qu'un concept de phase. Dès lors, la distinction entre les concepts de phases et de substances devient complètement artificielle. Pour répondre à cette difficulté de l'essentialisme, Wiggins objecte qu'il ne s'agit ici que d'histoires fantastiques, c'est-à-dire d'histoires qui ne respectent pas les catégories métaphysiques du monde actuel, monde dans lequel les hommes ne se transforment jamais en statues de sel. Wiggins évacue ainsi le problème, sans doute un peu vite, en objectant que les histoires ne relèvent pas toutes de descriptions de mondes possibles. Mais il semble qu'il fasse ici une confusion entre les mondes logiquement possibles, dans lequel les métamorphoses sont possibles, et les mondes nomologiquement possibles dans lesquels ces métamorphoses sont impossibles. Si les histoires fantastiques impliquent la stipulation de substances tout aussi fantastiques, elles ne sont néanmoins pas nécessairement absurdes.

1.3.3 Théories endurantiste et perdurantiste

Nous n'avons pas trouvé d'explication satisfaisante aux paradoxes de l'identité objectuelle. D'où viennent ces divers problèmes ? Rappelons que, afin d'éviter un conflit entre identité numérique et changement qualitatif, nous avons pris soin de relativiser temporellement le principe d'indiscer-

nabilité des identiques. Mais nous n'avons cependant pas encore précisé ce qui doit au juste être relativisé dans l'usage du principe. Or il n'y a, semble-t-il, que trois possibilités : soit il s'agit de l'identité elle-même, soit il s'agit des propriétés, soit il s'agit des objets. Considérons chacune des hypothèses et voyons comment nous échappons au paradoxe des coïncidents temporaires.

Qu'est-ce qui, au juste, doit être temporellement relativisé dans l'usage du principe d'indiscernabilité des identiques ? Admettons qu'il s'agisse de l'identité. Cette solution a été non seulement envisagée, mais plus encore défendue [17] par George Myro ((MYRO, 1997)). Selon Myro, il y a deux genres de propriétés, celles qui ne sont pas temporellement dépendantes, comme être rond ou être à 1 km de la tour Eiffel, et celles qui le sont, comme être rond ce soir et être sur la Tour Eiffel à midi. Pour Myro, la loi de Leibniz ne doit être appliquée qu'au premier type de propriétés. Reprenons alors le paradoxe de la statue et du bloc d'argile. Selon la théorie de l'identité temporellement relativisée, la statue est bien identique au bloc de marbre. La destruction de la statue à t_2 relève d'une propriété temporelle. Elle ne nécessite donc pas d'être aussi attribuée au bloc de marbre. Il n'y a alors guère de paradoxe : on préserve l'identité du bloc et de la statue.

Cependant, il faut reconnaître le coût de cette solution. La théorie de l'identité temporellement relativisée laisse la possibilité qu'une chose a soit identique à une chose b au moment t_1, puis que a devienne numériquement différente de b au moment t_2. Sous cette perspective, l'identité est vue non comme une relation permanente, mais comme une relation occasionnelle. Il serait toutefois étonnant, voire absurde, qu'un principe logique soit valable à un moment donné et qu'il ne le soit plus à un autre. En outre, une telle conception rejoint la thèse d'identité relative en réduisant considérablement l'efficacité du principe d'indiscernabilité des identiques, qui semble bien trop important pour qu'on cède ainsi à ce genre d'argumentations. Il semble donc préférable de ne pas relativiser l'identité et d'envisager une autre voie. On pourrait toujours échapper au paradoxe à la manière du bateau de Thésée en affirmant qu'il s'agit ici d'artefacts qui dépendent de nos intentions.

Théorie de l'endurance

Si l'identité n'est pas relative, considérons alors l'hypothèse que ce sont les propriétés qui doivent être temporellement relativisées. Comment expliquer qu'un individu puisse successivement avoir, assis, une forme courbe, puis, debout, une forme droite ? Comment attribuer des propriétés contradictoires au même individu sans proférer d'absurdités ? Bien sûr, comme nous l'avons précisé, le principe d'indiscernabilité n'exige pas qu'un individu possède de manière synchronique des propriétés contradictoires, mais seulement qu'il les ait de manière diachronique. Mais il reste encore à savoir ce que signifie pour un individu d'exemplifier une propriété *à un moment donné*.

David Lewis ((LEWIS, 2007, pp. 202-205)) désigne cette question sous le nom du problème des *propriétés intrinsèques temporaires*. De manière générale, nous dirons qu'un individu est dit posséder une propriété *intrinsèque* [18] si cette propriété ne dépend d'aucun autre objet que cet

[17] Remarquons que la thèse de Myro relève d'un cas particulier de relativisation de l'identité, dont la thèse plus générale a été défendue avant Myro par Peter Geach. Nous reviendrons amplement sur la théorie de l'identité relative par la suite.

[18] Définir la notion de propriété intrinsèque n'est pas une affaire simple. Lewis propose trois définitions différentes :

- « A sentence or statement or proposition that ascribes intrinsic properties to something is entirely about that

individu. Autrement dit, il s'agit d'une propriété que cet individu possède par lui-même et qu'il peut posséder indépendamment de tout ce qui n'est pas lui :

(Propriété intrinsèque) Une propriété P est intrinsèque à un individu i si l'exemplification de P par i ne dépend d'aucune autre existence que celle de i et de ses parties propres [19].

À l'inverse, une propriété extrinsèque est une propriété qui dépend d'une autre chose que l'individu ciblé :

(Propriété extrinsèque) Une propriété P est dite extrinsèque à un individu i si l'exemplification de P par i implique l'existence de quelque chose différent de i et de ses parties propres.

Par exemple, avoir une masse de 70 kg est une propriété intrinsèque d'un certain objet, alors qu'avoir un poids voisinant 630 N est une propriété extrinsèque [20]. La différence de sens entre intrinsèque et extrinsèque peut aussi être exprimée de manière modale : si, en dehors de l'objet visé, tous les autres objets du monde venaient à disparaître, cet objet conserverait toutes ses propriétés intrinsèques, et aucune de ses propriétés extrinsèques.

Moore (MOORE, 1922) soutient que les propriétés intrinsèques ne jouent pas toutes le même rôle à l'égard de l'objet qui les instancie. Les propriétés intrinsèques naturelles, appelées aussi propriétés purement qualitatives d'un objet, sont plus rares que ses propriétés intrinsèques. Être identique à Socrate est sans doute une propriété intrinsèque de Socrate, mais ce n'est pas une propriété naturelle de Socrate, il s'agit d'une propriété dite « impure », et plus particulièrement,

---

thing; whereas an ascription of extrinsic properties to something is not entirely about that thing, though it may well be about some larger whole which includes that thing as part. » (LEWIS, 1983a, p. 197)

- « A thing has its intrinsic properties in virtue of the way that thing itself, and nothing else, is. Not so for extrinsic properties, though a thing may well have these in virtue of the way some larger whole is... » (LEWIS, 1983a, p. 197)
- « If something has an intrinsic property, then so does any perfect duplicate of that thing; whereas duplicates situated in different surroundings will differ in their extrinsic properties. » (LEWIS, 1983a, p. 197)

Selon les auteurs, la notion d'intrinsécalité amalgame plus ou moins, ou au contraire se distingue nettement des notions d'essence, de qualités non relationnelles, de qualités internes, de propriétés naturelles. Pour un aperçu des diverses conceptions de l'intrinsécalité, voir (MARSHALL & WEATHERSON, 2018). Pour ce qui nous concerne, nous nous contenterons de la définition donnée ci-après, qui suffit largement à notre exposé.

[19] On rappelle qu'une partie propre d'un objet est une partie quelconque distincte de cet objet.

[20] Notons qu'il ne faut pas amalgamer les propriétés intrinsèques/extrinsèques et les prédicats monadiques/relationnels. Le deuxième couple de termes a une signification purement syntaxique ; il désigne des prédicats respectivement à une et à plusieurs places. Or un prédicat monadique n'est pas nécessairement une propriété intrinsèque, et un prédicat relationnel n'est pas nécessairement une propriété extrinsèque. Par exemple, la relation qu'une chose entretient à ses parties n'est pas extrinsèque, mais elle est pourtant relationnelle. Plus généralement, on peut trouver des exemples de propriétés à la fois relationnelles et à la fois intrinsèques. On peut citer en exemple la relation d'aimer quelqu'un, utilisée de manière réflexive lorsque cet amour renvoie de façon narcissique à sa propre personne. On a bien une prédication relationnelle concernant une relation intrinsèque. On remarquera en outre que les propriétés intrinsèques semblent être autant qualitatives que quantitatives puisque s'y trouvent des grandeurs, des degrés et des intensités. Mais cela peut toutefois poser problème : est-ce que la masse d'un objet, sa taille, ne dépendent pas de la totalité de l'univers ? Pour cette raison, les auteurs préfèrent généralement restreindre les propriétés intrinsèques aux seules propriétés qualitatives. Néanmoins, cette question peut aussi poser au niveau des qualités : n'y a-t-il pas, au fond, que des propriétés extrinsèques ?

d'une propriété haeccéitique en ce qu'elle fait relève d'une identité non qualitative. Robert Adams a proposé deux façons de caractériser les propriétés pures. La première, sémantique, consiste à dire qu'une propriété qualitative est pure [21] si et seulement si elle peut être exprimée dans un langage ne faisant pas mention de termes référentiels tels que les noms propres, les expressions indexicales, les descriptions définies, les adjectifs ou les verbes dérivés de noms propres comme « kafkaïen » ou « boycotter ». La deuxième définition est métaphysique. Elle consiste à décrire les propriétés purement qualitatives comme des propriétés 1) excluant les propriétés d'*haecceité*, 2) qui ne sont pas liées à un ou plusieurs particuliers et 3) qui ne sont pas liées à un ensemble défini de façon extensionnelle par des individus qui en seraient membres. On dira qu'une propriété qualitative est impure [22] si elle n'est pas purement qualitative. Ainsi, « être rond » et « être à moins d'un mètre d'une fenêtre » expriment respectivement des propriétés pures intrinsèques et extrinsèques. En revanche, « habiter au 221B Baker Street » et « s'être passé durant la bataille de Marathon » expriment des propriétés qualitatives impures en ce qu'elles font référence à des particuliers.

Relativiser temporellement les propriétés exige dès lors de les considérer (quasiment) toutes comme des propriétés extrinsèques, à savoir, comme des propriétés relatives à des moments. Envisager que les propriétés sont relatives au temps, c'est dire par exemple que Platon était « imberbe-enfant » et « barbu-adulte ». Lewis objecte qu'envisager qu'il n'y a que des propriétés extrinsèques, corrélées au temps, est absolument contre-intuitif. Lewis souligne en effet qu'être rond n'est pas une propriété extrinsèque, dépendante du temps, mais au contraire, une propriété intrinsèque : « si nous savons ce qu'est une forme, nous savons que c'est une propriété et non une relation » ((LEWIS, 2007, p. 312)). Certes, le changement occasionne des variations extrinsèques, mais il est également vrai qu'il existe des altérations purement intrinsèques. Or, en rabattant les propriétés sur les relations, les seules propriétés intrinsèques qui subsistent sont celles qui ne dépendent pas du temps, comme l'identité à soi ou l'appartenance spécifique.

Lewis désigne la conception de la persistance qui considère que les propriétés sont toujours temporellement relatives sous le nom d'« endurantisme ». Selon la théorie endurantiste, un individu est intégralement présent à chaque instant de son existence. Lewis rapproche les individus endurants des universaux :

> L'endurance correspond à la manière dont un universel, si une telle chose existe, serait entièrement présent là où et au moment où il est instancié. (LEWIS, 2007, p. 310)

De ce point de vue, les seules propriétés intrinsèques que possède un objet sont celles qui ne sont pas temporaires, à savoir, des propriétés permanentes comme le fait d'être un objet ou l'identité à soi. Ce n'est donc pas l'objet lui-même qui change, mais ses propriétés. Or Lewis fait remarquer que le changement d'un objet devrait être alimenté par la transition de ses propriétés intrinsèques. L'endurantisme conduit donc, en quelque sorte, à nier le changement en dénaturant les propriétés. La propriété d'être imberbe, remarque Lewis, ne devrait pourtant pas dépendre d'un instant particulier (comme par exemple être imberbe-enfant). Or, en relativisant les propriétés au temps, l'endurantisme conduit à l'idée que, de Platon enfant au Platon adulte, rien ne change, puisque l'un a les mêmes relations que l'autre aux instants considérés : Platon à l'âge adulte a autant la propriété d'être imberbe-enfant que Platon à l'âge d'enfant possède celle d'être barbu-adulte. Or,

[21] On dira aussi, de manière équivalente, « purement qualitative ».

[22] Ou qu'elle est « non purement qualitative ».

pour parler d'un quelconque changement, il est nécessaire qu'une certaine propriété intrinsèque à l'enfant ne soit plus possédée par l'adulte.

Théorie de la perdurance

La théorie de l'endurance implique le recouvrement, c'est-à-dire l'idée qu'à deux moments différents du temps où existe un individu, il existe nécessairement une partie commune à ses deux occurrences temporelles, à savoir, l'individu en question. Si on veut s'extraire des difficultés soulevées par cette perspective, il faut alors concéder que ce ne sont pas les propriétés des individus qui doivent être relativisées au temps, mais leurs parties constitutives. C'est la solution que choisit Lewis, sous le nom de *perdurantisme*. Elle consiste à admettre que les individus sont composés de parties temporelles tout comme ils sont composés de parties spatiales. Il faut alors admettre que les choses ne sont jamais présentes intégralement à un moment particulier du temps, mais uniquement par phases temporelles. Ainsi, non seulement un objet quelconque ne peut être présent à différents endroits au même moment, mais de plus, il ne peut être qu'en partie présent au même endroit. Cela revient finalement à analyser la persistance d'une chose de manière quadri-dimensionnelle [23]. La temporalité est vue comme une extension aux dimensions temporelles. Les individus ne sont pas seulement étendus dans l'espace, mais également dans le temps. Lewis substitue à l'image héraclitéenne des individus charriés dans le mouvement du fleuve *Devenir* l'image d'une route traversant différents villages. Cette route est telle que chacune de ses parties est présente à l'endroit où elle est. En revanche la route n'est jamais entièrement présente dans aucune de ses parties. Ainsi, contrairement au point de vue endurantiste, les parties d'un individu ne sont pas toutes présentes à un moment donné :

> Une chose perdure si et seulement si elle persiste en ayant différentes parties temporelles, ou stades, à différents moments du temps, bien qu'aucune de ses parties ne soit entièrement présente à plus d'un seul moment du temps, tandis qu'elle endure si et seulement si elle persiste en étant entièrement présente à plus d'un moment du temps (LEWIS, 2007, p. 309)

On peut se représenter l'existence des choses comme des vers qui traversent le temps (*temporal worms*), de la même manière qu'on se représente habituellement les choses comme des sommes de parties spatiales. Le temps n'est pas un milieu traversé par un individu tout entier : il est au

[23] On doit à Quine les idées séminales concernant le perdurantisme. Dans *Méthodes de logique* ((W. v. O. QUINE, 1980, p. 237)), Quine parle de « stages » ou encore de « moments-particules » pour désigner des parties temporelles. La conception perdurantiste apparaît aussi chez Nelson Goodman, qui écrit, dans *La structure de l'apparence* que : « Nous n'éprouvons pas le besoin d'hypostasier un noyau sous-jacent d'individualité afin d'expliquer comment un pied et un plateau, qui diffèrent si fortement, peuvent appartenir à une table unique. Et pourtant, si nous considérons la table en des moments différents, on dit parfois qu'il nous faut alors nous demander ce qui persiste à travers ces sections temporellement différentes. La réponse est que l'unité est sous-tendue par ces différents éléments, plutôt qu'elle ne les sous-tend : il se trouve que ces sections sont d'une extension temporelle, et non plus spatiale, moindre que l'objet tout entier, mais elles ont néanmoins avec lui la même relation d'élément à totalité plus large. Et comme auparavant, parce que ces éléments ont certaines caractéristiques et certaines relations entre eux, la totalité qu'ils forment est ce que nous appelons une chose, et plus précisément une table. Nos tables, nos bateaux à vapeur et nos pommes de terre sont des événements dont la dimension spatiale est petite, comparée à leur vaste dimension temporelle [...] Le bateau-à-vapeur pendant une heure est un événement spatialement aussi grand que le bateau mais temporellement plus petit. » (GOODMAN, 2004, p. 124)

contraire une dimension dans laquelle se déploient ses parties. Là où l'endurantiste considère que les choses occupent l'espace en traversant intégralement le temps, le perdurantiste considère qu'elles occupent à la fois l'espace et le temps. De ce point de vue, un individu est un *processus*. Il s'étend dans le temps, plutôt qu'il ne l'habite.

Puisqu'une chose qui persiste dans le temps n'est qu'en partie présente à un moment donné, des propriétés instanciées à un moment peuvent bien différer de celles instanciées à un autre moment tout en concernant néanmoins le même individu. Platon-enfant a la propriété intrinsèque d'être imberbe, Platon-adulte a celle d'être barbu. D'après le principe de distinction des différents, cela induit que Platon-adulte soit différent de Platon-enfant. Mais cette conclusion est assumée par le perdurantiste pour qui la distinction de parties temporelles ne pose pas plus de problème que la distinction des parties spatiales. Cela n'est pas plus étonnant que la constatation que le bras gauche de Platon n'est pas identique à son bras droit. En revanche, si les parties temporelles Platon-enfant et Platon-adulte diffèrent, elles concernent cependant le même individu, à savoir Platon, sans être identiques à lui. Là encore, ce n'est pas plus étonnant que le fait que les bras droit et gauche appartiennent au même individu sans lui être identique. La transition d'une propriété intrinsèque à l'autre est expliquée par le fait que ce sont des parties temporelles différentes du même individu qui instancient ces propriétés contradictoires. Le jugement que la personne qui philosophe au bord de la rivière est la même que celle qui y jouait enfant peut alors se paraphraser de la manière suivante : la personne dont fait partie la phase Platon-adulte est la même que celle dont fait partie la phase Platon-enfant [24].

La théorie de la perdurance permet ainsi de résoudre le paradoxe de Théon et Dion et celui de Lumpl et Goliath. Par exemple, dans le cas du paradoxe de Chrysippe, puisque les choses persistent sans que leurs parties soient toutes présentes au même moment, Dion ne partage en fait, avec Théon, que ses parties post-opératoires. Il y a bien deux individus, qui coïncident spatialement seulement après l'opération. Les parties temporelles qui appartiennent à Dion après l'opération appartiennent aussi à Théon, de la même manière qu'un pont ou qu'une route peut appartenir à deux communes. Mais Dion est cependant distinct de Théon, parce qu'avant l'opération, ils n'ont pas toutes leurs parties temporelles en commun.

Paradoxe de Dishpan et Plastic

La solution passant par la théorie perdurantiste s'applique aussi pour résoudre le paradoxe Lumpl et Goliath. Effet, ces derniers ne partagent pas toutes leurs parties temporelles puisque Lumpl survit à Goliath : la statue et le bloc d'argile qui la constitue ne terminent pas leur carrière ensemble. En revanche, dans le cas d'une coïncidence permanente, l'explication ne semble plus tenir. En effet, le paradoxe de Lumpl et Goliath peut être rendu plus complexe encore par l'hypothèse supplémentaire que les deux objets naissent et disparaissent simultanément, autrement dit que leurs carrières soient parfaitement coextensives. Admettons par exemple, à la suite de Lewis ((LEWIS,

[24] Précisons que pour le présentiste, seuls les objets présents existent. Il n'y a donc pas lieu de parler de propriétés temporellement différentes dans le temps parce que le passé et le futur n'ont aucune sorte d'existence. Ainsi, un objet ne trouve sa réalité et sa consistance que dans le moment présent. À proprement parler, un individu est présentement imberbe ou barbu, mais il n'y a pas de sens à dire qu'il est barbu-adulte alors qu'il était imberbe-enfant, car sa seule existence est celle du moment présent. Il n'y a donc pas de contradiction entre les propriétés, puisque qu'à tout moment, un individu possède telle propriété, mais pas la propriété contradictoire.

2007)), qu'une poubelle appelée « Dishpan », moulée à partir d'un bloc de plastique appelé « Plastic » soit synthétisée par injection instantanée de plastique dans un moule, et détruite ensuite par incinération. Dans ce cas de figure, Dishpan et Plastic posséderaient des carrières exactement identiques. Ainsi, nous n'aurions plus l'occasion de distinguer ces deux choses par d'éventuelles différences historiques. En revanche, il y aurait encore une autre manière de discerner Dishpan de Plastic, celle-ci, modale. En effet, on pourrait encore constater que Dishpan *pourrait* survivre à une petite perte de matière, sans que cela soit le cas de Plastic, sans quoi il ne serait évidemment plus le même bloc de matière.

Pour répondre à cette difficulté, Allan Gibbard ((GIBBARD, 1975)) propose d'abandonner la thèse de nécessité de l'identité, en admettant le principe contraire qu'il y a des identités contingentes. De ce point de vue, il serait vrai que, quoique Dishpan et Plastic soient actuellement identiques, ils auraient pu néanmoins être différents.

De son coté, Lewis soutient que le nom de Dishpan n'est pas référentiellement transparent. En effet, le nom « Dishpan », en tant que nom d'objet, occupe une fonction qui n'est pas uniquement référentielle. Il s'agit aussi d'une corrélation à une contrepartie qui la représente dans d'autres mondes. L'idée de Lewis consiste donc à bloquer l'identité transmondaine : la seule identité est toujours intramondaine, les choses sont *world bounded*. Lewis est alors amené à professer une théorie des doubles, baptisée « théorie des contreparties » que nous examinerons plus en détails par la suite.

1.3.4 L'identité est-elle vague ?

Ostensions, hypostases et quantifications

Revenons sur le bases historiques qui ont déterminé l'orientation actuelle de l'ontologie. Dans l'article « Identité, ostension et hypostase », Quine soutient l'idée qu'un objet est un processus temporel constitué de phases. Examinant le problème d'Héraclite (on ne se baigne pas deux fois dans le même fleuve), Quine affirme en effet qu'« un fleuve est un processus temporel, et les phases fluviales en sont ses parties momentanées. » ((W. V. O. QUINE, 2003a, p. 106)). On peut donc bien se baigner deux fois dans le même fleuve, disons, le Caÿster, mais pas dans les mêmes parties temporelles du fleuve. Vu comme un processus et non comme un porteur de propriétés endurant, il n'est plus besoin de rendre dépendant l'objet de sa constitution matérielle à un moment t quelconque du temps. Sa constitution peut ainsi se voir complètement modifiée sans qu'il ne cesse de s'agir du même objet. Quand nous désignons du doigt un fleuve à deux moments différents du temps, nous ne montrons pas deux phases identiques du fleuve, mais deux phases *apparentées* au même objet. L'ostension associe ainsi l'identité à la parenté de phases, et comme le souligne Quine, « l'attribution d'identité est essentielle car elle fixe la référence de l'ostension » ((W. V. O. QUINE, 2003a, p. 106)). Ainsi, en désignant deux phases temporelles différentes du fleuve, je constitue par cet acte un objet, le fleuve, auquel il me sera encore possible de référer plus tard, les phases étant « passées ».

Si nous désignons certaines parties temporelles qui ne sont pas apparentées au même fleuve, par exemple si nous désignons des phases de la Seine et des phases du Rhône, nous désignons alors un objet encore plus grand que ces fleuves particuliers, un objet spatialement dispersé, que

nous désignons sous le nom générique de fleuve. Quand ce que nous voulons dire de certaines régions spatio-temporelles ne relèvent plus de distinctions entre des parties d'un même objet, « nous simplifions le discours en rendant ses objets aussi grands et peu nombreux que possible - en prenant les diverses surfaces importantes pour des objets singuliers » ((W. V. O. QUINE, 2003a, p. 112)).

Comme le reconnaît Quine, un tel point de vue est fortement rattaché à la conception de Hume considérant que l'idée d'un objet externe découle en vérité d'une erreur d'identification. L'identité provient d'un jugement erroné, elle est déduite de ressemblances qui coexistent entre des parties temporelles qui sont néanmoins des faits *distincts*. Dans le monde de Hume, toute chose est évanescente et disparaissante, rien ne demeure, tout n'est que diversité et multiplicité, rien ne reste identique. Ce sont seulement nos intérêts pratiques qui nous engagent à réifier des phases successives en « objets », c'est-à-dire à référer à certains groupes privilégiés de phases temporelles parmi toutes les associations possibles. En nous référant à certaines associations privilégiées de parties temporelles plutôt qu'à une multitude de parties temporelles distinctes, « le sujet discuté gagne en simplicité formelle à être représenté comme un objet singulier [...] au lieu d'une multiplicité d'objets a, b, etc. se tenant dans une relation de parenté [...] » ((W. V. O. QUINE, 2003a, p. 111)).

De ce point de vue, la présumée singularité des objets ordinaires que sont les tables, les chats et les fleuves serait un simple mirage conceptuel. L'identité d'une chose matérielle n'est pas un fait métaphysique qui se tiendrait entre des parties temporelles différentes, mais un jugement erroné relevant de la pure psychologie. Désigner une relation entre des parties temporelles différentes, c'est en fait inventer « un nouvel objet non momentané qui sert de support à notre affirmation d'identité » ((W. V. O. QUINE, 2003a, p. 107)). D'un tel point de vue, les objets ne sont pas vus comme l'ameublement du monde, ils ne sont pas des choses extérieures données d'emblée par delà nos intentions. Il s'agit au contraire d'invention « humaines, trop humaines ». Sous une telle perspective, les individus sont des créations linguistiques plutôt que des faits.

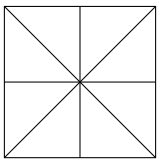
Quand diverses parties temporelles sont suffisamment similaires, l'ostension répétée de ces phases permet, par induction, la création mentale d'un objet qui sert d'explication à la parenté des diverses phases désignées :

> [...] le concept d'identité accomplit une fonction centrale dans la spécification par ostension des objets spatio-temporellement larges. Sans l'identité, n actes d'ostension spécifient simplement jusqu'à n objets, tous d'une extension spatio-temporelle déterminée. Mais si nous affirmons l'identité de l'objet d'une ostension à une autre, nous permettons à nos n ostensions de faire référence au même grand objet, et nous fournissons ainsi à notre interlocuteur un fondement inductif, à partir duquel il peut deviner la portée voulue de l'objet. L'ostension pure plus l'identification apportent, avec l'aide d'un peu d'induction, l'extension spatio-temporelle. (W. V. O. QUINE, 2003a, p. 109)

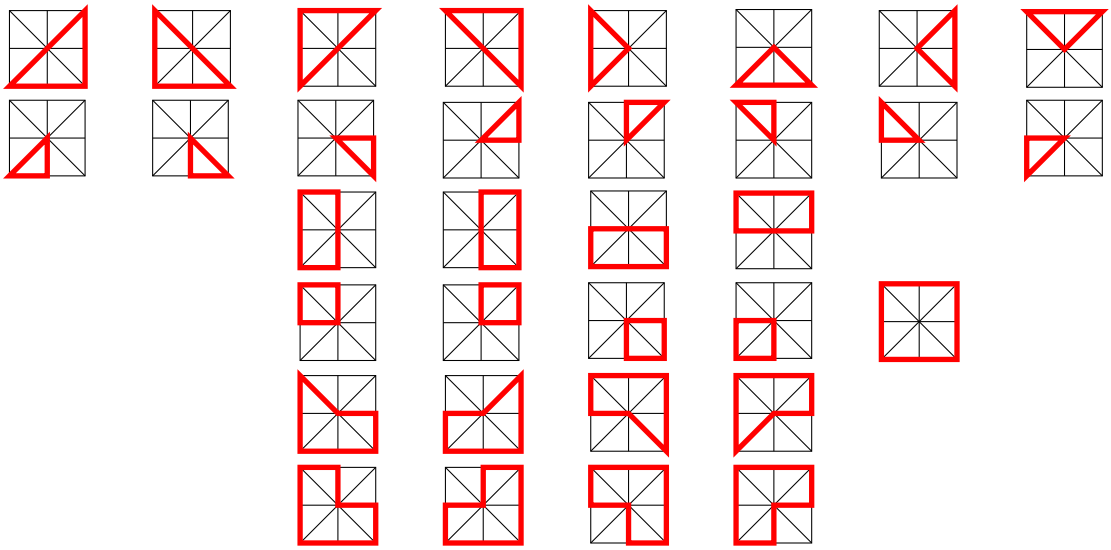
En outre, si l'identité n'est pas une relation factuelle mais résulte d'un processus sélectif d'inductions qui dépendent de facteurs purement psychologiques, l'identité peut alors se dire d'un objet dans certaines circonstances mais pas dans d'autres. C'est alors la notion d'objet elle-même qui porte les conditions, purement épistémiques, de l'identification. Selon que l'on désigne un fleuve à deux moments, on peut en effet ou désigner *deux* parties aqueuses différentes, ou désigner *un seul* objet s'étendant sur une partie spatio-temporelle plus importante, comme par exemple

le fleuve. La même caractérisation semble s'appliquer aussi, souligne Quine, à un terme général concret comme « rouge ». En effet, nous pourrions parler de rouge pour indiquer la relation de parenté entre diverses parties spatio-temporelles suffisamment similaires quant à la rougeur. Nous désignerions par ce terme différentes régions spatio-temporelles, dont la succession induit de considérer un objet relevant d'une plus vaste étendue spatio-temporelle que celles indiquées par chacun de nos gestes. C'est à cette étendue que s'applique l'identité, et non aux différentes parties désignées. Cela conduit Quine à énoncer un principe d'identification des indiscernables : « les objets qu'on ne peut pas distinguer entre eux dans les termes d'un discours donné doivent être réputés identiques pour ce discours ».

Une conséquence de ceci est que nous pouvons mettre sur le même plan un objet singulier comme le Caÿster et un universel comme le rouge, ce qui conduirait à admettre « une égalisation générale des particuliers et des universaux » ((W. V. O. QUINE, 2003a, p. 112)). Il semble qu'il y ait une certaine liberté, pour ne pas dire une gratuité, dans nos choix d'intégration des parties spatio-temporelles : « tout n'est qu'intégration spatio-temporelle » ((W. V. O. QUINE, 2003a, p. 114)). Dans ce cas, peut-on uniformément rabattre les universaux sur les concrets particuliers ? Mais comme le montre Quine, il n'en va pas ainsi avec les termes généraux abstraits. Quine propose ainsi de considérer la figure ci-dessous :



On demande de dénombrer les zones convexes délimités par les segments tracés. De fait, on peut en compter 33, à savoir, 16 zones triangulaires, neuf zones rectangulaires, 4 zones pentagonales et 4 zones hexagonales, comme l'indique l'illustration ci-dessous.



Si nous faisons référence aux formes plutôt qu'aux régions spatiales, nous pouvons admettre qu'il n'y a en fait que 5 formes, à savoir, un triangle, deux rectangles (dont un carré), un pentagone et un hexagone. Or si l'on tente de considérer que l'universel qui semble impliqué par « forme triangulaire » soit toute la région obtenue en assemblant les zones triangulaires, nous sommes conduits à un problème. Car cet assemblage n'est finalement rien d'autre que le grand carré. Ainsi, nous sommes conduits au jugement erroné de l'identité entre forme triangulaire et forme carrée. Quine en conclut que

> [...] la théorie des universaux concrets qui marchait pour le rouge, s'effondre quand on la généralise [...] Par conséquent, nous sommes conduits à reconnaître deux types différents d'association : celles des parties concrètes dans un tout concret, et celle des instances concrètes dans un universel abstrait. Nous sommes conduits à reconnaître la divergence de « est » : « C'est le Caÿster » d'une part, « c'est un carré » d'autre part (W. V. O. QUINE, 2003a, p. 115-116)

Néanmoins, précise Quine, nous ne sommes pas obligés de reconnaître une référence au prédicat « carré » :

> Tout ce que nous pouvons exiger, quand nous expliquons 'est carré' ou une autre expression, c'est que notre interlocuteur apprenne à prévoir quand nous appliquerons l'expression à un objet et quand nous ne le ferons pas ; l'expression elle-même n'a pas besoin d'être le nom d'un quelconque objet isolé (W. V. O. QUINE, 2003a, p. 117)

Les ostensions qui introduisent un terme général (« carré ») sont donc différentes des ostensions qui introduisent un terme singulier (« la Seine ») en ce que les premières, contrairement aux secondes, n'impliquent pas l'identité des objets désignés entre les différents actes d'ostension. Cette différence fournit selon Quine un « algorithme » servant de critère d'entité : pour savoir si un terme désigne ou non une entité, il suffit de vérifier qu'il satisfait à la substitution des identiques *salva veritate*.

Paradoxes des amas pluriels

Remarquons cependant que sans le concept adéquat, des actes d'ostension demeurent toujours ambigus. Que montrons-nous quand nous désignons du doigt une partie de l'espace ? Si nos concepts proviennent de la répétition d'actes d'ostension, il faut alors admettre que les objets que nous conceptualisons sont vagues.

Comment savons-nous que nous avons affaire à une chose plutôt qu'à deux ? Disposons-nous de manière innée de la notion d'objet, ou la construisons-nous progressivement à partir du matériau de notre environnement ? Comment passons-nous de l'unité d'une partie à l'unité que forme un tout ? Qu'est-ce qui régit ces hiérarchies, et comment savons-nous où nous devons les arrêter ? Une branche est une partie d'un tout qu'est un arbre, mais un arbre est aussi une partie d'un tout qui est un paysage. Nous ne considérons cependant pas qu'un paysage est un objet au sens où nous disons d'un arbre qu'il est un objet.

Dans l'article *The problem of the many* ((UNGER, 1980)), Peter Unger invite le lecteur à penser à un unique nuage flottant dans un ciel par ailleurs complètement dégagé. Vu du sol, un tel nuage

pourrait sembler bien délimité. Mais à bien y réfléchir, ce n'est pas vraiment le cas. En effet, qu'est-ce qu'un nuage sinon une nuée de gouttes en condensation suspendues dans les airs ? La densité de cet amas n'est pas régulière, puisqu'elle décroît à sa périphérie, tant et si bien qu'à un certain seuil, il devient difficile de déterminer si une goutte fait encore ou non partie de cet amas, ou si elle est seulement « proche » du nuage. Ce qui fait dire à Unger qu'il n'y a pas une seule mais de nombreuses surfaces qui s'avèrent être de bons candidats pour démarquer les frontières du nuage. Décrivant l'expérience d'Unger, David Lewis écrit que :

> Therefore many aggregates of droplets, some more inclusive and some less inclusive (and some inclusive in different ways than others), are equally good candidates to be the cloud. Since they have equal claim, how can we say that the cloud is one of these aggregates rather than another? (LEWIS, 1993, p. 164)

Ce genre de problème n'est d'ailleurs pas restreint aux nuages gazeux. Considérons le cas d'une table particulière. Nous pouvons la montrer et dire d'elle, voilà *une* table. Mais notons que notre désignation demeure vague : quand nous parlons d'*un* certain amas de matière, nous désignons en fait de *nombreux* amas. En effet, considérons à nouveau la table sous un angle microscopique : cette poussière qui s'en détache, mais qui est encore adjacente, fait-elle ou non partie de la table ? De fait, tout objet macroscopique enveloppe dans sa désignation une collection de blocs de matière différents et partiellement coïncidents. Remarquons encore que le même problème se pose pour les organismes, tant il est difficile de déterminer à partir de quel seuil un atome d'oxygène fait ou non partie du corps, ou à partir de quel seuil un fragment de la désquamation n'en fait plus partie.

De ce point de vue, lorsque nous désignons des objets ordinaires ou des organismes, croyant parler de choses particulières, nous parlerions en fait, par effet de vague, plus vraisemblablement d'une multiplicité de choses différentes mais partiellement coïncidentes. Nous croyons avoir affaire à une seule chose, mais nous parlons en fait de nombreuses choses. Ce problème motive Unger à éliminer de l'ontologie les objets ordinaires et les organismes. Partant de la constatation qu'une désignation d'un objet ordinaire n'est jamais singulière mais toujours plurielle, les théories de Peter Unger ((UNGER, 1979b)) et de Peter Van Inwagen ((P. INWAGEN, 1990)) consistent à supprimer sa dénotation, faisant disparaître l'objet ordinaire dans les vapeurs des nuages d'atomes. L'objet n'est alors qu'un mirage exercé par la présence de nombreux simples.

Considérons par exemple une réunion arbitraire de deux objets, comme par exemple, la réunion des régions spatiales que recouvrent un stylo et une table. Nous appellerons cette « chose » du nom de « stylo-table ». De fait, il paraît assez peu ordinaire de considérer qu'une telle chose constitue véritablement *un objet*. Nous serions plutôt enclins à affirmer qu'il y a là, en fait, non pas un objet, mais bien *deux* objets, ou, s'il l'on veut faire mention de l'unité, *la* réunion de deux objets. Néanmoins, Quine soutient la thèse que n'importe quelle réunion de régions spatiales, aussi éclatée et dispersée fût-elle, constitue bien un objet. On sera alors quitte de distinguer d'une part des objets dits « naturels » ou « ordinaires » qui correspondent à ceux que nous avons l'habitude de délimiter en tant que tels, et d'autre part des objets « non-ordinaires » comme les stylo-table.

Les objets non-ordinaires ne se limitent d'ailleurs pas seulement aux associations surréalistes de régions spatiales, mais s'étendent également aux réunions de parties temporelles. Nous pourrions avoir également un objet qui est cette table jusqu'au dimanche 8 mars 2020 à 8h00 (heure

française) puis qui est ce stylo après cette date. Nous voilà ainsi ramenés à nouveau au problème des doubles par métamorphoses.

La conception de la notion d'objet chez Quine est essentiellement « technique » par opposition à la conception « conservatiste » qui n'admet que les objets ordinaires. Pourquoi cet écart entre objet au sens technique de Quine et objet au sens ordinaire du terme ? Pourquoi avons-nous le sentiment que la table-stylo ne relève pas d'un authentique objet ? D'où provient cette inclinaison ? N'est-ce qu'un fait de langage dépendant d'habitudes enracinées dans nos manières de considérer le monde ?

La question de l'unité d'un objet passe par la question de ses frontières. Généralement, on distingue deux types de frontières, d'une part celles qui sont instituées, comme le sont les frontières d'une nation, et d'autre part celles qui sont naturelles comme on le dit habituellement des limites d'un organisme vivant [25]. Le libéralisme méréologique incite sans doute à considérer que toutes les frontières sont conventionnelles. Cela occasionne un nouveau problème, discuté notamment par Gareth Evans ((EVANS, 1978)), Timothy Williamson ((WILLIAMSON, 1988)) et Pascal Engel ((ENGEL, 2004)) : ces frontières sont également vagues, ce vague pouvant relever des choses elles-mêmes (vague ontologique), des conditions d'emploi des termes (vague sémantique) ou encore des conditions d'application des propriétés descriptives (vague épistémique).

Gareth Evans a présenté un argument, très discuté, sur l'impossibilité des objets vagues. Gareth Evans introduit son argument ainsi :

> It is sometimes said that the world might itself be vague. Rather than vagueness being a deficiency in our mode of describing the world, it would then be a necessary feature of any true description of it. It is also said that amongst the statements which may not have a determinate truth value as a result of their vagueness are identity statements. Combining these two views we would arrive at the idea that the world might contain certain objects about which it is a fact that they have fuzzy boundaries. But is this idea coherent? (EVANS, 1978, p. 208)

L'argument d'Evans, qui est présenté de manière formelle dans son article, peut être résumé de la façon suivante : admettons que Paris soit un objet vague en raison des limites floues de la circonscription. Dès lors, le rôle de Paris pourrait être tenu par différents objets partageant entre eux des limites communes sans qu'on puisse exactement les départager. Soit Paname l'un de ces objets. Puisque nous avons admis que Paris est un objet vague, il vient qu'il est identique seulement de manière indéterminée à Paname. En revanche, Paname est identique de manière déterminée à Paname, par la réflexivité de l'identité. Dans ce cas, on peut en déduire que Paris et Paname sont discernables au moins par une propriété : Paname a la propriété d'être identique de manière déterminée à Paname, alors que nous avons vu que cette propriété manque à Paris. Par application du principe de diversité des discernables, on peut en déduire que Paris n'est pas identique à Paname.

L'argument d'Evans a été contesté. David Lewis a montré que les objections provenaient souvent d'une mauvaise compréhension des prétentions d'Evans, qui n'aurait, selon lui, jamais voulu dire qu'il n'existe pas d'énoncés vagues :

[25] Barry Smith et Achille Varzi proposent d'utiliser les termes *bona fide* et *fiat* pour distinguer entre frontières naturelles et frontières de convention ((SMITH & VARZI, 2000))

> The correct interpretation is that Evans trusts the reader - unwisely! - to join him in taking for granted that there are vague identity statements, that a proof to the contrary cannot be right, and that the vagueness-in-describing view affords a diagnosis of the fallacy. His point is that the vague-objects view cannot accept this diagnosis, because it says that a name like 'Princeton' rigidly denotes a certain vague object. In fact, the vague-objects view does not afford any diagnosis of the fallacy, so it is stuck with the unwelcome proof of an absurd conclusion, so it is in bad trouble. (Or better, what is in trouble is the value-objects view combined with the view that vague identity yields identity statements with indeterminate truth value.) On this interpretation, every bit of what Evans says fits into place. However, he has left some important things unsaid. (LEWIS, 1988, p. 129)

De son côté, Pascal Engel ((ENGEL, 2004)), affirme que ce que montrent réellement les arguments d'Evans, c'est qu'il peut y avoir des objets vagues sans nécessairement qu'il y ait du vague dans la relation d'identité.

Dénombrements d'entités non-ordinaires

Si l'on admet que deux parties temporelles quelconques déterminent un individu, autrement dit, si l'on adhère au principe de somme méréologique non restreinte, alors la somme de la main gauche de de François Hollande et de la jambe droite de Nicolas Sarkozy constitue bien un individu. Mais cette conséquence s'avère sans doute très peu intuitive et, tout du moins, très peu désirable. On peut donc tenter de restreindre le principe méréologique par une relation. Mais la continuité spatio-temporelle et la continuité de constitution à elles seules ne permettent pas de circonscrire les limites, au sens métaphysique, d'un individu. En effet, là encore, on peut aboutir à des individus bien étranges. Considérons par exemple les parties temporelles d'un disque vinyle en les sommant de la manière suivante : on considère d'une part la somme des parties temporelles du disque qui se trouvent à l'intérieur de la pochette papier (appelons-la « disque-dedans »), et d'autre part la somme des parties temporelles du disque qui se trouvent à l'extérieur (appelons-la « disque dehors »). En sortant progressivement le disque de sa pochette comme représenté sur la figure 1.5, ces deux sommes méréologiques produisent bien deux individus au sens où ces sommes respectent le principe de continuité spatio-temporelle. D'une part, la position et le volume de matière de disque-dedans évolue de manière continue, et, dans le même temps, la position et le volume de *disque-dehors* s'accroît de manière tout aussi continue [26]. Si les relations de continuité spatio-temporelles et de continuité matérielles suffisaient à rendre compte de l'individuation métaphysique des objets, nous devrions alors admettre l'existence, au sens métaphysique du terme, des deux individus disque-dehors et disque-dedans. Mais nous sommes certainement peu disposés à reconnaître que de telles sommes constituent de vénérables individus. Il s'agit plutôt de sortes de chimères métaphysiques bâties sur le langage en tentant d'individuer les choses *malgré elles*.

Une telle expérience de pensée pourrait avoir pour effet d'inciter à considérer que les disques n'ont pas de conditions d'identité, au sens métaphysique du terme. Comme l'indique Hirsch, dans un certain langage, nous pourrions postuler des disques-dedans qui ne seraient pas identiques aux disques-dehors. Cela encouragerait manifestement une conception conventionnaliste de l'identité : deux choses pourraient être identiques dans un langage, et différentes dans un autre.

[26] Cet exemple est adapté de celui proposé par Eli Hirsch dans ((HIRSCH, 1992, p. 32)).

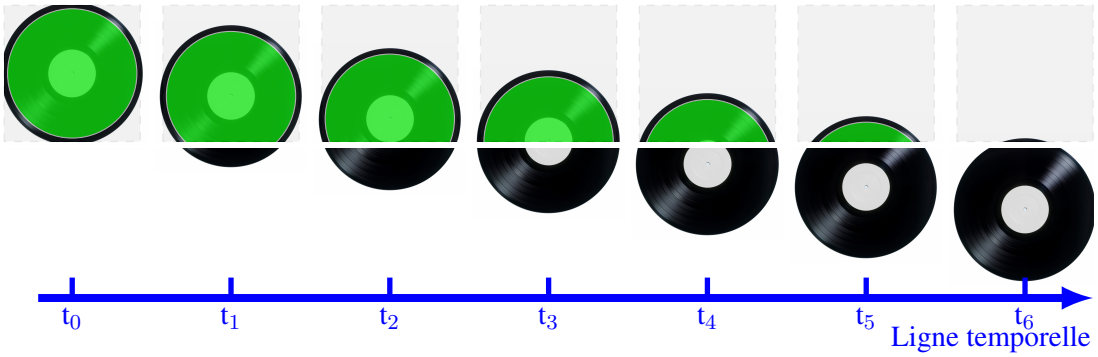


FIGURE 1.5 : Deux objets non-ordinaires : « disque-dedans » et « disque-dehors »

Il n'est sans doute pas la peine de rechercher des particuliers dérogeant aux critères habituels d'unité parmi des objets aussi extra-ordinaires que les disques-dedans. Nous avons à notre portée des entités que le langage ordinaire admet au rang des particuliers. Parmi elles figurent les ombres, les flaques, les nuages, les trous, les bosses. Il semble en effet assez certain que ce genre de choses s'avèrent être des particuliers. Nous pouvons les désigner directement, nous pouvons tendre la main vers et dire : « Voilà *une* ombre ! », « Voilà une flaque ! » ou encore « Voilà un trou ! ».

Ce qui différencie par exemple les critères d'unité et d'identité des ombres de ceux des particuliers ordinaires, c'est qu'elles peuvent se superposer, coïncider, se désolidariser, fusionner ou encore fissionner. Une ombre peut en cacher une autre. Deux ombres superposées devraient alors compter plus d'une fois, plus précisément, autant de fois qu'il y a d'objets dont elles sont les projections. Nous pouvons alors avoir le cas de deux particuliers situés exactement au même endroit.

Le fait que l'ombre paraisse connexe et ciné-cohésive en dépit de son immatérialité implique qu'elle est souvent prise pour une partie propre de l'objet, comme le montrent les superstitions ancrées sur le rapport du soi à l'ombre dans de nombreuses civilisations. De nombreux auteurs ont associé le thème de l'ombre à celui du double, notamment, parmi les plus connus, Chamisso dans *L'Étrange histoire de Peter Schlemihl* (CHAMISSO, 1874), Hofmannsthal dans *La femme sans ombre* (qui servit de livret à l'opéra de Richard Strauss) (HOFMANNSTHAL, 1992), Andersen dans *L'ombre* (ANDERSEN, 1876) et Abe Kôbô dans *Le Tanuki de la tour de Babel* (KÔBBÔ, 1994). Selon la plupart des interprétations, la disparition de l'ombre signifie un amoindrissement de la substance de celui qui en est victime. Dans *Le Double*, Otto Rank associe directement l'ombre à l'âme (RANK, 1973). Mais on peut aussi faire l'hypothèse contraire que l'ombre n'est pas une partie immatérielle de l'être, mais au contraire un révélateur de la matérialité des êtres physiques. Sa disparition implique une sorte d'effacement plutôt qu'une perte de substance immatérielle. Les personnages qui perdent leur ombre ne perdent pas une partie de leur substance mais perdent le témoignage de leur existence matérielle : ce n'est pas leur individuation qui est affectée par cette disparition, ce sont les relations publiques qu'ils entretiennent avec autrui. Dans les romans cités, l'enjeu de la perte de l'ombre n'est d'ailleurs pas ontologique, mais social.

Même si l'on a bien affaire à une chose étendue dans le temps et l'espace, et en dépit du fait que ce genre de particuliers puisse être provisoirement saisi comme une unité, une ombre ne

dispose pas de conditions assurant une identité temporelle suffisamment stable. L'unité temporelle de l'ombre a un caractère précaire. L'existence temporelle d'une ombre peut être discontinue, et sa persistance problématique.

À ce sujet, Roberto Casati souligne que notre schéma général du dénombrement recourt à un modèle spécifique de choses-à-compter, les cailloux, dont il n'est pas inutile de rappeler ici qu'il est la source étymologique du mot « calcul » :

> Les cailloux, (et avec eux toutes les choses matérielles) sont notre modèle des choses qui se comportent « bien » et, assurément, se comportent tout autrement que les ombres : deux cailloux ne peuvent occuper la même place ; un caillou n'est plus le même si nous le cassons en plusieurs morceaux. Si nous fusionnons deux cailloux, nous en créons un troisième en faisant disparaître les deux autres. Le caillou semble se déplacer de façon continue dans l'espace. Si nous laissons un caillou quelque part, nous le retrouvons au même endroit, et si nous ne le retrouvons pas, nous savons que quelqu'un l'aura déplacé. (CASATI, 2002, p. 60-61)

Comme l'écrit Casati, les cailloux se « comportent bien » parce qu'ils possèdent des caractéristiques spatio-temporelles permettant de faciliter leur dénombrement. Au contraire des ombres et des flaques d'eau, le déplacement de cailloux n'affecte pas leur effectif. Les cailloux ne se superposent pas. Ils ne fusionnent ni ne fissionnent, ne donnent donc pas lieu à des paradoxes prêtant à affirmer que deux choses au moment t en sont une seule au moment t'. Enfin, les cailloux persistent continûment dans le temps et l'espace.

La question est donc de savoir pourquoi nous donnons l'avantage aux critères qui s'avèrent pertinents pour les cailloux plutôt qu'à ceux qui conviendraient mieux aux ombres et aux nuages. Il semble évident que des aspects pratiques entrent en compte : nous aurions moins de difficultés à heurter l'ombre d'un camion que le camion lui-même.

1.4 La théorie conventionnaliste de l'identité

Selon Stéphane Chauvier,

> Être un individu, tout comme être un objet, être une propriété ou être un genre, c'est, non pas exclusivement certes, mais au moins préalablement, être *pensé* ou *signifié* d'une certaine manière, d'une manière logiquement voire cognitivement déterminée ou typique. [...] Nous appliquons le concept d'individu à des composantes de la réalité, non sur la base de leur présentation phénoménale, mais sur la base du mode de présentation catégoriale typique de certains termes de première intention. Quelque chose est un individu ou peut être pensé comme un individu parce qu'on y pense, en première intention, d'une certaine manière. (CHAUVIER, 2010, p. 6)

L'individuation repose-t-elle sur une présentation catégoriale ? Mais comment construisons-nous de telles catégories si elles ne préxistent pas, autrement dit, s'il ne s'agit pas de reconnaissances ? S'agit-il d'une présentation seulement linguistique ? Dans cette section, nous présenterons la perspective conventionnaliste de l'identité. Nous distinguerons le conventionnalisme fort,

pour lequel l'unité est un fait seulement linguistique, du conventionnalisme modéré pour lequel l'unité possède une certaine réalité psychologique (mais pas nécessairement métaphysique). Nous adopterons la version modérée.

1.4.1 Objections à la théorie platonicienne de Frege

Objections au statut platonicien des nombres

Les preuves que donne Frege pour montrer que le nombre n'est pas une propriété tirée des choses par la sensation semblent définitives. Mais cela n'implique pas de souscrire nécessairement à l'idée que le nombre est une réalité en soi.

Tournons-nous vers l'idée, développée par John Stuart Mill, que la propriété d'être deux et celle d'être douze sont abstraites de l'observation répétée de collections de paires et de douzaines :

> Le fait énoncé dans la définition d'un nombre est un fait physique. Chacun des nombres, deux, trois, quatre, etc., dénote des phénomènes physiques et connote une propriété physique de ces phénomènes. Par exemple, deux dénote toutes les paires, et douze toutes les douzaines d'objets ; lesquels connotent ce qui en fait des paires et des douzaines, c'est-à-dire une propriété physique. On ne niera pas, en effet, que deux pommes ne puissent être physiquement distinguées de trois pommes, deux chevaux d'un seul, etc., que ce sont là des phénomènes différents, visibles et tangibles. (J. S. M. MILL, 1866, p. III. 24. 5)

Le nombre provient bien d'une réalité physique, mais, *pace* Locke, il n'est pas une propriété agissant sur les sens comme être jaune ou être courbe, il s'agit d'une relation entre différentes choses associées ou regroupées.

Frege conteste cependant cette caractérisation en constatant que l'étendue d'application de l'idée de nombre outrepasse les seules choses sensibles. De fait, les nombres s'appliquent aussi bien aux amas de cailloux qu'aux événements et qu'aux concepts. C'est d'ailleurs ce que remarquait aussi Locke, soutenant que « le nombre s'applique en effet aux hommes, aux anges, aux actions, aux pensées, à toute chose qui existe ou peut être imaginée » (LOCKE, 2001, Livre II, chapitre XVI, §1). Selon Frege, le nombre n'est donc pas une abstraction résultant de sensations émanant de phénomènes physiques car, écrit-il,

> il serait bien étonnant qu'une propriété abstraite des choses extérieures puisse être transposée sans altération de sens aux événements, aux représentations, aux concepts. C'est comme si l'on voulait parler d'un événement fondant, d'un concept salé, d'une représentation bleue, ou d'un jugement épais. (FREGE, 1969a, p. 151)

Si le nombre deux naissait de l'impression sensible d'un amas de deux choses comme, disons, l'observation de deux cailloux, la même impression sensible devrait alors aussi résulter de la réunion de deux concepts arbitrairement choisis, comme « être à l'heure à son rendez-vous » et « résoudre un problème de trigonométrie ». Cela paraît absurde.

On peut cependant répondre à l'objection de Frege que si le nombre provient bien d'une abstraction tirée de l'expérience sensible d'une réunion de choses, il gagne ensuite, en tant qu'abstraction, une sorte d'autonomie eu égard aux choses matérielles d'où il émane. C'est par une forme

de « comme si », une fiction, qu'on en vient à dire qu'il y a deux concepts comme l'on dit qu'il y a deux cailloux, en étendant l'idée à des choses plus abstraites.

La conception frégenne du nombre a été vivement contestée par Ludwig Wittgenstein. Comme le souligne Jacques Bouveresse dans *La Force de la Règle* (BOUVERESSE, 1987), pour Wittgenstein, faire découler la notion de nombre de celle de successeur, comme le fait Frege, est d'emblée suspect. Car il ne va pas de soi que ce que l'on fait pour passer de 1 à 2 soit identique à ce que l'on fait pour passer de 2 à 3. Car dans un cas, on pourrait dire qu'il s'agit de la multiplication par 2, alors que dans l'autre, il s'agit de la multiplication par $\frac{3}{2}$. Constatons ainsi que passer d'un nombre à un autre peut conduire à faire à chaque fois des choses différentes. Or c'est la notion de *règle* qui détermine la signification de « faire la même chose ». C'est parce que nous appliquons une règle que nous en retrouvons, à chaque fois, les conséquences. Comme le remarque Jacques Bouveresse,

> L'impression de déjà connu que l'on éprouve en présence du cas nouveau et la propension à appliquer à nouveau l'expression concernée, sont, en réalité, une seule et même chose. (BOUVERESSE, 1987, p. 39)

On ouvre ainsi une critique de la notion frégenne de nombre par des arguments constructivistes. Dans la conception ordinale, favorisée notamment par Dedekind, Cantor et Cassirer, l'équipotence n'est pas un prérequis de la notion de nombre. Le nombre est considéré selon sa place dans une suite. Plus encore, il est complètement défini ainsi, perdant au passage tout contenu ontologique. Autrement dit, alors que pour Frege, il existe des objets, les nombres, qui s'ordonnent selon une suite régulière, selon Ernst Cassirer, c'est au contraire la place qu'occupent les nombres dans la suite en question qui les détermine comme objets :

> [...] there is a system of ideal objects whose whole content is exhausted in their mutual relations. The "essence" of the numbers is completely expressed in their positions. And the concept of position must, first of all, be grasped in its greatest logical universality and scope. (CASSIRER, 1953, p. 40)

Du côté de la psychologie, Jean Piaget a montré que l'apprentissage du dénombrement ne consiste aucunement à les extraire des objets, mais bien plutôt à effectuer des actions « telles que réunir, ordonner, mettre en correspondance, etc. » (PIAGET, 1968). Par exemple, pour dénombrer les éléments d'une collection, on peut les apparier un à un aux mots énumérés lors d'une comptine comme « un, deux, trois, etc. ». L'arrêt de cet algorithme survient lorsqu'il n'y a plus d'objet à faire correspondre et le dernier mot énuméré indique alors le cardinal de la collection des choses dénombrées. On peut aussi faire appel à la seule perception visuelle en appréhendant directement la disposition des éléments et en les identifiant à des figures dont la correspondance au cardinal est déjà connue et mémorisée, comme les points sur la face d'un dé [27]. Et bien sûr, il y a aussi l'appariement avec les doigts de la main. La comptine, les figures ou les doigts de la main servent alors de collection paradigmatique pour retrouver le nom propre correspondant au nombre cardinal recherché.

[27] Les didacticiens des mathématiques désigne cette procédure sous le nom de méthode par *subitizing*.

Il peut être contesté que de tels arguments ne disqualifient pas la conception de Frege parce qu'ils se rapportent non à la nature des nombres, mais aux manières dont les personnes en construisent la représentation. Il n'en reste pas moins que, contre l'approche frégéenne qui considère le nombre comme une entité séparée, on peut toujours favoriser une conception comme celle de Cassirer, qui identifie le nombre à un réseau de relations.

Cassirer demande :

> La dérivation de la série des nombres est-elle faite à partir du concept de classe ou bien se meut-elle en cercle dans la mesure où elle présuppose tacitement des concepts empruntés au domaine qu'elle entreprend de réduire ? (CASSIRER, 1953, p. 102)

L'argument le plus fort employé par Cassirer contre la conception cardinale du nombre est que les classes qui sont dites de même cardinal doivent d'abord apparaître comme diverses, elles doivent manifester l'idée d'une pluralité, elles doivent être vues comme différentes, ce qui ramène de manière frontale la notion de nombre qui était censée être réduite à l'équipotence.

La conception idéaliste de l'unité chez Berkeley

La conception de l'unité chez Berkeley est directement reliée à la définition des objets ordinaires, qui sont considérés comme des agglomérations d'idées sensibles. Une cerise, écrit Berkeley, n'est finalement rien d'autre qu'un amas d'impressions sensibles ou d'idées qui sont corrélées entre elles par l'esprit en une unité (BERKELEY, 2009c, p. 229) en raison de la répétition d'impressions récurrentes. De ce point de vue, « chose » signifie déjà « unité ». L'unité est inscrite dans la définition même de la chose, elle est essentielle à l'objet, elle n'est pas une propriété qui pourrait librement s'y ajouter ou s'y soustraire. En ce sens, elle ne peut être une propriété de la chose. Or, Berkeley souligne que la raison d'être d'une même dénomination, autrement dit, d'un même regroupement d'idées sensibles au sein d'une même unité, est purement conventionnelle :

> Now, this naming and combining together of ideas is perfectly arbitrary, and done by the mind in such sort as experience shows it to be most convenient: without which our ideas had never been collected into such sundry distinct combinations as they now are. (BERKELEY, 2009b, p. 47)

Du point de vue de Berkeley, dénombrer est une affaire de convention. Nous pouvons toujours compter des choses d'une certaine manière mais aussi d'une autre. Berkeley en conclut que le nombre n'est pas une idée résultant des choses, mais qu'il dépend étroitement de la manière dont l'esprit conçoit les choses. L'unité, conteste Berkeley, n'est pas une abstraction qui émanerait de l'expérience comme le serait une couleur ou une forme. Il s'agit d'une entité purement mentale, une « créature de l'esprit ». Nous avons ainsi affaire à une chose ou à plusieurs non en vertu de ce qu'il y a, mais en rapport à notre manière de les considérer.

Ainsi, dans le *Traité concernant les principes de l'entendement humain*, Berkeley souligne que le nombre « is entirely the creature of the mind, even though the other qualities be allowed to exist without » (BERKELEY, 2009a, p. 87), et dans *L'essai pour une nouvelle théorie de la vision*, il écrit que :

> [...] it ought to be considered that number (however some may reckon it amongst the primary qualities) is nothing fixed and settled, really existing in things themselves. It is entirely the creature of the mind, considering either an idea by itself, or in any combination of ideas to which it gives one name, and so makes it pass for an unit. According as the mind variously combines its ideas, the unit varies, and as the unit, so the number, which is only a collection of units, also varies. We call a window one, a chimney one, and yet a house in which there are many windows and many chimneys has an equal right to be called one, and many houses go to the making of one city. In these and the like instances it is evident the unit constantly relates to the particular draughts the mind makes of its ideas, to which it affixes names, and wherein it concludes more or less as best suits its own ends and purposes. Whatever, therefore, the mind considers as one, that is an unit. Every combination of ideas is considered as one thing by the mind, and in token thereof is marked by one name. (BERKELEY, 2009b, p. 47)

Il apparaît donc que Berkeley considère à la fois que différentes idées sensibles peuvent être associées en une seule idée d'objet, et à la fois qu'un objet puisse être identique à lui-même. L'idée d'unité est en quelque sorte à la fois inextricable et à la fois contradictoire à l'idée d'identité numérique, puisque une même idée peut être décomposée en une multitude d'autres. Berkeley n'est pas réellement soucieux de rendre compte d'une différence entre l'identité numérique et le caractère purement arbitraire de l'unité, autrement dit, il ne semble pas distinguer le fait pour une chose d'être la même et le fait de parvenir à combiner plusieurs choses en une seule. En effet, Berkeley parle indifféremment de « différentes choses réputées être une » (BERKELEY, 2009c, p. 83) ou « décrites comme une seule et même chose » (BERKELEY, 2009b, p. 23). D'une part, des idées numériquement différentes peuvent être prises pour une seule chose, puisque leur agglomération en unité ne dépend que de convenances. Mais d'autre part, on peut aussi reconnaître une seule et même idée. D'un tel point de vue, il y aurait chez Berkeley un balancement entre l'identité et le double, toute chose pouvant toujours être vue comme simple ou comme double. Comme le remarque Donald Baxter,

> Considering two things as one thing is considering them no longer to be two. So it is considering them no longer to be distinct, but rather to be identical. Conversely, considering them to be identical is considering them as one thing. (BAXTER, 1991, p. 94)

Objections à la notion grammaticale d'objet

Si la notion frégéenne d'unité peut être remise en question, c'est sans doute aussi le cas de la notion d'objet. Nous avons vu qu'elle est, chez Frege, de part en part sémantique. Néanmoins, Frank Ramsey a objecté qu'il n'est en fait pas possible de distinguer entre sujet et prédicat au seul moyen du langage (RAMSEY, 1990)). Ramsey constate en effet que toute expression contenant un sujet et un prédicat peut-être reformulée en une nouvelle expression, ayant la même signification, mais où sujet et prédicat seraient échangés. L'expression « Socrate est sage » peut par exemple être reformulée par « la sagesse est une qualité de Socrate ». En outre, les formulations réflexives semblent parfaitement rétives à l'interprétation grammaticale des particuliers. Car une phrase comme « Brutus a tué César » a le même sens que « César a été tué par Brutus ». On peut

ainsi tout autant l'interpréter comme une affirmation concernant Brutus que comme une affirmation au sujet de César : « there is no essential distinction between the subject of a proposition and its predicate, and no fundamental classification of objects can be based upon such a distinction » (RAMSEY, 1990).

De son côté, dans *Les Individus*, Peter Strawson suggère l'idée qu'on peut cependant conserver l'esprit de la définition frégéenne de la notion d'objet, à condition de ne pas la restreindre à la seule grammaire. Le langage permet d'identifier des particuliers non seulement grâce aux désignations directes, aux descriptions définies, aux noms propres, mais aussi parce que nous ramenons les objets qui nous entourent à un espace communément partagé, et auquel nous pouvons nous référer :

> Nous pouvons, les uns à l'égard des autres, préciser sur quelles choses particulières porte notre discours parce que nous pouvons intégrer les comptes rendus et les histoires de chacun dans un unique tableau du monde ; le cadre de ce tableau est un cadre unitaire spatio-temporel doté d'une dimension temporelle et de trois dimensions spatiales. (STRAWSON, 1959b, p. 41-42)

Pour Strawson, la toile qui permet de dessiner le tableau du monde, le cadre de référence des identifications est l'ensemble des entités matérielles. Mais il subsiste une condition pour la possession d'un tel cadre, qui est la capacité de pouvoir ré-identifier certains éléments de ce cadre nonobstant les discontinuités inévitables dans l'observation d'une chose étendue dans le temps.

Strawson introduit ainsi une composante pragmatique dans la définition linguistique d'objet : un objet se distingue d'un prédicat, non pas en raison de sa place dans la grammaire, mais parce qu'en parler nécessite préalablement une identification par l'énonciateur. Les objets ressortissent à des termes dont les références sont *saillantes*, autrement dit, qui émergent sur le fond du monde. Ainsi, comprendre la phrase « Brutus a tué César » nécessite de savoir qui est Brutus et qui est César, ou tout du moins parvenir à les identifier. Or l'identification s'applique d'une manière complètement différente quand il s'agit de propriétés ou d'objets. Ce n'est pas la grammaire qui permet de distinguer un individu comme Brutus d'une propriété comme celle d'avoir tué César, mais la manière dont des locuteurs utilisent ces mots avec l'intention de communiquer :

> L'introduction identifiante d'un particulier ou d'un universel dans le discours implique que l'on sache quel est le particulier ou l'universel signifié par l'expression introductrice. Savoir quel est le particulier ainsi signifié implique connaître, ou parfois - dans le cas de l'auditeur - apprendre, à partir de l'expression introductrice utilisée, un fait empirique quelconque qui suffit à identifier ce particulier [...]. Mais savoir quel universel est signifié n'implique pas, de la même manière, la connaissance d'un fait empirique quelconque : simplement, une connaissance du langage. (STRAWSON, 1959b, p. 208)

Constatons alors qu'il existe de nombreux cas pour lesquels l'identification d'un certain particulier nécessite celle d'un autre particulier. Par exemple, l'identification d'une ombre dépend en grande partie de l'identification de l'objet qui la projette, et l'identification de l'assassin de Kennedy dépend bien sûr de l'identification de Kennedy. Strawson appelle « particuliers privés » la classe des particuliers dont l'identification des membres dépend de l'identification d'une deuxième classe sans qu'il soit possible d'identifier autrement l'un de ses membres. Dans cette classe on trouve par exemple les groupes de sensations, les événements mentaux, et les données sensibles

qui dépendent des particuliers que sont les personnes. Ensuite, Strawson détermine une classe privilégiée de particuliers qui sert de fondement à l'identification des classes privées. Ces particuliers, Strawson les appelle des « particuliers de base ». Il confère ce rôle aux entités matérielles, entendues comme choses qui résistent au toucher, autrement dit des particuliers qui ne sont pas seulement visibles comme les gaz colorés ou les rayons de lumière. Le rôle de particulier de base est également étendu à la notion de personne. De ce point de vue, de nombreux objets ne peuvent être identifiés que de manière épistémique, et non pas en vertu d'une relation métaphysique.

Objections à la thèse de l'identité objectuelle

Peut-on généraliser cette dernière remarque à l'ensemble de tous les individus ? N'y a-t-il, au fond, que des particuliers privés ? Il s'agit de se demander s'il est bien légitime de considérer un principe métaphysique d'identité. On pourrait en effet rejeter la dimension ontologique de l'identité.

De fait, une objection à la conception relationnelle de l'identité consiste à demander s'il est vraiment raisonnable de parler de relation pour désigner une propriété nécessairement réflexive, autrement dit, pour désigner quelque chose qui ne se tient qu'entre une chose et elle-même, c'est-à-dire pour qualifier une relation qui, justement, ne relie jamais « deux » éléments, mais seulement un seul. C'est un fait que certaines relations puissent s'avérer réflexives de manière contingente, comme par exemple « être responsable de ». Mais c'est autre chose d'affirmer d'une propriété qui est nécessairement réflexive qu'il s'agit d'une relation. Aussi, peut-être devrions-nous considérer que l'identité est une propriété monadique plutôt qu'une relation dyadique. Mais peut-être ne relève-t-elle ni d'une propriété ni d'une relation. C'est notamment ce qu'indique Betrand Russell qui écrit que :

> The question whether identity is or is not a relation, and even whether there is such a concept at all, is not easy to answer. For, it may be said, identity cannot be a relation, since, where it is truly asserted, we have only one term, whereas two terms are required for a relation. And indeed identity, an objector may urge, cannot be anything at all: two terms plainly are not identical, and one term cannot be, for what is it identical with? (RUSSELL, 2010, p. 65)

Saul Kripke a ardemment défendu la thèse de l'identité objectuelle contre ce genre d'argument. Il suggère que même si l'on considère que le « est » de l'identité n'est qu'une relation linguistique de co-référence, autrement dit, une relation entre les termes désignant un même objet plutôt qu'entre les objets eux-mêmes, on est toujours en droit d'introduire par stipulation une nouvelle relation, appelons-la, comme il le suggère, « schmidentité », relation qui vaudrait entre un objet et lui-même. Dans ce cas, la question de savoir si Ciceron est schmidentique à Tully ressurgirait :

> Quiconque y pense sérieusement verra, à mon avis, que la conception de l'identité comme relation entre des noms, loin d'être le seul moyen de résoudre les problèmes en question, ne permet même pas de les résoudre ; qu'il faut donc l'abandonner, et concevoir l'identité comme la relation qu'entretient une chose avec elle-même. (KRIPKE, 1982, p. 97)

Bien sûr, toute la question réside dans le fait de savoir si stipuler une telle relation est ou non

légitime. Kripke le reconnaît d'ailleurs, et, commentant son propre argument, écrit en note de bas de page que :

> Ce procédé ne convaincra évidemment pas un philosophe qui prétendrait qu'un langage ou un concept artificiel de ce type est logiquement impossible. Dans le cas présent, certains philosophes ont cru qu'une relation, étant essentiellement à deux termes, ne peut pas mettre en rapport une chose avec elle-même. Cette position est purement et simplement absurde. Quelqu'un peut parfaitement être à lui-même son pire ennemi, plus sévère critique, et ainsi de suite. Certaines relations sont réflexives, comme par exemple la relation « pas plus riche que ». L'identité ou la schmidentité n'est que la plus petite relation réflexive. (KRIPKE, 1982, p. 97)

Considérons les exemples mentionnés par Kripke. Notons qu'ils mentionnent justement des relations qui ne sont pas seulement réflexives, ou qui ne le sont que de manière très métaphorique. En les réinterprétant, il apparaît immédiatement l'affirmation d'une identité. La première peut en effet être formulée : « être le pire ennemi d'un tel et être identique à celui-ci », et la seconde : « être le critique d'un tel et être identique à lui ». Quant au dernier exemple, il consiste simplement à affirmer qu'il existe toutes sortes de relations réflexives, ce qui ne fait pas de doute. De tels exemples ne sont donc pas du tout convaincants, et ne fournissent pas une exemplification de la relation recherchée.

C'est sans doute de manière très métaphorique que Kripke utilise la formulation « plus petite relation ». Il est certain qu'il ne veut pas faire mention d'une relation d'ordre, au sens mathématique du terme, et qu'il faut entendre sa proposition ainsi : une relation qui serait réflexive et qui ne serait que cela. Au sens extensionnel du terme, cela consiste à dire qu'on peut toujours faire voire la relation comme une fonction qui assigne à chaque monde possible un ensemble de couples d'objets, et qu'il suffit de considérer alors l'ensemble des couples redondant du type (x, x) pour caractériser l'identité.

Au sens ensembliste, c'est incontestable. Mais ce que prétend Kripke, ce n'est pas seulement définir une relation formelle, mais une relation *métaphysique*, objectuelle. Or la difficulté qui me semble la plus importante, c'est qu'on ne peut pas rendre compte d'une bijection entre les relations et les ensembles, ne serait-ce qu'en raison de l'asymétrie de certaines relations, asymétrie qui implique qu'il y a plus d'ensembles que de relations. Il n'est donc pas certain que l'ensemble des paires (x, x), (y, y), etc. représente une relation au sens métaphysique du terme [28].

Wittgenstein a notoirement contesté le caractère relationnel de l'identité : « Que l'identité ne soit pas une relation entre objets, c'est évident. » (WITTGENSTEIN, 1993, p. 5.5301). Ainsi, pour Wittgenstein, l'identité ne relève pas d'une relation entre une chose et elle-même, relation qui devrait être formulée par un principe dans un langage logique. Plus précisément, selon Wittgenstein, la bonne manière de stipuler l'identité consiste à n'utiliser qu'un seul symbole pour désigner une et même chose [29]. L'identité du signe ne doit en aucun cas servir de prétexte pour introduire un signe d'identité. Car constater que « $a = a$ », c'est dire deux fois la même chose, et au fond, nous

[28] Pour plus de précisions sur les difficultés générées par une telle tentative, on pourra se référer à (WEHMEIER, 2012).

[29] Par comparaison, on trouve également dans les écrits d'Heidegger une remarque semblable concernant la formalisation de l'égalité. Mais *a contrario* de Wittgenstein, Heidegger fait usage de cette distinction pour nier que l'égalité

pourrions nous contenter d'écrire simplement « a », comme l'encourage Wittgenstein à plusieurs reprises dans le *Tractatus Logico-Philosophicus* [30] :

> J'exprime l'égalité des objets par l'égalité des signes, et non au moyen d'un signe d'égalité. J'exprime la différence des objets par la différence des signes. (WITTGENSTEIN, 1993, p. 5.53)

Du point de vue de Wittgenstein, l'identité ne relie pas des choses, mais exclusivement des signes. L'identité est en fait toujours déjà incorporée à la signification d'un nom : deux signes sont équivalents s'ils ont même signification, autrement dit, s'ils sont synonymes. Ainsi, les énoncés d'identité ne sont pas des propositions élémentaires. Plus encore, ils n'ont pas même de sens :

> 4.243 - Pouvons-nous comprendre deux noms sans savoir s'ils désignent la même chose ou deux choses différentes - Pouvons-nous comprendre une proposition où apparaissent deux noms sans savoir s'ils ont même signification ou des significations différentes ? Si je connais la signification d'un mot anglais et de son équivalent allemand, il est impossible que je ne sache pas qu'ils sont équivalents ; il est impossible que je ne puisse les traduire l'un par l'autre. Des expressions comme « $a = a$ », ou celles qui en dérivent, ne sont ni des propositions élémentaires, ni même des signes pourvus de sens. (WITTGENSTEIN, 1993)

Aussi, Wittgenstein considère que le signe d'égalité marque la synonymie subsistant entre deux termes, mais non l'identité de la référence. L'identité n'est qu'un simple processus de substitution des termes d'une proposition sans changement de valeur de vérité. Le signe d'égalité marqué entre deux expressions ne signifie rien d'autre que la possibilité de substitution de ces expressions *salva veritate* (WITTGENSTEIN, 1993, 4.241 et 6.23).

Wittgenstein insiste particulièrement sur le fait qu'il semble n'y avoir que deux manières contraires mais tout aussi vaines d'appliquer l'identité logique aux entités : « Dire de deux choses qu'elles sont identiques est une absurdité, et dire d'une chose qu'elle serait identique à elle-même, c'est ne rien dire du tout »(WITTGENSTEIN, 1993, p. 310). D'une part, appliquer l'identité à *deux* choses est manifestement absurde. Affirmer d'une chose qu'elle est identique à une autre relève du non-sens. Mais d'autre part, appliquer l'identité à *une* chose est absolument insignifiant parce qu'affirmer d'une chose qu'elle est identique à elle-même est une sorte particulièrement vide de pléonasme. La portée rationnelle de l'identité semble alors confinée à une forme de sèche et vaine redondance. L'identité logique semble ou inapplicable ou sans intérêt.

---

numérique soit l'identité et pour mettre en avant une notion métaphysique d'identité. Précisément, Heidegger constate que la formulation de l'identité par le signe d'égalité nécessite l'utilisation de deux termes ($a = b$) ou de deux occurrences du même terme ($a = a$). Or il constate aussi que pour désigner l'identité d'une chose, un seul terme suffit. Il en conclut que l'égalité n'est tout simplement pas l'identité : « La formule « A = A » indique une égalité. Elle ne présente pas A comme étant le même. » (HEIDEGGER, 1957, p. 257). Selon Heidegger, l'identité logique est une forme trompeuse qui « voilerait » ce que le principe d'identité veut véritablement signifier, à savoir, « la restitution de la chose à elle-même ». Le but que se propose Heidegger est opposé à celui de Wittgenstein : il ne s'agit pas de simplifier la notation, mais d'évincer complètement le rôle que joue la logique formelle dans la quête de l'identité, afin de la remplacer par le langage ordinaire. Cependant, on peut mettre en doute l'idée que ce dernier abriterait « la maison de l'Être », et il y a plus encore, beaucoup de raisons de croire qu'il loge des amphibologies menant à de bien spécieux discours.

[30] Voir aussi les aphorismes (WITTGENSTEIN, 1989, aph. 5.52, 5.531, 5.532, 5.5321).

Doit-on se résoudre à l'idée que l'identité logique n'est rien d'autre qu'une énonciation vaine ou absurde, en tout cas, sans conséquence en termes de connaissance car bornée à la seule constatation que chaque chose coïncide parfaitement et exclusivement avec elle-même ? Si l'identité échappe constamment aux choses, abrite-t-elle encore quelque mérite cognitif concernant le monde ? Comme le souligne Clément Rosset dans *Le démon de la tautologie* (ROSSET, 1997), l'affirmation que « a est identique à a » possède en réalité une vertu qu'on retrouve de façon plus générale dans la tautologie : sa sécheresse épistémique est compensée par son caractère résolument inattaquable. La vérité logique exprimée par « $a = a$ » est d'une évidence incontestable, et c'est ce qui en fait sans doute la force. Cette évidence de l'identité s'oppose aux illusions du double. Car, remarque Rosset [31],

> la définition même de l'illusion que de ne jamais se résoudre, ou se résigner, à l'application stricte du principe d'identité (et c'est pourquoi je pense que l'illusion est toujours marquée au coin du double, d'une duplication hallucinatoire de l'unique qui constitue précisément sa 'duplicité'.) (ROSSET, 1997, p. 11)

Ainsi, le mérite de l'identité logique réside en ce que son énonciation n'excède nullement la part de ce qu'on peut espérer dire avec justesse d'une chose, parce qu'elle commence et s'achève dans l'énonciation de la coïncidence de cette chose avec elle-même. On peut ainsi opposer au caractère *idiot* - c'est-à-dire, au sens étymologique du terme, *simple, trivial* -, qui caractérise inévitablement tout énoncé d'identité, le caractère mensonger et illusoire de nos représentations lorsqu'elles s'écartent plus ou moins de cette relation triviale.

Revenons au cas des Dupondt sur la lune. Le ressort comique tiré par Hergé provient d'un désaccord entre une conception immanentiste de l'unité et une vérité arithmétique élémentaire. On trouve déjà la même question posée dans l'*Hippias Majeur*, où Platon expose un sophisme similaire. Dans la pièce de Platon, Socrate indique à Hippias qu'à eux deux, ils sont à la fois *un* (groupe de deux personnes) et à la fois *deux* (individus). En effet, Socrate interroge :

> Chacun de nous deux, séparément, n'est-il pas un ? et cette qualité d'être un, n'est-il pas dans le cas de la posséder ? (...) Mais est-ce que par hasard, pour la raison que tous deux ensemble nous sommes pairs, est-ce que, pour cette raison, chacun de nous, séparément, est pair ? (...) Sommes-nous deux, l'un et l'autre pris ensemble ? alors forcément chacun de nous est aussi forcément deux, sommes-nous chacun deux ? alors, tous deux ensemble aussi, nous sommes forcément un. (PLATON, 1955, 302a)

Les prémisses du raisonnement spécieux exposé par Platon reposent sur la thèse que les nombres sont dans les choses, *indépendamment* des autres objets du monde. De ce fait, l'erreur commise par les Dupondt, lorsqu'ils affirment « nous sommes seuls » s'avère finalement être de considérer que l'unité d'une chose est une propriété intrinsèque plutôt qu'extrinsèque : le nombre serait instancié par la chose même qu'il qualifie, ceci, indépendamment des autres choses qui l'environnent. De ce point de vue, bien sûr erroné, un Dupont et un Dupond, cela ne ferait qu'une seule

[31] C'est également ce que constate Nietzsche dans un passage de *La Philosophie tragique à l'époque des Grecs* : « s'il refuse de se contenter d'une vérité sous forme de tautologies, c'est-à-dire de cosses vides, (l'homme) échangera éternellement des illusions contre des vérités. (NIETZSCHE, 1975, p. 280) ».

personne, et non deux. Il est alors possible de répondre à l'objection de Frege ainsi : le nombre n'est certes pas une propriété intrinsèque d'une chose, en revanche, il peut être vu comme une propriété relationnelle.

1.4.2 La thèse de l'identité relative de Peter Geach

L'argument de la loi de Buridan

Parmi les arguments contre l'idée que l'identité soit une relation objectuelle, on trouve dans la littérature ce que Peter Geach désigne sous le nom de « loi de Buridan » (P. T. GEACH, 1987, p. 9). Cette loi selon stipule que :

(Loi de Buridan) : La référence d'un nom dans une proposition ne saurait changer en fonction de la valeur de vérité de la-dite proposition.

Pour reprendre l'exemple donné par Geach, si je dis par exemple que « Fifi est un chien », la proposition exprimée est soit vraie, soit fausse, mais sa valeur de vérité ne devrait rien changer au fait qu'elle s'applique à Fifi. Pour voir ceci plus clairement, considérons le petit jeu logique introduit par Geach. Admettons que le nom « A » soit utilisé pour vous nommer quand il entre dans une proposition vraie, et qu'il soit utilisé afin de nommer l'âne Brownie quand il entre dans une proposition fausse. Que faisons-nous quand nous affirmons « Brownie est un âne » ? Si la proposition exprimée par cette phrase s'avère fausse, alors « Brownie » est utilisé dans une proposition fausse, et nous pouvons en conclure par hypothèse que le terme désigne l'âne Brownie, et donc que la proposition exprimée par « Brownie est un âne » est vraie. Mais alors, le nom « Brownie » est utilisé pour exprimer une proposition vraie, et dans ce cas, il fait référence à vous, donc vous êtes un âne. Geach conlut :

> Obviously you are indeed a donkey if you agree to such a stipulation. Buridan rightly says that such a stipulation is to be rejected out of hand: the matter of what 'A' stands for must already be settled before we can even consider whether 'A is a donkey' is true or false. (P. T. GEACH, 1987, p. 10)

On voit bien, en effet, que déterminer la valeur de vérité d'une expression n'est possible que si les références des termes sont préalablement fixées. Si l'on attribue une propriété à une chose, on le fait précisément au sujet de cette chose, et non de manière générale.

Comme y invite Kai Wehmeier (WEHMEIER, 2008), nous pouvons rapprocher ceci des propositions existentielles négatives. La proposition exprimée par la phrase « a n'existe pas » est soit vraie, soit fausse. Dans le premier cas, elle ne concerne aucun objet. Dans le second, elle en concerne un. On voit ainsi que sa valeur de vérité fait fluctuer l'objet sur lequel porte l'énoncé, ce qui est évidemment une entorse au principe de Buridan. C'est justement ce qui fait dire à Frege que l'existence n'est pas une propriété objectuelle, mais une propriété de second ordre, une propriété conceptuelle [32]. Or nous pourrions revendiquer le même genre de situation pour les énoncés

[32] Nous reviendrons avec plus d'insistance sur cette question dans le chapitre 3.

d'identité. Si un énoncé d'identité $a = b$ est vrai, il ne concerne qu'une seule personne. Autrement dit, les deux termes mis en relation co-dénotent le même objet. En revanche, si l'énoncé est faux, il ne concerne pas le même objet. On voit donc qu'il y a ici encore une entorse au principe de Buridan qui peut encourager à considérer que l'identité n'est pas une relation objectuelle mais conceptuelle. On pourrait bien sûr objecter qu'il reste possible de conférer à l'identité une place spéciale, à savoir, d'être la seule relation à violer la loi de Buridan, mais il est clair qu'une telle solution paraît complètement taillée sur mesure et donc peu convaincante.

La théorie métaphysique de l'identité relative

Peter Geach a souligné à différentes occasions les liens qui unissent la théorie de l'identité relative et la théorie de la cardinalité de Frege. Geach s'étonne même que Frege, qui a bien montré le caractère sortalement relatif de nos dénombrements, n'ait pas fait le pas supplémentaire d'affirmer le caractère sortalement dépendant de l'identité :

> Frege sees clearly that "one" cannot significantly stand as a predicate of objects unless it is (at least understood as) attached to a general term; I am surprised he did not see that the like holds for the closely allied expression "the same". (P. T. GEACH, 1962, p. 64)

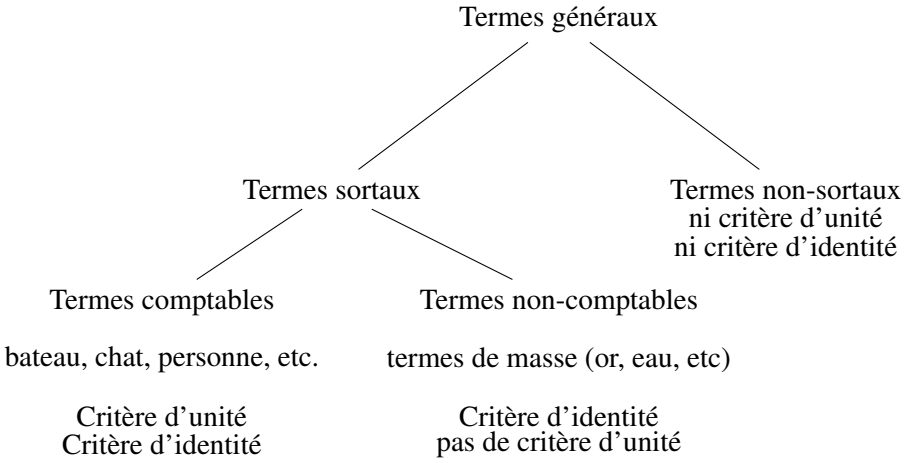
Rappelons que nous avons déjà dit que, pour les partisans de l'identité objectuelle, la connaissance des critères d'individuation d'une chose dépend de la capacité à employer le concept sortal sous lequel tombe cette chose. La dépendance est donc, dans ce cas, seulement épistémique, indépendamment de la question de savoir si les sortes relèvent d'entités objectives comme les universaux ou si elles n'ont d'existence que conceptuelle.

Comme nous l'avons déjà mentionné, selon les sortalistes, pour tout particulier donné, il existe un sortal « ultime » qui en délivre les conditions d'identité. Il s'agit alors de distinguer d'une part les termes sortaux de phase qui peuvent diversement se dire d'un même individu, comme par exemple le terme « enfant », et d'autre part les sortaux « ultimes » qui fixent les conditions d'identité du particulier considéré, comme par exemple « humain ».

Pour revenir au paradoxes de la statue et du bloc d'argile, sous cette perspective, il n'y a pas une chose qui serait à la fois une statue et un bloc d'argile, car pour des conditions de persistance différentes comme celles qui régissent la sorte « statue » et la sorte « bloc d'argile », on doit avoir deux choses différentes. En effet, les termes « statue » et « bloc d'argile » renvoient à des sortaux ultimes. Dès lors, il convient de dissocier deux entités coïncidentes, d'une part une statue et d'autre part un bloc d'argile. Bien que ces deux entités cohabitent sur le même emplacement spatio-temporel, les partisans de la thèse de dépendance sortale refusent de les identifier, dupliquant ainsi chaque artefact par un bloc de matière qui le compose.

Comme les conditions d'identité d'une chose ne sont généralement pas les mêmes que les conditions d'identité de sa constitution matérielle, cela implique plus généralement la thèse de désolidarisation de l'identité et de la constitution stipulant que l'identité n'est pas la constitution matérielle. Or il s'agit précisément de ce que rejette Geach, pour qui l'existence d'un sortal ultime n'a pas plus de sens que celle d'une identité absolue. Comment reconnaître qu'un prédicat F est un terme sortal ? Selon Geach, nous disposons de deux caractéristiques. D'abord, il permet de

déterminer les conditions d'identité de ce qu'il subsume. Ensuite, un tel terme est reconnaissable à ce qu'il fournit des informations permettant de répondre à la question « combien y a-t-il de F ? ». Ainsi, « dure » et « ronde » ne sont pas des termes de sortes, car ils ne déterminent pas les conditions d'identité d'une chose qui peut demeurer la même tout en changeant de densité ou de forme. Une question comme « combien d'eau » sur la table pourrait s'avérer sans réponse. On opposera ainsi les termes de sorte (comme les cailloux) aux termes de masse (comme l'eau). On a alors une répartition des termes généraux selon le diagramme suivant.



Reconnaître une catégorie « F » présuppose là aussi une identification, puisqu'elle correspond à une segmentation du monde, que celle-ci soit réelle ou conventionnelle. Nos concepts véhiculent ainsi différents cribles qui régissent nos aptitudes à dénombrer, ce qui semble être une proposition raisonnable. Mais Geach ne se contente pas de soutenir une thèse épistémique de dépendance sortale, il propose une version encore plus forte, stipulant que ce n'est pas la connaissance des conditions d'individuation d'une chose qui est sortalement dépendante, mais directement les conditions d'individuation elles-mêmes. Selon Geach, le rapport de dépendance des individus aux sortes n'est donc pas seulement épistémique, il est métaphysique.

Du point de vue formel, l'idée consiste à indexer chaque signe d'égalité par un sortal F. Remarquons que cette stratégie maintient l'identité relative comme une relation d'équivalence, car elle demeure bien réflexive, symétrique et transitive :

- $\forall x\ (x =_{\mathrm{F}} x)$
- $\forall x\ \forall y\ (x =_{\mathrm{F}} y) \rightarrow (y =_{\mathrm{F}} x)$
- $\forall x\ \forall y\ \forall z\ (x =_{\mathrm{F}} y \wedge y =_{\mathrm{F}} z) \rightarrow (x =_{\mathrm{F}} z)$

De même, l'indiscernabilité des identiques prend la forme suivante :

$$\forall x\ \forall y \left(x =_{\mathrm{F}} y \rightarrow \forall \mathrm{P}\ (\mathrm{P}x \leftrightarrow \mathrm{P}y) \right)$$

Dans ce cas, on peut admettre la possibilité suivante : $\exists x\ \exists y \left(x =_{\mathrm{F}} y \wedge x \neq_{\mathrm{G}} y \right)$.

(Thèse de relativité sortale de l'identité) Il est possible qu'il existe deux objets x et y dont chacun relève à la fois des sortes F et G et tels que x est le même F que y et x n'est pas le même G que y.

Cette thèse permet à Geach de résoudre les paradoxes de constitution matérielle. Prenons le cas de la statue et du bloc d'argile qui la compose. En perdant un bras, l'objet reste le même sous la sorte statue, mais diffère sous la sorte bloc d'argile. Considérons encore le cas de deux statues sculptées successivement dans le même bloc d'argile. Nous avons là des individus qui tombent à la fois sous le concept sortal de statue et sous le concept sortal de bloc d'argile, et qui, en outre, sont des individus identiques sous l'aspect de bloc d'argile, mais des individus différents sous l'aspect de statues. D'un point de vue épistémique, cette théorie confirme la double possibilité d'appréhension de l'objet, comme statue, et comme bloc de matière.

Cette solution paraît tentante en ce qu'elle répond à l'idée que toute chose est toujours identifiée sous un certain aspect, par un mode de présentation. Néanmoins, elle contredit directement le principe d'indiscernabilité des identiques. En effet, d'après ce principe, si un objet x est le même qu'un objet y sous la sorte F, x et y sont indiscernables, et doivent donc aussi être identiques en rapport à la sorte G. Faisons apparaître le problème d'une manière plus formelle. Admettons qu'il existe deux objets x et y tels que $x =_{\mathrm{F}} y$ et $x \neq_{\mathrm{G}} y$:

(1)	$x =_{\mathrm{F}} y$	hypothèse
(2)	$x \neq_{\mathrm{G}} y$	hypothèse
(3)	$x =_{\mathrm{G}} x$	par réflexivité de l'identité
(4)	$x =_{\mathrm{F}} y \rightarrow (\mathrm{P}x \leftrightarrow \mathrm{P}y)$	par application de l'indiscernabilité des identiques à (1)
(5)	$x =_{\mathrm{F}} y \rightarrow (x =_{\mathrm{G}} x \leftrightarrow x =_{\mathrm{G}} y)$	par (4), en posant P := $(x =_{\mathrm{G}})$
(6)	$x =_{\mathrm{G}} y$	Par (1), (3) et (5)
(7)	$x =_{\mathrm{G}} y \wedge x \neq_{\mathrm{G}} y$	Contradiction par (2) et (6)

Pour défendre la théorie de l'identité relative en bloquant l'issue de l'argument, on doit alors rejeter le principe d'indiscernabilité des identiques, afin de bloquer le passage de (4) à (5). La théorie de Geach veut rendre compte du caractère perspectiviste de l'identité : une identité est toujours affirmée à propos d'une sorte d'objet. Pour ce faire, la thèse de relativité sortale de l'identité conteste l'idée que l'identité implique la ressemblance parfaite. L'identité relative implique au contraire une indiscernabilité qui n'est que relative à un certain aspect, celui qui est induit par la sorte F sous laquelle la chose est considérée. L'indiscernabilité n'est donc pas totale, mais ne touche qu'un ensemble de propriétés à l'exclusion d'autres qui peuvent demeurer différentes.

Dès lors, il devient malheureusement très délicat de déterminer les propriétés que nous pouvons dériver d'une identité entre deux choses. Cette lourde entorse à la caractérisation classique de l'identité a pour conséquence que la théorie métaphysique de l'identité relative est peu acceptée par les philosophes, la plupart préférant s'en tenir à la version épistémique de la dépendance sortale. En refoulant la conception perspectiviste de l'unité sur l'identité logique, plutôt qu'en questionnant la légitimité de l'identité ontologique, la théorie de Geach conduit en effet à une révision de la logique bien trop coûteuse.

Nous retiendrons cependant de la théorie de Geach l'effort qu'elle produit pour rendre compte du fait qu'un objet nous apparaît toujours sous un certain aspect, et que ses conditions d'identification dépendent de cet aspect. Mais nous ne rejetterons pas pour autant le caractère absolu de l'identité logique. Ce que nous devons faire, ce n'est donc pas relativiser l'identité logique aux sortes, mais contextualiser l'application que nous faisons de l'identité. Cela implique de renoncer à des critères métaphysiques d'individuation, autrement dit, de renoncer à l'identité objectuelle, au profit de thèses plus conventionnalistes.

1.4.3 Conventionnalisme de l'unité

Segmentations métaphysiques et épistémiques

Les paradoxes évoqués précédemment ressortissent aux questions de l'individuation, autrement dit, ils portent sur les questions métaphysiques afférentes à l'unicité et à l'unité. A première approximation, le point de vue métaphysique sur l'identité pourrait nous amener à croire que les choses sont à la fois uniques et à la fois multiples, qu'elles sont identiques ou distinctes seulement de manière vague. Cela pourrait inciter à la croyance que les choses sont fondamentalement doubles parce qu'elles ne peuvent ni se tenir à la catégorie de l'unique, ni à celle du pluriel. Une telle conception de l'identité semble parfaitement résumée par le fragment d'Héraclite : « De toutes choses l'un, et de l'un toutes choses » (CONCHE, 1986, fragment 127). Car de ce point de vue, l'unité d'une chose provient paradoxalement de la pluralité de ses constituants. Or ces derniers peuvent à leur tour être saisis comme des unités et ainsi de suite. L'unité et la multiplicité semblent alors être les pôles contraires d'un processus dialectique sans fin se jouant à l'intérieur des choses.

Mais plutôt que de céder à ce genre d'argument, nous pourrions avoir de bonnes raisons de penser que l'individuation n'est pas un fait métaphysique mais ressortit à de simples conventions de langage.

La conception déflationniste de l'ontologie est renouvelée aujourd'hui par des philosophes comme Eli Hirsch (HIRSCH, 1992) ou comme Mati Eklund (EKLUND, 2009), qui affirment que Carnap aurait en réalité été partisan d'un pluralisme linguistique mettant l'accent sur le fait qu'il existe divers systèmes linguistiques et que la signification et la valeur de vérité d'un énoncé peut varier selon le système linguistique retenu. De ce point de vue, la notion d'existence ne serait pas une notion absolue, mais relative à des schèmes linguistiques. Cela implique que les discussions ontologiques, au sens externe de Carnap, ne sont pas creuses, elles sont seulement sans fondement. En particulier, Hirsch défend la thèse appelée variance quantificationnelle qui affirme que les disputes au sujet de ce qu'il y a sont sont des disputes purement verbales. Ce qui n'est pas reconnu comme une entité par tel langage peut cependant l'être par un autre langage. Selon Hirsch, ce qui est préconisé par Carnap, ce n'est pas vraiment d'évincer les questions métaphysiques hors de la philosophie, mais de les considérer comme des façons de parler étroitement liées à des conventions linguistiques. La défense d'un égalitarisme des quantifications se trouve aussi exprimée chez Hilary Putnam, notamment quand ce dernier soutient par exemple que :

> [A]ll situations have many different correct descriptions, and [...] even descriptions that, taken holistically, convey the same information may differ in what they take to be 'objects'

> [...] There are many usable extensions of the notion of an object [...] The logical primitives themselves, and in particular the notions of object and existence, have a multitude of different uses rather than one absolute "meaning". (PUTNAM, 1987, p. 71)

Dans *The Concept of Identity*, Eli Hirsch pose la question de savoir pourquoi le langage ordinaire contient des critères d'unité. En effet, nous disons par exemple, voilà un stylo, voilà une table, mais nous ne parlons généralement pas de leur somme méréologique, le stylo-table. S'agirait-il de quelque chose comme une entité fictionnelle ? Nous pouvons bien sûr prononcer les mots « la table et le stylo forment une unité ». Mais est-ce suffisant pour considérer le stylo-table comme une chose authentique ? Hirsch fait remarquer que le langage naturel, s'il permet d'un côté de parler d'objets ordinaires comme les disques, est d'un autre côté rétif à reconnaître l'existence d'objets plus fantaisistes comme les disques-dedans ou les disques-dehors. Mais rien n'empêche l'existence d'autres langages, associés à d'autres critères d'identité, langages qui donneraient raison à l'existence de choses comme les disques-dedans. Quand un usager du langage naturel dirait qu'on sort un disque de sa pochette, un usager de cet autre langage pourrait dire qu'un disque-dehors s'accroît progressivement quand, dans le même temps, s'amenuise un disque-dedans.

Plus précisément, Hirsch imagine ainsi le « M-langage » qui ne serait associé à aucun critère habituel de persistance, mais qui aurait néanmoins la même capacité que le langage naturel (N-langage) pour parler de ce qui est. Toute affirmation dans le N-langage au sujet de la persistance d'une chose serait retraduite dans le M-langage en expliquant comment sont reliées entre elles certaines phases temporelles. Lorsque nous disons, dans le N-langage, qu'*un* (même) disque a été sorti de sa pochette, nous pourrions dire dans le M-langage qu'une phase de disque était dans une phase de pochette, et ensuite qu'il y a eu une succession de phases de disques-dedans et de phases de disques-dehors, et enfin, qu'est apparu une phase de disque à l'extérieur de la pochette. Selon Hirsch, le N-langage ne serait pas plus véridique que le M-langage, il exprimerait les mêmes états de choses, mais en les décrivant de manières différentes. En ce sens, les discussions ontologiques ne sont pas nécessairement creuses, mais elles sont sans fondement dans la mesure où différentes ontologies comme l'universalisme méréologique ou le nihilisme méréologique sont toutes aussi légitimes les unes que les autres.

Cette thèse, qualifiée de variance quantificationnelle, relève sans doute d'une forme de conventionnalisme. Il reste à savoir alors quelles sont les raisons qui nous font préférer l'usage d'un langage dans lequel nous disposons de critère de persistance plutôt que l'usage d'un langage dans lequel nous n'en disposons pas. Ensuite, on peut se demander pourquoi notre langage véhicule certains critères d'unité plutôt que d'autres.

> [...] there is in fact no reason why our language had to contain just those criteria of unity, but that this is nothing more than an arbitrary convention of language. Our ordinary body-mindedness, according to this "conventionalist" position, is merely one scheme for conceptually dividing the world into units, and any number of other schemes might have done just as well. The scheme that we have gets passed on from generation to generation, in the manner suggested by Quine. (HIRSCH, 2011, p. 10)

Pourtant, Hirsch prétend que sa thèse n'est pas conventionnaliste au sens où la réalité ne serait que pure convention. En revanche, elle implique que les critères d'unité ne sont pas des propriétés métaphysiques des choses, mais des cribles cognitifs.

Affirmer que l'identification d'un objet nécessite toujours une attribution conceptuelle préalable implique que si l'on désigne une portion de l'espace-temps au seul moyen d'un geste ou d'un terme indexical comme « ça ! », il n'est cependant pas possible de suivre la carrière temporelle de ce qui est désigné. Car nous devrions pour cela inférer que ce qui est désigné relève d'une certaine catégorie de choses. En effet, qu'est-ce qui est pointé par un acte de désignation qui indique l'emplacement d'une table ? Est-ce le plateau ou les pieds de la table ? Sont-ce ses molécules ? Ou seulement la forme « ronde » ? Peut-être même n'est-ce que la portion d'espace, indépendamment de l'objet concret qui s'y trouve. Comme l'écrit Leszek Kołakovski :

> The picture of reality sketched by everyday perception and by scientific thinking is a kind of human creation (not imitation) since both the linguistic and the scientific division of the world into particular objects arise from man's practical needs. In this sense the world's products must be considered artificial. In this world the sun and stars exist because man is able to make them his objects, differentiated in material and conceived as "corporeal individuals". In abstract, nothing prevents us from dissecting surrounding material into fragments constructed in a manner completely different from what we are used to. Thus, speaking more simply, we could build a world where there would be no such objects as "horse", "leaf", "star", and others allegedly devised by nature. Instead, there might be, for example, such objects as "half a horse and a piece of river", "my ear and the moon", and other similar products of a surrealist imagination. (KOŁAKOVSKI, 1969, pp. 47-48)

Nous avons déjà mentionné, dans l'introduction générale, un petit exercice de gymnastique orthoptiste destiné à la rééducation des patients sujets à un strabisme divergent. Il s'agit d'observer deux images, séparées et tendues à bout de bras, en focalisant le regard de manière à ne voir qu'une seule image plutôt que deux. Autrement dit, il s'agit de faire converger les deux images de façon à ce qu'elles se superposent (voir figure 1.6). Pour différentes raisons pratiques, il nous arrive ainsi de percevoir, d'imaginer ou de concevoir que ce qu'on tient dans certaines circonstances pour deux choses ne compte dans d'autres que pour une seule. Et inversement, il nous arrive aussi de considérer que ce qu'on tient parfois pour une chose compte en d'autres circonstances, pour deux choses. Nous pouvons compter mille pissenlits dans un même champs, parce que nous pourrions en arracher les mille tiges. Mais un biologiste n'en compterait sans doute qu'un seul [33], relevant d'une certaine entité parsemée témoignant d'un même code génétique entre ses différentes occurrences spatiales non-connexes.

On peut donc souscrire à l'idée que l'individuation est purement conventionnelle, qu'elle n'est qu'un fait de langage, et non une réalité métaphysique. Parmi les partisans historiques de cette thèse, Nietzsche soutient que le langage est exclusivement métaphorique. Les mots sont des conventions, ils découpent le monde selon des lignes qui sont historiques et non éternelles. Les concepts masquent la réalité par une volonté, par un parti pris proprement humain. Il transforme un ordre de chose en un autre.

> Ce qui nous sépare le plus radicalement du platonisme et du leibnizianisme c'est que nous ne croyons plus à des concepts éternels, à des valeurs éternelles, à des formes éternelles, à des âmes éternelles, et la philosophie, dans la mesure où elle est scientifique et non dogmatique,

[33] A ce sujet, on pourra se référer à (PRADEU, 2008).

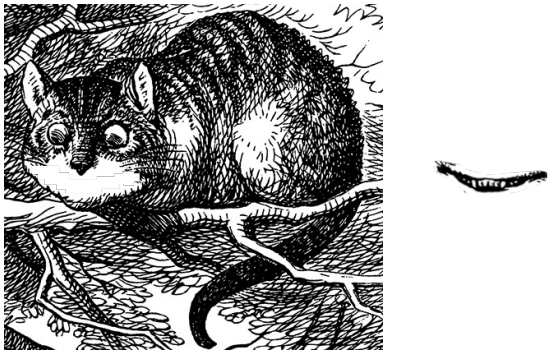


FIGURE 1.6 : Réunir les deux images en une seule

> n'est pour nous que l'extension la plus large de la notion d'« histoire ». L'étymologie et l'histoire du langage nous ont appris à considérer tous les concepts comme *devenus*, beaucoup d'entre eux comme encore en devenir, de telle sorte que les concepts les plus généraux étant les plus faux, doivent aussi être les plus anciens. (NIETZSCHE, 1965, p. I, 112)

Les concepts généraux ne sont pas des mots qui indiqueraient des choses du monde ou des propriétés métaphysiques. Leur fausseté et leur ancienneté s'accroît en rapport avec leur généralité. Aussi, Nietzsche en conclut que :

> Il est infiniment plus important de connaître le nom des choses que de savoir ce qu'elles sont [...]. Il suffit de forger des noms nouveaux, de nouvelles appréciations et de nouvelles probabilités pour créer à la longue aussi des « choses nouvelles ». (NIETZSCHE, 1982, § 258)

Toute philosophie doit ainsi se faire à coups de marteau, autrement dit, elle doit consister à fendre nos concepts afin de voir ce qu'il contiennent « d'humain, trop humain ». De ce point de vue, le langage est un double, un usurpateur qui prétend occuper la place du réel.

Clément Rosset semble suivre la même ligne de pensée lorsqu'il affirme une imperméabilité quasi-complète du langage à l'égard du réel. Il est une exception notable toutefois, la tautologie qu'exprime l'énonciation que « le réel est le réel ». Comme l'écrit Rosset dès les premières lignes de *Le démon de la tautologie* :

> J'appellerai ici réel, comme je l'ai toujours fait au moins implicitement, tout ce qui existe en fonction du principe d'identité qui énonce que A est A. J'appelle irréel ce qui n'existe pas selon ce même principe : c'est-à-dire non seulement tout ce qui ne fait parade d'existence que sous le mode de l'imaginaire ou de l'hallucination, mais aussi et plus précisément ce qui semble bénéficier du privilège de l'existence mais qui se révèle illusoire à l'analyse pour ne pas répondre rigoureusement au principe d'identité (soit tout A qui ne se résume pas à l'A qu'il est, mais connote de certaine façon un B qu'il pourrait être aussi, voire plus subtilement un autre « A » [...]). (ROSSET, 1997, p. 11)

Ce qui est intéressant dans la conception de Rosset, c'est qu'il montre qu'on peut accréditer l'idée que le Réel est une entité impossible à saisir, sans nécessairement souscrire à un pluralisme

des réalités [34]. Il y a bien une unique réalité, mais nos mots échouent à la qualifier. Cet échec épistémique à saisir le réel, sans pour autant en évincer la place, pourrait nous inciter à considérer que ce que nous appelons individus relève de segmentations conventionnelles plutôt que de segmentations métaphysiques.

Dans les *Fondements de l'arithmétique*, Frege soutient qu'*une* paire de bottes et *deux* bottes constituent *le même* phénomène physique et sensible (FREGE, 1969a, p. 152). Or nous devons pourtant constater une différence numérique puisque nous disons d'un côté *une* paire et d'un autre, *deux* bottes. Ailleurs, Frege écrit

> While looking at one and the same external phenomenon, I can say with equal truth both "It is a copse" and "It is five trees", or both "Here are four companies" and "Here are 500 men". (FREGE, 1969a, Section 84)

Une telle manière de voir les choses pourrait encourager, *pace* Frege, un point de vue conceptuel de la notion d'unité. En outre, comme nous l'avons vu, des philosophes comme Geach ou Wiggins prétendent que nous ne pouvons pas avoir un accès direct aux particuliers. Leurs théories respectives préconisent que la référence à un individu se fait toujours par la médiation d'un sortal. Plus précisément, ils affirment que nous ne pouvons tracer la carrière d'une chose à travers le temps qu'à la condition de la reconnaître d'emblée comme *une certaine sorte* de chose. Chez ces philosophes, la notion de sorte est rattachée à une conception réaliste des universaux. Mais la question peut sans doute se poser dans ces termes en adhérant à un strict point de vue conceptualiste pour lequel les prédicats ne sont que des entités mentales, voire pour un point de vue nominaliste pour lequel ils ne relèvent que d'entités linguistiques.

Le conventionnalisme modéré

La question est alors de savoir si nous pourrions ou non percevoir quelque chose qui échapperait complètement à nos concepts. Y a t-il des modes de référence qui sont indépendants de toute catégorisation sortale ?

Stéphane Chauvier constate qu'il y a deux manières de poser la question de l'individuation. La première consiste à concevoir l'individu comme une unité numérique, unité appartenant à une sorte donnée, que Chauvier désigne sous le nom d'« hénade ». C'est ce qui se produit quand on désigne quelque chose comme étant *un* bateau, plutôt que *du* bateau. La seconde manière de concevoir l'individu consiste à le voir comme « un absolu singulier, fermé sur le reste du monde » (CHAUVIER, 2010, p. 3), c'est-à-dire comme une « monade ». C'est ce qui se produit lorsqu'on pointe un bateau en disant *ce bateau*. Or selon Chauvier, il n'est pas raisonnable de penser parvenir à sélectionner une monade (*ce* bateau) sans avoir préalablement attrapé celle-ci comme une certaine sorte de chose (bateau), pour qui on ne peut envisager « un penseur qui serait sensible à l'individualité des présences, mais pas à leur identité spécifique ou générique, quelqu'un qui verrait des individus nus, sans essence ou nature. » (CHAUVIER, 2010, p. 3).

À titre d'illustration, rappelons que Lovecraft, qui se refuse généralement à décrire précisément les monstrueuses créatures qui peuplent ses récits, fait un usage immodéré de l'adjectif

[34] La conception défendue par Rosset semble d'ailleurs exactement opposée à un pluralisme, tant on pressent chez cet auteur l'idée que le Réel ne peut pas être saisi de manière fragmentaire, mais seulement comme une totalité hénadique.

« indicible » pour en désigner l'aspect. Ainsi, dans la nouvelle justement intitulée *L'indicible*, Lovecraft désigne la créature qui vient de tourmenter les protagonistes de sa nouvelle de la manière suivante : « C'était partout – une sorte de gélatine – de gelée – et qui pourtant avait des formes – mille formes si horribles... dépassant toute description. Il y avait des yeux – et une souillure... C'était la fosse, c'était le maelström, c'était l'abomination ultime, [...] *c'était l'indicible* !... » (LOVECRAFT, 2000, p. 174). Ce qui ne peut être dit, parce que ne tombant sous aucun concept, parce qu'échappant aux dimensions habituelles de l'espace, étant « partout », peut cependant encore être indiqué par des indexicaux.

John Campbell constate qu'il y a bien des choses qui ne sont pas d'emblée rapportées à des concepts, car notre connaissance du monde est d'abord une expérience perceptive. Campbell cite le cas des couleurs :

> At an early stage, then, children do not yet have colour concepts. But since they have colour vision, they can use the colour of an object as one of the characteristics that differentiates it from its background; they can use colour as an object-defining characteristic without yet being able to verify propositions about the colour of the object. (CAMPBELL, 2002, p. 30)

L'identification visuelle, rappelle en effet Campbell, est antérieure à la connaissance conceptuelle. Les gestes de désignation et les termes indexicaux comme « ceci » ou « cela » sont les outils qui nous permettent une focalisation de l'attention. Cette dernière offre la capacité de tracer un objet sans pour autant que nous en ayons une connaissance conceptuelle.

Mais que serait un objet lovecraftien échappant à toute catégorisation sortale ? Un objet qui ne relèverait ni de l'animal, ni de l'humain, ni de l'inanimé ? Pour être pensé, une telle chose devrait encore être vue *comme un objet*. Cependant, s'il parait vraisemblable que l'acquisition de bon nombre de nos concepts passe par des ostensions répétées, il reste à savoir comment nous acquérons le concept même d'objet à partir de telles répétitions. Il semble difficile d'admettre que l'unité soit exclusivement un fait linguistique, un trait culturel que nous pourrions échanger pour d'autres manières d'appréhender le monde, tant nos schémas cognitifs apparient l'unité à la cohésion motrice et à la connexité.

Les termes « objet », « chose » sont généralement traités comme des termes catégoriques, qui ne spécifient pas une sorte particulière, autrement dit, des termes qui ne permettent pas le dénombrement. Mais il y a sans doute différentes acceptions du terme « objet ». Du point de vue commun, un objet physique est un objet qui peut être déplacé en décrivant une trajectoire continue à travers un espace à trois dimensions. Selon Fei Xu (XU, 1997), on peut considérer que les enfants aussi bien que les adultes conçoivent la notion d'objet physique comme un terme sans doute plus général que « chat », « bateau » ou « personne », mais qui relève néanmoins d'une sorte. Les objets sont des choses qui, contrairement aux ombres, peuvent être heurtées, voire soulevées.

Dans l'article intitulé *From Lot's Wife to a Pillar of Salt : Evidence that Physical Object is a Sortal Concept* (XU, 1993, p. 366) Fei Xu fait remarquer que, vu d'un point de vue ontologique, la sémantique développée par Russell, qui réduit la catégorie logique des noms propres aux déictiques, pourrait impliquer une quantification sur des particuliers nus :

> The logic of sortals contrasts with standard model-theoretic semantics, in which the logical form for 'Some person is tall' in predicate calculus is '$\exists x($ person$(x) \wedge$ tall(x)'. In this

> formulation, x is a bare particular - it is an individual with no properties of its own but it stiIl supports properties such as 'being a person' or 'is tall'. (XU, 1993, p. 366)

En écrivant par exemple « $\exists x$ Px », on attribuerait à un certain particulier x, préalablement dépourvu de qualités, la propriété P. Le sortaliste s'oppose à une telle conception : les individus qui font l'objet de quantifications sont nécessairement préalablement identifiés par le fait de posséder quelque propriété. Macnamara et La Palme (MACNAMARA, REYES & LA PALME REYES, 1994) proposent par exemple de formaliser l'assertion « cet homme est grand » par l'expression formelle $\exists(x : \text{homme})(\text{ grand}(x : \text{homme}))$. De ce point de vue, il n'y a pas de x qui soit indépendant d'une certaine sorte de choses.

Déterminer si la notion d'identité est seulement conceptuelle, voire conventionnelle passe sans doute par le fait de savoir comment les enfants apprennent de tels critères. En psychologie du développement, des investigations ont été faites afin de déterminer quels critères sont employés pour individuer des objets et décider si nous avons affaire à une chose ou à deux choses. Les premiers travaux remontent à Jean Piaget (PIAGET, 1970) et T.G.R Bower (BOWER, 1974), ils sont développés actuellement par Olivier Houdé (HOUDÉ, 2004). De nombreuses expériences ont ainsi été menées sur des bébés pour évaluer leurs réactions face à des expériences troublantes de dénombrement. Le paradigme expérimental de l'habituation consiste à mesurer la durée de fixation oculaire de l'enfant.

D'après Jean Piaget et Alina Szeminska (PIAGET & SZEMINSKA, 1997), les expériences menées sur des enfants âgés de 2 à 4 ans montrent qu'à cet âge, la seule causalité qui est attribuée aux objets demeure « psychomorphique » (PIAGET, 1970, p. 26). Les concepts de classe et de sous-classe sont progressivement construits par le sujet connaissant. Si des enfants en bas-âge parviennent à différencier deux ensembles différents (comme, d'une part, des jetons ronds et rouges, et d'autre part des jetons bleus et carrés), leur capacité à distinguer les relations entre les différents individus d'une même classe est en revanche plus lente dans son élaboration. La différence entre un individu, représentant d'une classe, et la classe elle-même demeure pour l'enfant confuse pendant une certaine période. Piaget constate par exemple que :

> En de nombreuses situations de la vie courante [l'enfant] aura peine à distinguer en face d'un objet ou d'un personnage x s'il s'agit d'un même terme individuel x demeuré identique à lui-même ou d'un représentant quelconque x ou x' de la même classe X : l'objet demeure ainsi à mi-chemin de l'individu et de la classe par une sorte de participation ou d'exemplarité. Par exemple, une petite fille Jacqueline, voyant une photographie d'elle plus jeune dit que "c'est Jacqueline quand elle était une Lucienne (= sa sœur cadette)". (PIAGET, 1970, p. 150)

La conception de Piaget indique que l'enfant constitue d'abord la notion de classe, et seulement après qu'il constitue celle d'individu persistant dans le temps. Pour Piaget, la notion d'objet n'est pas celle d'une entité autonome indépendante du sujet, elle doit être rapprochée des *sense-data* : un enfant n'est pas surpris de la disparition d'un objet dont la trajectoire l'amène derrière un écran occultant.

Cette conception est battue en brèche depuis sa formulation. Les expériences de Piaget ont été contestées notamment par T.G.R. Bower :

> The evidence that infants do segregate their environments into units is clear. A large number of studies on the eye-fixation behavior of infants has shown that infants will fix on the external contours of objects in their visual field. If the objects are moved, the infants will track them. If after moving together, the contours of an object break and begin to move independently, very young infants will display massive surprise. This indicates that the common motion (common fate) had specified for them a single unit. (BOWER, 1974, p. 102)

Bower a montré que l'identité relève de la continuité spatio-temporelle des positions d'un objet en déplacement. En revanche, la catégorie sortale de l'objet ne rentre pas dans les conditions d'identification de l'enfant de bas-âge. Selon Bower, une balle qui est substituée, durant son déplacement, à un lapin en peluche, continue cependant d'être identifiée par l'enfant à la peluche. Quand un objet est déplacé suivant une trajectoire qui l'amène derrière un écran, puis que l'écran est abaissé en révélant à sa place un objet d'un autre type, le changement ne semble pas provoquer d'effet de surprise chez le bébé spectateur. La continuité de forme, ou de ressemblance ne semble donc pas jouer le rôle d'un critère d'identification.

Les résultats des expériences de menées dans les années 1980 par Strauss et Curtis (STRAUSS & CURTIS, 1981) semblent indiquer que des bébés ont la capacité pré-langagière à discriminer parfaitement entre 2 et 3 objets, que cette capacité devient seulement partielle avec 3 et 4 objets, et qu'ils s'avèrent incapables de distinguer entre 4 et 5 objets. Karen Wynn (WYNN, 1992) prétend montrer avec des expériences menées sur des bébés que des individus âgés entre quatre et cinq mois ont déjà l'intuition qu'un objet ajouté à un objet est égal à deux objets. Les expériences de Wynn consistent à déplacer, devant des enfants, des marionnettes sur un théâtre, objets qui sont ensuite cachés derrière un écran relevé, occultant la scène. Un opérateur retire ou ajoute subrepticement une marionnette derrière l'écran, puis abaisse ce dernier. On procède alors à la mesure de la réaction de surprise de l'enfant. Wynn conclut de telles expériences que des enfants de quatre à cinq mois possèdent des représentations du nombre qui ne sont pas seulement réductibles à une simple dichotomie entre l'unique et le multiple.

Les conclusions qu'avance Wynn doivent sans doute être relativisées, car l'étonnement peut en effet résulter d'autres facteurs que d'un jugement de distinction numérique, autrement dit d'une capacité innée à l'arithmétique, mais par exemple aussi à l'appréhension perceptive de la persistance des objets matériels. En effet, pourquoi ne pas interpréter les résultats de Wynn comme un raisonnement perceptif du bébé plutôt qu'un raisonnement cognitif comme elle le suppose ? Jacqueline Bideaud (BIDEAUD, 2001), interprète les résultats de Wynn à partir d'un mécanisme particulier dont disposeraient les bébés, le *subitizing*, qui leur permettrait une reconnaissance immédiate de configurations spatiales liées à de petits entiers. Le bébé ne compterait pas, mais associerait les cardinaux à des formes privilégiées comme par exemple la configuration des points sur les faces d'un dé.

Eli Hirsch (HIRSCH, 1992) soutient l'idée que la notion d'individu doit être première, qu'il s'agit d'une notion innée, elle n'est donc ni objectuelle ou ni purement conventionnelle. L'appréhension de l'unité d'un objet, ou plus simplement, l'appréhension d'un objet en tant que tel est une capacité innée. Selon Hirsch les critères d'unité relèvent de la constitution de notre appareil cognitif. Les objets ordinaires respectent des critères spatiaux de connexité (ils ne sont pas séparés d'eux-mêmes) et de cohésion dynamique (leurs parties se meuvent de manière dépendantes).

Nous sommes inclinés à identifier un même et unique objet parce que ses parties sont connexes et qu'elles se déplacent en conservant leur structure. Selon Hirsch, ce sont ces deux facteurs qui nous font dire qu'il y a là un objet et non deux. D'un côté, la connexité assure à l'objet qu'il ne soit pas éclaté dans l'espace, mais ramassé en parties entretenant des rapports de proximité entre elles. D'un autre côté, la cohésion dynamique assure en outre que ces rapports soient entretenus de manière temporelle. C'est donc la synchronicité de déplacements des différentes parties d'une chose qui fait qu'elle est identifiée comme *un* objet plutôt que deux.

Selon Hirsch, la notion d'unité doit ainsi être rapportée à notre capacité à concentrer notre attention sur un objet, capacité qui apparaît déjà chez les bébés comme le montre l'observation de leur regard. Lorsqu'on déplace un objet devant eux, les bébés ont généralement tendance à le suivre des yeux plutôt que de focaliser leur attention sur des portions arbitraires de l'espace. La capacité à maintenir notre concentration sur un objet en déplacement semble innée, et Hirsch la rapproche de ce que Quine appelle « un sens inné de la similarité ». En effet, le bébé qui suit du regard un objet tend à associer mentalement différentes portions de l'espace-temps en les corrélant en raison des similarités complexes qu'elles manifestent, et qu'il apprendra plus tard à désigner sous le nom d'« objet mouvant ».

De fait, les segmentations du monde en certaines unités, désignées comme « objets » sont sans doute encouragées par divers phénomènes. Parmi ceux-là, la connexité, la séparabilité, la continuité spatio-temporelle et la continuité matérielle, mais aussi la synchronicité de déplacement des parties figurent sans doute parmi les raisons les plus fortes de préférer certains décomptes à d'autres. Le changement d'unité par la modification de l'ensemble unité est d'ailleurs un enjeu majeur de l'apprentissage du dénombrement. Par exemple, pour de nombreux élèves du cycle 3, il reste difficile de déterminer à quoi correspond au juste la moitié de l'ensemble des jetons représentés ici :



Les réponses attendues, qu'il s'agisse de la moitié de l'ensemble des parties,



voire de la moitié des parties de l'ensemble :



cèdent souvent la place à la moitié d'une des parties de l'ensemble, l'élève associant de manière inappropriée le tout à la connexité d'une des parties :



Cela montre l'importance qui est attribuée, parfois de manière démesurée, à la notion de connexité dans les processus de dénombrement, et cela indique certainement le caractère primitif de cette notion.

Selon le conventionnaliste, si nous faisions usage d'autres critères d'unité, notre expérience du déplacement d'objets pourrait certainement s'avérer très différente de celle que nous faisons avec nos critères habituels. C'est justement ce que révoque Sydney Shoemaker : de tels critères ne sont pas des conventions, mais des manières innées de voir le monde, et nous ne pouvons donc pas nous en défaire si facilement au profit d'autres manières de compter les choses. On trouve en effet chez Shoemaker un argument rattachant directement la notion d'unité, ou d'objet, à celle de cohésion motrice. Puisque nous faisons, écrit Shoemaker, l'expérience directe du déplacement d'un objet, c'est que nous sommes capables de rendre compte de cet objet en tant qu'unité relativement stable dans le temps :

> It is a striking fact that motion, though it involves the persistence through time of the moving object, is often directly observed rather than inferred [...] And I think it is partly because there is an experience of motion that spatiotemporal continuity occupies the central role it does as a criterion of identity [...] It does not seem to be just a matter of convention that we use spatiotemporal continuity as a criterion of identity. On the contrary, when I see motion (as opposed to inferring it) there seems to be no way in which I could describe what I see except by saying "It (or: something) is moving," and in saying this I imply the persistence of something through time. (S. SHOEMAKER, 1963, pp. 203-204)

De ce point de vue, la cohésion motrice, c'est-à-dire la synchronicité de déplacement des parties d'un objet indiquerait notre tendance spontanée à considérer cet objet comme doté d'une unité.

On peut trouver des antécédents à cette notion psychologique d'unité dans la *Théorie de la Forme* de Wolfgang Köhler, où l'auteur soutient que nos représentations sont naturellement et spontanément organisées en genres distinctifs d'unités. Les expériences de Köhler restreintes à des figures fixes semblent en effet indiquer que nous avons une manière instinctive de dénombrer certaines réunions de choses. Considérons les deux figures suivantes :



Pourquoi, demande Köhler, voit-on spontanément *deux* groupes de trois cercles plutôt que *six* cercles ou plutôt que *trois groupes* de deux cercles (KÖHLER, 1964, p. 147) ? De fait, la proximité spatiale des taches produit un effet de cohésion.

Pour Köhler, ce phénomène implique que nos critères d'unité ne peuvent pas relever seulement du langage. Nous aurions des tendances innées à regrouper des portions spatio-temporelles en objets, et à ne pas accepter que d'autres assemblages relèvent d'objets. La notion d'objet ne ressortirait donc pas seulement de pures conventions véhiculées au long de l'histoire par les sociétés humaines. Elle relèverait de fonctions cognitives. La notion d'objet ordinaire serait ainsi

expliquée par des dispositions innées à voir le monde d'une certaine manière, sans pour autant doter cette façon d'une prérogative ontologique. Nous compterions les choses de la façon dont nous le faisons ordinairement parce que notre entendement est structuré pour cette manière-là de compter les choses. Köhler dépeint sa thèse ainsi :

> Since early childhood we have often observed that sets of sensations which have approximately the same color, and differ in this respect from their environment, tend to behave as units, i.e., to move and be moved, to appear and disappear, at the same time. Such is the case with stones, with papers, with plates, with shoes, with many animals, with the leaves of plants [...]. It is only an example of the well known generalizing power of memory if, as a result of such experiences, we treat all homogeneously colored areas as units. (KÖHLER, 1964, p. 141)

Le constat d'unité provient, souligne Köhler, d'une habitude psychologique dû au caractère homogène que prend l'apparence d'un objet, autrement dit, la ressemblance qui se manifeste entre les parties contiguës, et l'aspect disruptif avec le fond de son environnement qui nous pousse à assimiler celui-ci à une unité. En outre, l'unité est étroitement associée à un comportement typique des objets : à savoir, la synchronicité de déplacement, l'apparition et la disparition des parties constitutives.

Par cette caractérisation basée sur la psychologie de la perception des objets en mouvement, Köhler réfute ainsi à la conception conventionnaliste forte :

> On the desk before me I find quite a number of circumscribed units or things: a piece of paper, a pencil, an eraser, a cigarette, and so forth. The existence of these visual things involves two factors. What is included in a thing becomes a unit, and this unit is segregated from its surroundings. In order to satisfy myself that this is more than a verbal affair, I may try to form other units in which parts of a visual thing and parts of its environment are put together. In some cases such an attempt will end in complete failure. (KÖHLER, 1964, pp. 137-138)

L'unité n'est pas un fait de langage, dit Köhler, mais une expérience, ou plus exactement, un « fait visuel » (KÖHLER, 1964, p. 138). Nous expérimentons visuellement qu'un stylo est un objet, mais que la somme méréologique stylo-table n'en est pas un, parce que nous pouvons déplacer le stylo indépendamment de la table.

L'unité d'un objet a donc une réalité qui n'est pas seulement linguistique, mais réelle, caractérisée par l'effet disruptif qu'il occasionne sur le fond de son environnement, par ses propriétés extrinsèques.

Les conventions de langage contribuent cependant à l'élucidation de la sorte d'objet auquel nous avons affaire. On pourrait certes imaginer un scénario dans lequel un stylo serait collé à une table, un peu à la manière dont les restes de repas dans les œuvres de Daniel Spoerri. Mais dans ce cas, nous devrions sans doute reconnaître que le stylo et la table y auraient perdu chacun leur autonomie pour participer à un tout que recouvre un nouvel objet, disons un « stylo-table », relevant d'une catégorie qui n'est plus celle du stylo ni celle de la table, mais par exemple celle d'une œuvre, c'est-à-dire, d'un nouvel objet, moins commun qu'un stylo et qu'une table. Dans ce cas, ce n'est pas l'unité des objets qui est remise en jeu, mais les concepts afférents à de tels objets.

Les paradoxes que nous avons relevés semblent ainsi indiquer que l'individualité (être un particulier d'une certaine sorte) n'est pas une propriété appartenant aux choses elles-mêmes mais une caractéristique *conférée* aux particuliers. Elle résulte d'une manière particulière de penser à un particulier.

1.5 Conclusion

Nous l'avons vu, l'enquête philosophique au sujet des objets tente de répondre à deux attentes différentes. La première attente, qui a une vocation métaphysique, consiste à déterminer ce qui constitue les unités ontologiques du monde. La seconde attente se rapporte à un enjeu épistémique consistant à déterminer ce qui fait l'unité d'une représentation, c'est-à-dire ce qui fait l'unité dans le mode de présentation d'un objet.

Nous avons vu que la réponse à la question de l'identité nécessitait de préciser la sorte d'objet que nous visons. Le problème épistémique de l'identification et le problème métaphysique de l'individuation temporelle d'une chose résident à la fois dans la question de la persistance d'une unité comme réunion de parties constituantes et à la fois dans la question de la nature de cette unité. Les partisans de l'identité métaphysique prescrivent de considérer que la nature de ces unités relèvent des sortes, qui sont des réalités en soi, là où nous proposons de les considérer seulement comme la sélection contextuelle de ressemblances, d'aspects similaires manifestés entre des choses.

Ce chapitre nous a ainsi entraîné dans une discussion mêlant identité métaphysique et unité de dénombrement. Deux principales questions ont émergé. La première est celle de l'agglomération non restreinte : la réunion quelconque de deux objets quelconques compose-t-elle encore un objet, comme l'affirment les partisans d'un certain libéralisme méréologique ? Par exemple, si deux jumeaux Pierre et Jean comptent chacun pour un objet, alors la somme méréologique Pierre-et-Jean constitue-t-elle encore un nouvel objet ? Accepter cette idée conduit à une sorte d'unité toujours répétée quand on réunit deux objets dont la somme fait encore un objet, ce qui semble contredire le sens commun qui y verrait une pluralité d'objets plutôt qu'un seul objet singulier. On peut aussi poser la même question au sujet des sommes de parties temporelles. Des parties temporelles quelconques composent-elles encore un objet ? La réunion des parties temporelles de cette table avant le 1 janvier 2020 et des parties temporelles de ce stylo après cette date forme-t-elle un objet ? La seconde question est relative aux sections arbitraires. Elle consiste à demander si n'importe quelle partie spatiale ou temporelle d'un objet constitue elle-même un objet. L'anse d'une tasse de café est-elle un objet ? La moitié gauche d'un livre est-elle un objet ? Et la main d'un individu [35] ?

Nous avons vu qu'une réponse possible était le conventionnaliste fort, qui affirme que ce que

[35] Du point de vu du droit français, la notion de personne est absolument irréductible à celle de chose. En conséquence, une personne n'est donc jamais propriétaire de son corps : elle *est* son corps. En revanche, une partie coupée et détachée d'une personne perd son statut juridique premier pour en gagner un nouveau en devenant une chose. Comme le fait remarquer l'historien du droit Jean-Pierre Baud, elle peut dès lors faire l'objet d'une revendication de propriété, et voir conséquemment différents prétendants en réclamer la pleine légitimité, et pas seulement la personne dont elle a été détachée du corps : « Puisque, dans se globalité, le corps de l'homme s'identifie à sa personne, et puisqu'un élément corporel détaché du corps est nécessairement une chose, la doctrine française établit ainsi qu'un élément corporel ne devient une chose qu'au moment où il est séparé du corps. Mais alors, si une main coupée ne devient une chose qu'au moment de son amputation, elle peut donc être appropriée par le premier qui s'en saisit. Il s'agit du premier occupant sur les choses qui n'appartiennent à personne. » (BAUD, 1993, p. 15).

nous considérons comme des sortes d'objets persistants, comme un stylo, comme une table, ou comme une œuvre plastique ne relève en réalité que de conventions. Mais rien de tel n'existe en dehors des masses de matière qui les composent, et qui peuvent toujours être rendues plus ou moins volumineuses. En revanche, comme nous l'avons vu, le conventionnalisme modéré n'exclut pas un cadre ontologique : ce n'est pas parce que nous pouvons compter les choses de différentes manières que nous devons les compter de n'importe quelle manière. Notre cognition est fortement imprégnée par les principes de séparabilité, de connexité et de synchronicité motrice. Nous avons une certaine liberté à découper de manière cohérente le monde par le langage, mais seulement en suivant certains plans de coupe.

On pourrait juger encore trop fort un tel conventionnalisme, et lui préférer une forme moins extrême, considérant que le découpage de la réalité vient partiellement de nous et partiellement du monde lui-même. Un telle approche existe effectivement dans la littérature philosophique contemporaine. Elle se manifeste par exemple à travers ce que David Armstrong qualifie d'universaux rares (*sparse universal*) (D. M. ARMSTRONG, 1978), que David Lewis appelle de son côté les propriétés naturelles (*natural properties*) (LEWIS, 1983c) et que Theodore Sider désigne sous le qualificatif d'aimants référentiels (*reference magnetism*) (SIDER, 2011). Si l'on admet une quelconque structure métaphysique du monde épistémiquement abordable, il reste alors la question de savoir dans quelle mesure notre connaissance s'amalgamerait autour de telles structures qualitatives. Imaginer que la réalité est dotée d'une structure métaphysique qui viendrait agréger nos conceptions autour d'elle, d'une manière toute analogue à celle dont la limaille de fer s'organise autour d'un aimant est certes une jolie image. Mais les conséquences tirées sont, en quelque sorte, réversibles. De fait, les aimants ont autant la faculté d'attirer que de repousser. On peut tout aussi bien imaginer que le monde est structuré d'une manière absolument repoussante pour la connaissance humaine qui chercherait à tout prix à en éviter la saisie par des artifices trompeurs, par une forme de duperie universelle par le double.

Probablement que le stylo et la table n'ont pas plus de réalité métaphysique que le stylo-table. Seuls existent les différents volumes de matière que nous désignons par ces termes. La restriction de constitution que nous imposons à ce que nous considérons comme de vénérables objets est intentionnelle, et non métaphysique, mais elle demeure bridée par des principes physiques. La continuité spatio-temporelle et la continuité matérielle fournissent de bons cribles d'identification d'un même volume de matière, et la connexité et la synchronicité de déplacement des parties constitutives en fournit un bon crible d'unité.

Nous concluons donc en faveur d'une conception conventionnaliste modérée de l'identité. L'unité d'un objet a donc une réalité qui n'est pas seulement linguistique, et son identification est tributaire d'un effet disruptif qu'occasionne l'objet sur le fond de son environnement, par ses propriétés extrinsèques, par la synchronicité motrice de ses parties constitutives, etc. Autrement dit, la notion d'unité n'est pas seulement acquise de manière culturelle, mais prédisposée par notre appareil cognitif en vertu d'une sélection de ressemblances entre les différentes phases d'un événement perçu comme par exemple le déplacement d'un objet.

Comme nous l'avons vu, des paradoxes émergent des conceptions métaphysiques de l'identité et de l'unité, qui induisent des doubles, des entités superposées quoique numériquement différentes. La théorie essentialiste de Wiggins prétend par exemple s'extraire des paradoxes de l'identité, mais elle le fait au prix d'objets numériquement distincts, quoique complètement coïncidents.

Dans le chapitre suivant, nous allons considérer plus généralement l'idée de doubles parfaits et examiner la question de savoir si de telles choses ont une quelconque place dans l'ontologie.

Chapitre 2

Métaphysique du double

> Et dans chaque anneau de l'humaine existence absolument parlant, il vient toujours une heure où, d'abord à un seul, ensuite à plusieurs, puis à tous, se révèle la plus puissante pensée, celle de l'éternel retour de toutes choses – et c'est à chaque fois pour l'humanité l'heure de *Midi*.
>
> Friedrich Nietzsche, *Fragments posthumes* (NIETZSCHE, 1982)

> L'idée du « retour éternel » est une des plus profondes, des moins discutées, et des plus fécondes.
>
> Abel Rey, *Retour éternel et philosophie de la physique* (REY, 1927, p. 267)

> Ma raison est assez bien convaincue, dit la marquise, mais mon imagination est accablée de la multitude infinie des habitants de toutes ces planètes, et embarrassée de la diversité qu'il faut établir entre eux ; car je vois bien que la nature, selon qu'elle est ennemie des répétitions, les aura tous fait différents.
>
> Fontenelle, *Entretiens sur la pluralité des mondes* (FONTENELLE, 1852, p. 92)

> You pretend you're not Laura Palmer ?
>
> *Twin Peaks* de David Lynch (LYNCH & FROST, 2017)

On peut critiquer une argumentation en refusant d'emblée l'argument, en montrant par exemple qu'une des prémisses est erronée ou qu'une étape de la démonstration ne respecte pas un certain principe logique. Une autre manière de critiquer un raisonnement consiste au contraire à l'admettre d'abord, et à le pousser jusqu'à ses derniers retranchements afin de faire saillir des incohérences.

C'est la méthode que je me propose d'employer dans ce qui suit. Dans le chapitre précédent, nous avons opté pour une conception conventionnaliste plutôt qu'objectuelle de l'identité. La conception objectuelle est, comme nous l'avons vu, fondée sur la recherche de critères d'identité tirés de propriétés. Dans ce qui suit, j'admettrai, avec le réaliste des universaux, l'idée que les prédicats, ou du moins que certains prédicats renvoient à des réalités objectives, afin de faire jaillir ensuite des paradoxes.

L'objet de ce second chapitre est de considérer le double dans ses rapports à l'identité objectuelle. Que seraient des doubles ontologiques si de telles choses existaient ? L'enjeu est de montrer qu'il y a, en philosophie, un problème de doubles, en décrivant les différents types.

Poser la question de l'existence de doubles relève d'une question ontologique : y a-t-il, au sein des choses qui sont, des choses qui sont des doubles ? Mais cette question paraît d'emblée articuler l'ontologie à la métaphysique, puisque qu'il s'agit d'une part de savoir s'il y a un certain type d'entités dans notre catalogue des choses qui existent (question ontologique), et d'autre part de déterminer si ces choses peuvent être d'une certaine manière, à savoir, si elles peuvent être indiscernables entre elles (question métaphysique).

Le plan de ce chapitre est le suivant : nous commencerons, dans la première section, par analyser de manière générale l'idée de mondes indiscernables dans les scénarios cosmogoniques puis dans les théories modales des mondes possibles. Nous aborderons ensuite, dans la deuxième section, la question de l'identité croisée et des contreparties et le problème du paradoxe de Chisholm. Suivra, dans la troisième section, une discussion du principe d'identité des indiscernables, en déployant trois catégories d'indiscernables ontologiques qui apparaissent dans les paradoxes d'identité, à savoir, les répliques pures, les doubles purs et les coïncidents purs.

2.1 Des mondes multiples ?

Le double a évidemment un rapport à la répétition : répétition dans le temps, répétition dans l'espace. Mais, d'un point de vue philosophique, la question du double est plus précisément encore celle de savoir *ce qui se répète*. Est-ce la *même* chose, ou est-ce autre chose ? Pour élucider cette question, nous approchons la question métaphysique du double en commençant par analyser l'idée de mondes multiples. L'objet de cette première section est de déterminer quels sont au juste les problèmes du double en philosophie, à travers différentes illustrations interrogeant d'abord la notion de monde. Cette notion est en effet un vecteur d'entrée propice à l'analyse de la question du double selon les interrogations suivantes :

1. Y a t-il plusieurs mondes ?
2. Y a t-il des mondes indiscernables ?
3. Le monde actuel contient-il des indiscernables ?
4. Le monde actuel aurait-il pu contenir des indiscernables ?

En examinant ces quatre questions, je développerai, au fil de ce chapitre, diverses instances de doubles : les doubles-mêmes, les contreparties, les mondes haeccéitiques, les répliques pures,

les doubles purs et les coïncidents purs. Ces questions, qui sont ici considérées selon un niveau métaphysique, seront déclinées d'abord en fonction de leur incidence physique, puis selon leur incidence logique.

2.1.1 L'espace physique et le double

Inclusivité et exclusivité

Combien de mondes y a-t-il ? La question étant posée ainsi, la réponse semble évidente : un seul, bien sûr ! celui que nous concédons être « notre monde » - et dire « notre », c'est déjà commettre un pléonasme. N'avons-nous pas déjà suffisamment affaire avec un monde ? N'est-ce pas bien assez que ce monde-là ?

La possibilité de parler du monde au pluriel semble *a priori* contestable. Le monde n'est-il pas la totalité de ce qui est, excluant la possibilité d'un autre monde qui ne coïnciderait pas pleinement avec celui-ci ? Car le monde est sans doute quelque chose d'extrêmement inclusif. Il est tout ce qu'il y a, comme le rappellent d'ailleurs les étymologies grecque (*cosmos*) et latine (*mundus*) du mot. Or le caractère absolument inclusif du monde implique son exclusivité. Puisqu'il englobe tout, le monde est unique, il est seul et sans rival. C'est déjà, rappelons-le, le constat qu'Aristote formule dans le *Traité du ciel*, où il rappelle qu'un usage de la notion de monde qui ne serait pas strictement rapporté à sa singularité caractéristique serait manifestement inapproprié. Aristote souligne en effet qu'« Il n'y a [...] aucune nécessité à conclure [...] qu'il existe plusieurs mondes, ni qu'il peut en exister plusieurs, s'il est vrai que ce monde-ci est composé de l'ensemble de la matière, comme cela est le cas. » (ARISTOTE, 2004, 278 a 10 27). Le monde, en tant que tel, est l'individu *maximal*. Dire qu'il existe des mondes *maximaux* numériquement différents est une contradiction. L'exclusivité du monde découle logiquement de son inclusivité exhaustive. L'unicité du monde est une conséquence de sa nature globale. Admettons en effet qu'il y ait non un seul, mais au moins deux mondes maximaux w et w'. Le monde w incluant tout ce qui existe, il devrait en particulier contenir tout ce qui existe aussi dans w'. Donc w' lui-même. Et le même raisonnement s'appliquerait aussi de manière symétrique pour montrer que w' devrait contenir w. Autrement dit, de manière extensionnelle, il vient que $w = w'$.

Pourtant, d'un point de vue linguistique, le concept de monde semble pouvoir être employé d'une manière fort similaire à celle des concepts sortaux comme « charançon » ou « personne ». En effet, la grammaire n'exclut pas que nous parlions de « différents mondes » ou de « ce monde », comme nous disons aussi qu'il y a *des* charançons sur l'hortensia ou comme nous parlons de *cette* personne plutôt que telle autre. Nous ne disons pas en revanche qu'il y a *du* monde - sinon dans un sens très différent, à savoir, pour indiquer la fréquentation d'un lieu, mais en tout cas, nous ne le disons pas comme nous disons qu'il y a *du* chocolat sur la nappe. À en croire la grammaire, la notion de monde semble ainsi relever de l'individuation. Ainsi, si le monde est exclusif, s'il n'y a qu'un seul et unique monde, cela implique aussi que le terme, employé pour désigner un individu (*ce* monde) indique un genre très particulier d'individu. Car il s'agit alors d'un individu exclusif en son espèce, sans rival possible, il est, comme l'écrit Stéphane Chauvier (CHAUVIER, 2010), un individu *unique en son genre*.

Si le monde est la totalité de ce qu'il y a, il est sans doute possible alors d'en rapprocher

le concept à celui de *réel*. Car comme le rappelle Clément Rosset, la signature du réel est son caractère singulier, proprement unique, non duplicable, « le réel étant précisément ce qui, pour être sans double, demeure réfractaire à toute entreprise d'identification » (ROSSET, 1979, p. 22). Or toute représentation du réel s'écarte nécessairement du réel en tant qu'elle n'est, justement, que représentation. En ce sens, elle constitue toujours une sorte de duplication. Partant, il n'y a donc aucun moyen de réfléchir le réel, parce que le geste-même de le réfléchir conduit à la production d'une duplique. Aussi, le réel demeure-t-il invisible :

> L'objet réel est en effet invisible, ou plus exactement inconnaissable et inappréciable, précisément dans la mesure où il est *singulier*, c'est-à-dire tel que qu'aucune représentation ne peut en suggérer de connaissance ou d'appréciation par le biais de la *réplique*. Le réel est ce qui est sans double, soit une singularité inappréciable et invisible parce que sans miroir à sa mesure. (ROSSET, 1979, p. 15)

Toute représentation est toujours une réplique du réel, et pour cette raison, elle s'écarte nécessairement de son modèle. Le double a donc un rapport à l'usurpation, du moins, quand on cherche à faire coïncider le réel avec sa représentation.

L'appréhension du réel relève ainsi chez Rosset d'une ontologie négative, à la manière dont on parle aussi de « théologies négatives ». Elle n'en diffère que dans la mesure où ce que les médiévaux avaient coutume d'attribuer à Dieu est ici attribué au Réel : « plus un objet est réel, plus il est inidentifiable » (ROSSET, 1979, p. 33). L'objet singulier est un objet inidentifiable en ce sens que toute identification est déjà la production d'*autre chose*, à savoir, une représentation. L'identité métaphysique, la singularité, est donc la prérogative du réel, mais elle demeure absolument inconnaissable [1]. Selon Rosset, le Réel doit ainsi être considéré comme « *un ensemble non clos d'objets non identifiables* ». Le contraire du réel, insiste Rosset, ce n'est donc pas le néant, l'absence d'être, mais bien le *double* (ROSSET, 1979, p. 28).

De tels arguments pourraient inciter à considérer une thèse plus forte encore, à savoir, que le monde n'est pas seulement exclusif en tant que monde, mais exclusif de la façon la plus générale qui soit. De ce point de vue, rien d'autre que le monde n'existerait. Il n'y aurait, au sens proprement ontologique, ni des hortensias, ni des charançons, ni des nappes, ni du chocolat, ni des personnes, mais seulement une unique étendue diversifiée par ses modes d'être. On trouve cette conception chez différents philosophes, quoiqu'ils ne s'accordent pas complètement sur les détails. Mais néanmoins, il semble assez juste d'affirmer que cette intuition joue un rôle majeur dans les philosophies de Parménide, de Plotin, de Spinoza, ou encore, pour citer un auteur contemporain, de John Heil (HEIL, 2011). Tous semblent en effet considérer que le monde est le seul objet réel car, d'une part, le caractère inclusif de monde assure qu'il n'y a rien au-delà ou au-dessus de lui, et d'autre part, son caractère absolument exclusif assure également qu'il n'y ait rien « en deçà » de lui. Selon Parménide, l'*Un* est un tout sans parties, complètement dépourvu de pluralité. Dans le *Poème*, Parménide dit par exemple que ce qui existe,

[1] C'est ce que Platon décrit aussi dans un célèbre passage de *La République*, lorsque l'esclave, sorti de force de la caverne, se trouve aveuglé par la lumière. Car, demande Platon, « est-ce que, une fois venu au jour, les yeux tout remplis de son éclat, il ne serait pas incapable de voir même un seul de ces objets qu'à présent nous disons véritables ? » (PLATON, 1950b, VII, 516a).

> [...] échappant à la génération, est en même temps exempt de destruction : car il est justement formé tout d'une pièce, exempt de tremblement et dépourvu de fin. Et jamais il ne fut, et jamais ne sera, puisque au présent il est, tout entier à la fois, un et un continu. (PARMÉNIDE, 1986, p. 261, Parménide, B, VIII)

Selon cette perspective, que je qualifierai d'« hénadique », il n'y a, au fond, qu'une seule chose, à savoir, la totalité de ce qui existe. De ce point de vue, toute chose demeure toujours la même parce que tout ce qu'il y a, c'est seulement l'ensemble de tout ce qui existe, ensemble qui reste identique à lui-même quelles que soient les modifications internes qu'il puisse subir. Ce qui est, est un tout sans parties propres [2]. Aussi, la forme du monde peut bien changer, cela reste le même monde. L'unité ne concerne au fond que la totalité de ce qui existe. Ce monisme radical, qui dénie au monde toute partie propre constitutive, confère à l'univers la prérogative de l'unité. La thèse de Parménide implique que les choses ne peuvent jamais être doubles puisqu'elles n'atteignent jamais vraiment l'Unité que dans le Tout, qui est sans rival. Il n'y aurait donc pas de multiples choses, mais une seule. Si n'est que l'*Un*, alors le deux ne survient jamais.

À titre d'illustration, on retrouve cette idée exprimée dans le contexte littéraire du roman de Jean-Paul Sartre intitulé *La Nausée* à travers la célèbre expérience existentielle éprouvée par Roquentin au jardin public :

> (...) la diversité des choses, leur individualité n'était qu'une apparence, un vernis. Ce vernis avait fondu, il restait des masses monstrueuses et molles, en désordre - nues d'une effrayante et obscène nudité. (...) Tous ces objets... comment dire ? Ils m'incommodaient ; j'aurais souhaité qu'ils existassent moins fort, d'une façon plus sèche, plus abstraite, avec plus de retenue. (...) De trop : c'était le seul rapport que je pusse établir entre ces arbres, ces grilles, ces cailloux. En vain cherchais-je à *compter* les marronniers, à les *situer*... chacun d'eux s'échappait des relations où je cherchais à l'enfermer, s'isolait, débordait. (SARTRE, 1981, p. 151)

Roquentin a la révélation d'une réalité exclusivement constituée d'une unique grande « masse » non-individuée. Derrière l'apparence de la séparation des choses, de leur aspect discret se tiendrait la réalité d'une « boue » métaphysique continue [3]. L'intuition éprouvée par Roquentin, c'est qu'en tout lieu, c'est *de* l'existence qui se manifeste - celle du monde, du réel, plutôt que *des* choses. Là où le langage tenait les choses encore écartées les unes des autres, dans l'expérience de Roquentin, elles se fondent les unes dans les autres en une totalité englobante, une masse informe, une *pâte* des choses.

Néanmoins, certains philosophes contestent la thèse d'exclusivité du monde, qu'elle soit forte ou non, soit qu'ils admettent plusieurs mondes, soit qu'ils en contestent carrément l'existence. Parmi ceux qui soutiennent l'inexistence du monde, Markus Gabriel objecte que la notion d'univers entendue comme limite ultime de ce qui existe n'a tout simplement aucune sorte de signification. Gabriel soutient en effet que le monde vu comme la somme de tout ce qui existe devrait nécessairement se contenir lui-même. L'argument de Gabriel est donc une version ontologique du paradoxe logique de Russell.

[2] Une partie b d'un tout a est dite *propre* si et seulement si b est différente de a.

[3] À cet égard, rappelons que l'histoire de Sartre est située dans la ville imaginaire de *Bouville*.

> Je soutiens que l'univers n'est pas le tout. (...) Le monde est bien plus grand que l'univers. (...) Si des États, des rêves, des possibles non actualisés, des œuvres d'arts et plus particulièrement aussi nos idées sur le monde font partie du monde, le monde ne peut pas être identique au domaine d'objets des sciences de la nature. (...) On ne peut donc raisonnablement définir le MONDE que si on le décrit en disant qu'il englobe tout, qu'il est le domaine de tous les domaines. (...) Et c'est précisément le monde, cette entité qui est censée comprendre tout, qui n'existe pas et ne saurait exister. (...) je soutiens que tout existe, excepté le monde (GABRIEL, 2013, pp. 18-19)

Mais cet argument est basé sur l'idée de distinguer l'univers, autrement dit l'ensemble des particuliers concrets, et le monde considéré comme une collection plus grande étendue aux objets inexistants. Gabriel nie la réalité du monde non pour défendre une thèse irréaliste, mais bien au contraire pour soutenir un point de vue réaliste : « le principe qui énonce que le monde n'existe pas implique que tout le reste existe. » (GABRIEL, 2013, p. 9). L'ontologie de Gabriel est en quelque sorte un pan-réalisme de type meinongien, où tout ce qui peut faire l'objet d'une intention est comptabilisé de la même manière qu'un authentique objet physique [4]. Mais rien de nous oblige à considérer les choses ainsi. Nous y reviendrons dans la partie de ce travail consacrée aux fictions.

À côté de l'affirmation de non-existence du monde, d'autres philosophes prétendent qu'il existe au contraire une pluralité de mondes. Comme nous l'avons dit, lorsque nous parlons du monde, nous faisons référence à quelque chose d'extrêmement inclusif : tout ce qui existe appartient au monde. Et de là, il semble nécessaire d'en déduire son unicité. Nous pouvons néanmoins envisager différentes manières de ne pas déduire la propriété d'exclusivité du monde de son caractère inclusif. L'une d'entre elle, métaphysique, est la théorie pan-réaliste fondée sur la thèse indexicale de l'actualité du monde, que défend David Lewis dans *De la pluralité des mondes possibles* (LEWIS, 2007), que nous aurons l'occasion d'examiner en détail plus loin, et qui affirme que les mondes possibles sont des mondes concrets très inclusifs, mais qu'ils sont, entre eux, causalement isolés. Cet isolement n'est cependant pas contradictoire avec les sommes méréologiques qui recouvrent différents mondes. On peut alors considérer la somme méréologique maximale recouvrant la totalité des mondes. Une autre manière, constructionniste, qui est celle de l'irréalisme que Nelson Goodman défend dans *Ways of Worldmaking* (GOODMAN, 2006), consiste (apparemment) à stipuler une pluralité de mondes *construits*, sans admettre un monde englobant. Nous aurons également à éclaircir cette conception dans ce qui suit. Mais pour le moment, je souhaite m'attarder sur une troisième manière de concevoir des mondes numériquement différents.

Si le monde est la somme extensionnelle de la totalité des choses qui existent, dès lors, dire que différents mondes appartiennent à une même pluralité, c'est dire qu'il existe quelque chose dont ils sont les parties. En tant que parties simples, ces mondes ne sont alors pas des choses maximales, ils ne sont des mondes que d'une manière lâche de parler. Cette troisième façon de considérer une pluralité de mondes est métaphysiquement faible, puisqu'elle consiste à restreindre le caractère inclusif du monde à une notion seulement locale et relative, comme l'attestent par exemple des

[4] Réalisme que Pascal Engel qualifie de *kitsch*, parmi d'autres métaphysiques qui se revendiquent appartenir au courant du « nouveau réalisme » (ENGEL, 2015). Dans la dernière partie de ce mémoire, j'accorderai quelques commentaires à la notion de *kitsch*, qui est une instance du double se rapportant à l'idée d'une authenticité dégradée. Si le terme provient originairement de contextes mercantiles et esthétiques, je montrerai qu'il est en réalité motivé par une démarche ontologique qui déborde largement sur la question de l'identité personnelle.

expressions comme « le monde végétal », comme « c'est un petit monde fermé », ou comme « le monde de Balzac ». Ces manières d'employer le terme de monde nous indiquent une pluralité de mondes qui se jouxtent, elles favorisent l'idée d'une coexistence de mondes. Ces mondes peuvent par exemple servir d'objets de comparaisons, comme l'entreprend David Baker dans sa célèbre étude sociologique sur les groupes « déviants » des musiciens de jazz, par exemple lorsqu'il écrit que « *The world of the dance musician* is a highly differentiated one [5]. » (BECKER, 1963, p. 84)

Une manière de comprendre de tels usages du mot « monde » serait de considérer les mondes à la fois comme des totalités closes et comme des réalités cohabitant l'une à côté de l'autre. En outre, de tels mondes pourraient aussi être englobés les uns dans les autres, à la manière de *matriochkas*, ce qui pourrait d'ailleurs être confirmé par exemple par le fait que nous disons des mondes sous-marins qu'ils forment une partie d'un environnement plus vaste. Constatons d'abord que de tels usages semblent référer à une acception certes localement close, mais qui n'est pas complètement globalisante. Un tel point de vue correspond à la notion de *milieu biologique*, ou *Umwelt*, qui, comme le caractérise Jakob von Uexküll (UEXKÜLL, 1965), désigne l'univers propre d'une espèce vivante, c'est-à-dire l'univers auquel cette espèce ou cet individu confère un sens en lui imposant ses propres déterminations. En particulier, le monde d'un individu vivant est partagé en monde d'action (*Wirkwelt*) et en monde de perception (*Merkwelt*). On parlera de cette façon de *l'univers de la tique* en ce que son mode d'appréhension de la réalité est particulièrement sensible une petit nombre d'aspects, en l'occurrence, aux émanations d'acide butyriques qui proviennent des mammifères, et que son monde d'action consiste à attendre, puis à se laisser tomber, et enfin à sucer le sang. Mais la tique demeure aveugle à bon nombre d'autres réalités. On peut étendre le concept de milieu aux individus humains et parler ainsi, comme Baker, du « monde du Jazz ». La complexité liée à l'intention engage évidemment des milieux qui ne varient pas uniformément par l'espèce, mais par groupes d'individus. Le milieu de l'astronome est ainsi celui « où l'être humain a troqué ses yeux contre un gigantesque instrument optique » (UEXKÜLL, 1965), et ce monde sera différent du monde du bibliophile. Le règne des animaux humains est donc celui d'un *milieu de milieux* alimentés par la diversité des conceptions scientifiques, artistiques ou religieuses, par les pratiques professionnelles et les loisirs, par les modes de vie. En ce sens, sociologique, un monde est un ensemble dans lequel s'inscrivent des processus mêlant des entités matérielles dont l'une ou plusieurs d'entre elles désignent ce dont elles font elles-mêmes parties. Mais cela n'implique au fond qu'une sorte de démultiplication épistémique de l'univers en fragments épars, et une telle manière de considérer le monde ne saurait en aucun cas nuire à son caractère ontologiquement exclusif.

Même s'il est indéniable que les êtres vivants habitent des milieux différents, nous pouvons toutefois souligner que tous ces milieux appartiennent à une globalité partagée. En parlant de différents mondes épistémiques comme des milieux, il ne s'agit au fond que de décrire des manières diverses d'habiter un monde ontologiquement unique.

La répétition à l'identique des mondes de la physis

Avant de considérer, dans ce qui suit, les versions métaphysiques réalistes et irréalistes et les versions sociologiques de la thèse d'une pluralité de mondes, je souhaite d'abord m'arrêter sur les

[5] C'est moi qui souligne.

versions *physiques* qui démarquent une objectivité minimale au monde. Pour cela, il nous faut revenir aux premières conceptions philosophiques. Car loin de défavoriser l'idée d'une pluralité de mondes, l'émergence de la philosophie grecque en a même exacerbé la thèse à travers l'idée d'une *répétition de choses matérielles qualitativement identiques*. C'est d'abord l'idée d'un retour cyclique du monde, répandue dans le monde gréco-oriental par la médiation de mages qui l'auraient eux-mêmes empruntée aux traditions cosmogoniques et plus largement religieuses iraniennes. Mais chez les Grecs, l'idée de cycles périodiques prend une forme rationalisée, appuyée sur des arguments théoriques, particulièrement attachée à la recherche de fondements ontologiques de l'être par le biais d'éléments primitifs. De fait, dans la littérature philosophique et philologique, il est souvent souligné [6], que la thèse d'une répétition cyclique de l'univers joue chez Héraclite et chez les stoïciens un rôle prépondérant dans les problématiques logiques (question de l'identité) et morales. Et bien sûr, on retrouve chez Platon une division de l'être en deux mondes parallèles dont l'un fait figure d'archétype. L'idée est d'ailleurs théorisée comme jamais elle ne l'a été à travers la thèse dualiste des formes intelligibles célestes dont participent les formes sensibles terrestres. Il est également souvent mentionné le rôle des mondes répétés dans la tradition stoïcienne. Mais il faut encore compléter ces références en remarquant que l'idée de mondes répétés dans l'espace joue aussi un rôle fondamental dans la physique et l'éthique épicuriennes, à cette différence près qu'elle y est purgée de toute notion d'*authenticité*. Je vais m'employer, dans ce qui suit, à montrer l'originalité de la thèse atomiste du point de vue de la répétition à l'identique des mondes. Comme on aura l'occasion de le constater, cette originalité jouera un rôle central dans le contenu de ce travail.

Je dois donc commencer par évoquer les théories cosmogoniques iraniennes qui ont irrigué la conception rationnelle grecque. Les études anthropologiques que Mircea Eliade (ELIADE, 1969) a consacrées aux ontologies antiques indiquent que beaucoup d'entre elles admettent l'hypothèse de mondes répliqués dans l'espace et/ou dans le temps. Eliade s'appuie notamment sur les travaux de l'iranologue Henrik Samuel Nyberg (NYBERG, 1931, pp. 31-36) consacrés à la métaphysique des traditions mazdéennes. Ces civilisations soutiennent deux types de duplication du monde. La première sorte est temporelle. En effet, selon ces cosmogonies, les événements du monde sont répétés dans le temps de manière cyclique, sur une certaine période. L'observation répétée du cycle des saisons, et plus généralement, des cycles astronomiques (année solaire, année lunaire) et biologiques, favorisent certainement l'émergence d'une telle conception. Notons en outre que cette idée est d'ailleurs probablement à l'origine des premiers calendriers chaldéens (COUDERC, 1946). En tant que cosmogonies, c'est-à-dire en tant que mythes de la création du monde, ces traditions soutiennent l'existence d'un cycle princeps, celui de la création originelle, dont les cycles suivants ne sont que des imitations du cycle de la création, répétés avec plus ou moins de similarité. L'idée d'une répétition cyclique imparfaite est également reprise par Pythagore et dans les mythes eschatologiques de Platon, à travers la thèse des métempsychoses, autrement dit, à travers l'hypothèse d'un retour des mêmes âmes dans des corps différents.

Si la reproduction n'est pas parfaite, c'est qu'une valeur peut être attribuée entre l'événement princeps et ses répétitions temporelles. Comme l'écrit encore Eliade, « *c'est la première manifestation d'une chose qui est significative et valable*, et non pas ses épiphanies successives » (ÉLIADE, 1963, p. 50). Eliade oppose ainsi deux temporalités, d'un côté, le temps *sacré*, cosmogonique, et

[6] Par exemple chez (BRUN, 1958), (LONG & SEDLEY, 1987), (GOURINAT, 2002).

d'autre part, le temps *profane* des répétitions. Cet écart entre l'original et la copie soumet alors l'univers à la possibilité d'une dégradation progressive pouvant conduire finalement à un effondrement [7]. L'éventualité de la fin du monde est déjouée par des célébrations religieuses, généralement pratiquées à la fin d'un cycle temporel, et dont le processus consiste à reproduire les conditions de la création originelle afin d'en raviver la répétition à l'intérieur d'un nouveau cycle. L'imitation humaine des gestes de la création permet la régénération cyclique du monde. Ces rituels conduisent à penser une sorte de renouvellement continue des cycles qui s'oppose à l'idée d'un déroulement linéaire du temps dans une progression historique : « toute Nouvelle Année est une reprise du temps à son commencement, c'est-à-dire, une répétition de la cosmogonie » (ELIADE, 1969, p. 69). Selon Eliade, dans les cosmogonies païennes, le temps de la création est un temps qui n'a pas la même valeur que le temps écoulé depuis, ou plus exactement, le seul temps qui compte est celui du retour à l'origine. La notion d'origine n'indique ici pas seulement l'antécédence du cycle princeps, mais également le degré plus élevé de réalité que celui-ci manifeste en rapport aux cycles suivants. L'origine temporelle est donc associée de manière axiologique à une notion de prééminence archétypale, une *authenticité* ontologique.

L'authenticité peut être décrite par la deuxième sorte de duplication qu'on trouve dans les diverses cosmogonies païennes, qui est d'ordre spatial. Dans le système métaphysique iranien de tradition zervanite, toute chose terrestre possède un pendant, *céleste*, d'une nature similaire (NYBERG, 1931, p. 33). Chaque élément matériel qui appartient à un cycle spatio-temporel terrestre, ou « *gêtîk* », est intimement associé à une contrepartie idéalisée, « *mênôk* céleste ». Il y a donc une forme de plan de symétrie, dans l'espace, reliant le monde terrestre au monde céleste par le centre du monde, généralement matérialisé par un tumulus naturel (montagne) ou artificiel (*ziggourat*), sorte de cordon ombilical reliant les deux mondes et marquant un endroit sacré.

On trouve là encore un obstacle qui semble empêcher la stricte répétition à l'identique des deux mondes : entre un monde et son symétrique, il y a encore une différence, qui n'est pas, cette fois encore, seulement d'originalité - c'est-à-dire manifestant une primeur temporelle dans la succession des répétitions, mais qui est une différence d'*authenticité*. Autrement dit, au sens que je donnerai ici à cette notion, il s'agit d'une différence *qualitative* de degré de réalité, le monde céleste manifestant une suprématie dans l'ordre du réel [8].

On voit que ces religions insistent sur l'idée de *répétition* du monde. Même répétition d'événements dans le temps, même répétition de particuliers dans l'espace. Il faut néanmoins nuancer cette répétition étant donné que ce qui est répété l'est avec une intensité décroissante. La décadence progressive des répétitions limite leur similarité. Sans l'intervention des hommes, par la

[7] Notons au passage que les religions monothéistes ont renversé la progression païenne en proposant une cosmologie passant de la chute hors d'Eden à la « Révélation (*Apocalyspe*) ». Rappelons aussi que d'autres traditions, par exemple bouddhistes, qui affirment elles-mêmes le retour cyclique des événements, privilégient au contraire la fuite *hors du monde*. Elles incitent ainsi à s'émanciper de la répétition des réincarnations successives (*métempsychose*) par l'accès à l'extinction (*Nirvana*).

[8] Bien sûr, les concepts d'authenticité et d'originalité sont fortement liés, puisque dans la plupart des contextes, il est attribué à une réplique une valeur moins élevée que la chose qu'elle copie. Par exemple, et pour changer de contexte, la valeur du premier tirage d'une lithographie est plus élevée que celle des tirages suivants. Au sens où j'emploie ces mots, l'originalité et l'authenticité sont donc généralement des notions très fortement intriquées, mais qu'il paraît cependant utile de distinguer. On notera également l'importance de l'authenticité dans le rapport aux propriétés esthétiques. Walter Benjamin parle par exemple de l'« aura » d'une œuvre d'art pour qualifier son aspect non reproductible. Nous aurons l'occasion de revenir sur ces questions dans le chapitre 6 consacré aux objets esthétiques.

pratique de rituels, cette séquence finirait par s'interrompre. La répétition est donc, *in fine*, conditionnée par les intentions humaines. Dans ces traditions, les mondes répétés ne sont que l'écho lointain du monde de la création, qui subsiste aussi comme modèle idéal, archétype existant dans un monde céleste, symétrique au monde sensible [9]. Ce monde est à la fois considéré comme nécessaire, parce que créé, et archétypal, parce que original et authentique. Ces traditions favorisent ainsi l'idée qu'il n'y a finalement qu'un seul monde réel, mais que ce monde n'est pas le monde sensible [10].

L'évolution axiologique des répétitions d'un monde original et authentique, qu'elle soit à tendance décadentiste ou progressiste, constitue le propre de la vision religieuse du monde. D'un tel point de vue, l'espace et le temps ne sont pas homogènes : seul le temps de la création est réel, et seul l'espace du centre du monde est réel. Le temps présente ainsi des « ruptures et des cassures » (ÉLIADE, 1957, p. 25). Quant à l'espace, il s'agit d'une forme amorphe qui ne prend de réalité qu'à partir d'une singularité, un centre du monde :

> L'expérience religieuse de la non-homogénéité de l'espace constitue une expérience primordiale, homologable à « une fondation du Monde ». [...] Lorsque le sacré se manifeste par une hiérophanie quelconque, il n'y a pas seulement rupture dans l'homogénéité de l'espace, mais aussi révélation d'une réalité absolue, qui s'oppose à la non-réalité de l'immense étendue environnante. La manifestation du sacré fonde ontologiquement le monde. Dans l'étendue homogène et infinie, où aucun point de repère n'est possible, dans laquelle aucune *orientation* ne peut s'effectuer, la hiérophanie révèle un « point fixe » absolu, un « Centre ». (ÉLIADE, 1957, p. 26)

Nous avons vu que les cosmogonies religieuses orientales associent une valeur de réalité aux mondes par un degré d'authenticité (selon une répartition spatiale) et d'originalité (selon une répartition temporelle) de ce qui est répété. Seul le temps originel, et seul l'espace qu'il consacre ont une valeur positive de réalité. Le reste relève tout simplement de l'irréel. Ainsi, comme nous l'avons déjà remarqué, dans de telles traditions, ce qui est répété n'est pas strictement similaire à ce qu'il répète. Or, nous allons montrer que la vision atomiste s'oppose directement à cette idée, en ce qu'elle ne confère aucune valeur particulière à un monde plutôt qu'à un autre.

Notons qu'il y a bien, dans la mythologie grecque, une pluralité de mondes qui communiquent entre eux. L'univers y est partitionné en trois strates, l'Olympe des Dieux, la Terre des humains et l'Hadès des âmes errantes. Cette tripartition spatiale du monde est complétée par une quadripartition temporelle, les quatre âges d'or, d'argent, d'airain et de fer décrits par Hésiode (HESIODE, 1928). Mais il est utile de remarquer que ces mondes ne cèdent encore de place qu'aux seuls humains. En effet, toute chose existante y revêt une forme humaine, voire, super-humaine. Cet état de fait n'est pas souvent souligné dans la littérature, à l'exception notable de Nietzsche qui, dans *La philosophie tragique à l'époque des Grecs*, constate que, dans la perspective mythologique, la nature n'est en effet que « déguisement, mascarade et métamorphoses des hommes-dieux » (NIETZSCHE, 1975, p. 222). La généalogie de la morale doit être mise en parallèle à une généalogie des mondes, qui sont toujours des ensembles de valeurs, donc des créations humaines. La

[9] La symétrie étant entendue ici au sens large comme une certaine « harmonie », une correspondance entre des éléments corrélés.

[10] C'est bien sûr le point de vue que défendra, avec une grande véhémence, Platon.

mythologie ne serait au fond qu'une sorte de vaste constructionnisme plus ou moins conscient.

L'idée d'objectivité doit être associée à celle de mondes amoraux. Elle ressortit directement à un rejet de la double axiologie de l'authenticité et de l'originalité, rejet qui fera l'enjeu de la philosophie atomiste. La philosophie naissante, amorcée par Thalès, puis développée par les physiciens Anaximène, Héraclite, Empédocle et Leucippe [11], relève d'un acte de libération de la nature, affranchie du joug humain. Il s'agit de concéder à la nature (*physis*) une indépendance ontologique par le fait de la désolidariser, en partie, sinon complètement, de la sphère des volontés humaines, et donc de la sphère des valeurs. Comme on le sait, Thalès introduit à cet effet un constituant primitif, l'eau, destiné à rendre compte de la réalité fondamentale des choses. Or, comme le souligne encore Nietzsche, cette objectivité minimale introduite dans la conception du monde relève d'une sorte de grandiose *métamorphose* épistémique :

> Quand Thalès dit que « tout est eau », l'homme tressaille et quitte la situation de chenille tâtonnante et rampante propre aux sciences particulières ; il pressent la notion ultime des choses et dépasse par ce pressentiment la gaucherie vulgaire des niveaux inférieurs de la connaissance. (NIETZSCHE, 1975, p. 347)

L'introduction de l'idée de nature contribue à fonder une ontologie axiologiquement neutre des mondes, et permet ainsi de récuser la prééminence de l'originalité. Il faut maintenant considérer comment elle est également corrélée à la notion de répétition par ce que Nietzsche a qualifié de théorie de l'« éternel retour » (*Ewige Wiederkunft*) [12]. Cette théorie est d'abord due, philosophiquement parlant, à Héraclite d'Éphèse. Celui-ci considère une période, « La Grande Année », dont l'étendue est évaluée à 10 800 années, au bout de laquelle toutes les choses sont amenées à se répéter à l'identique. À côté de la théorie physique qui confère au feu le rôle de constituant primitif des choses, Héraclite introduit aussi l'idée d'une conflagration universelle. La thèse est que le monde, en bout de course, est amené à prendre feu de toute part. Il s'agit cependant moins d'une destruction universelle que d'un changement de forme. La conflagration universelle débouche en effet sur une renaissance, ou *apocatastase*, qui est une régénération de l'univers par son élément primitif. L'idée générale, vue à travers le prisme des conceptions des premiers philosophes-physiciens, est sans doute que tout ce qui vit est amené à s'user jusqu'à disparaître. Pour que l'univers puisse se maintenir indéfiniment dans le temps, il est donc nécessaire qu'il soit régénéré à partir de son élément primitif qui doit, pour cela, reprendre sa forme pure. Toute chose doit ainsi revenir à son état de nature afin de parvenir à se maintenir dans l'existence, et cette succession se produit une infinité de fois.

Notons que chez ces premiers physiciens, les éléments « naturels » primitifs sont choisis parmi des substance de masses, comme l'eau, le feu, la terre et l'air. Ils ne sont pas encore vus comme des substances *individuées*. C'est-à-dire que l'idée de choses indiscernables mais numériquement identiques n'est pas encore rendue possible comme elle le sera ensuite avec la notion d'atomes. Avant d'y venir, je souhaite d'abord discuter de la notion de mondes répétés dans la tradition stoïcienne, parce que celle-ci incarne le rejet des indiscernables à travers la question de l'identité

[11] J'emploie évidemment ici le terme « physicien » au sens antique du terme.

[12] Je prendrai dans ce qui suit l'expression d'éternel retour dans son acception la plus générale, qualifiant aussi bien le retour d'un individu dans le temps que sa répétition dans l'espace.

personnelle. Commençons par souligner que la doctrine de la palingénésie stoïcienne emprunte à Héraclite la théorie de la conflagration universelle. Comme en témoigne Stobée,

> Zénon, Cléanthe et Chrysippe pensent que la substance se transforme en feu, comme si elle redevenait semence, et que, à partir de là, il s'accomplit de nouveau la mise en ordre du monde telle qu'elle était précédemment. (STOBÉE, 1884, I, 20, p. 171)

La problématique stoïcienne est cependant développée à partir d'une question existentielle sur le moyen de mener une vie heureuse. Et de là, elle mène à la question de l'identité personnelle à travers les cycles d'existence. En effet, comme le rappelle Jean Brun (BRUN, 1958), dans son *De natura hominis*, Némésius prétendait que pour les stoïciens, « Il y aura à nouveau un Socrate, un Platon... cette restauration ne se produira pas qu'une fois mais plusieurs fois : ou plus précisément, c'est éternellement que les choses seront restaurées ». Constatons alors qu'avec la thèse de la répétition de l'univers, les stoïciens sont confrontés à d'évidents paradoxes. Car est-ce Socrate lui-même qui revient, ou un autre individu qui aurait toutes les qualités de Socrate sans lui être identique ? Pour qu'il y ait répétition, il faut manifestement qu'intervienne une différence numérique. Si x à t_1 est identique à y à t_2 pour $t_2 > t_1$, il appert que y ne saurait « répéter » x. Nous ne dirions pas cela, nous dirions simplement, « y, c'est x ». Nous parlons de « répétitions » pour désigner le retour de qualités similaires, pour parler de ressemblances, autrement dit, pour mentionner une identité qualitative et non pour parler d'égalités numériques liées à la persistance d'une chose dans le temps.

En quel sens peut-on alors encore considérer que les individus qui « reviennent » sont les mêmes ? Comment considérer que ce qui revient, c'est Socrate lui-même ? Comme le rappelle Origène (GOURINAT, 2002, p. 222), certains stoïciens soutiennent que ce n'est pas possible, que celui qui revient n'est pas Socrate, mais une duplique de Socrate, un double. Mais dans ce cas, dans quelle mesure ce double pourrait-il s'avérer être discernable de Socrate ? Alexandre d'Aphrodise considère l'idée d'individus répétés qui soient à la fois numériquement identiques et possiblement discernables :

> En ce qui concerne les êtres doués d'une qualité propre qui existent postérieurement, il y a des différences par rapport aux êtres antérieurs seulement en ce qui concerne certains accidents extérieurs, comparables aux changements qui affectent Dion lorsqu'il reste le même au cours de sa vie, et qui ne le rendent pas différent. En effet, il ne devient pas différent si par exemple il avait des verrues sur le visage et qu'ensuite il ne les a plus. Ils disent qu'il y a des différences comparables chez les êtres doués d'une qualité propre d'un monde par rapport à ceux d'un autre. (GOURINAT, 2002, p. 220)

Anthony Long et David Sedley (LONG & SEDLEY, 1987) relèvent quatre possibilités de relations entre Socrate et l'individu qui « revient ». Ces individus peuvent être :

- des individus numériquement identiques et (donc) absolument indiscernables ;
- des individus numériquement distincts (*different token*) et absolument indiscernables (*same type*) ;

- des individus numériquement identiques mais possiblement discernables (*inessential discernibility*) ;

- des individus numériquement distincts mais (seulement) légèrement discernables (*slightly discernible tokens from the same type*).

Si l'on considère, en outre, que les stoïciens admettent le principe d'identité des indiscernables, qui rappelons-le, affirme que des choses numériquement distinctes ne peut être parfaitement ressemblantes, on devrait d'emblée rejeter la seconde alternative sous menace de contradiction logique. Ainsi, pour certains commentateurs, ce sont bien les mêmes individus qui reviennent (hypothèse 1 ou 3). Alexandre d'Aphrodise écrit que :

> [les stoïciens] pensent que, après la conflagration, toutes les mêmes choses dans le monde se produisent de nouveau numériquement, de sorte que l'être doué d'une qualité propre est de nouveau le même que celui d'avant et qu'il existe dans ce monde, comme Chrysippe le dit dans son traité *Sur le monde*. (GOURINAT, 2002, p. 220)

Némésius souligne en outre que chez les stoïciens, la Grande Année coïncide avec la révolution cyclique des astres. Autrement dit, ce ne sont pas seulement les individus qui reviennent, mais la succession temporelles des *événements*. Tout se reproduit de la même manière :

> Les Stoïciens disent que, une fois les planètes revenues au même signe céleste en longueur et en largeur, où chacune était au commencement, dès que le monde avait été formé, il se produit à une période déclarée de temps une conflagration et une destruction des êtres, et que, de nouveau depuis le début, le monde revient dans le même état, puis, comme les astres qui suivent de nouveau le même cours, chacun des événements de la période précédente s'accomplit sans différence. (GOURINAT, 2002, p. 221)

Les individus sont régénérés, ils reviennent en dépit de leur non-persistance temporelle. Ici, c'est à nouveau l'idée religieuse d'une réincarnation d'une substance immatérielle qui ré-émerge. Comme l'écrit encore Origène :

> Socrate sera de nouveau fils de Sophronisque et Athénien, et Phénarété épousera de nouveau Sophronisque et l'enfantera de nouveau. [...] Socrate ressuscitera, issu du sperme de Sophronisque ; il sera formé dans l'utérus de Phénarété, il sera éduqué à Athènes et deviendra philosophe et sa philosophie antérieure ressuscitera et pareillement ne sera pas différente de sa philosophie antérieure. Et Anytos et Mélétos ressusciteront aussi, de nouveau accusateurs de Socrate, et le conseil de l'Aréopage condamnera Socrate. [...] Socrate portera des vêtements qui ne seront pas différents de ceux de la période précédente, dans une pauvreté qui ne différera pas, dans une cité d'Athènes qui ne différera pas de celle de la période précédente. (GOURINAT, 2002, p. 223)

L'idée est donc celle d'une métempsychose qui serait une réincarnation « à l'identique » non seulement des individus, mais aussi de tous les événements.

En insistant sur le rôle que joue l'admission du principe d'identité des indiscernables, si l'on accepte le retour cyclique des choses, il ne s'agit plus d'une répétition d'indiscernables numériquement différents, mais du retour d'un seul et unique individu. Des cycles temporels ne sauraient être parfaitement ressemblants sans être alors numériquement identiques. Il n'y a donc, au sens propre, qu'un seul cercle temporel. De fait, il y a une sorte d'incohérence entre le principe de répétition et le principe de l'éternel retour. Comme on le voit, la différence entre répétition et retour est occasionnée par l'admission ou non d'indiscernables. D'une certaine manière, ce qui revient en étant le même ne se « répète » qu'une seule fois [13]. Si les cycles reviennent, en revanche, ils ne se répètent pas, et s'ils se répètent, alors ils ne reviennent pas.

On voit que ces questions suscitent non seulement le problème de l'identité temporelle, mais également celui de l'identité transmondaine à travers la question des possibles [14].

À côté des discussions engendrées par les paradoxes stoïciens, la thèse des répétitions apparaît également chez les atomistes antiques. Avec l'atomisme, l'espace n'est plus référencé à un point fixe sacré, mais relève d'une étendue homogène et infinie dans laquelle se meuvent des entités simples, les seules qui soient responsables de la création du monde, dont certaines sont indiscernables [15]. L'atomiste met ainsi en exergue l'idée d'une répétition qui serait strictement *identique*, autrement dit, sans origine nécessaire et sans rapport à l'authenticité. Une répétition *parfaite*. Dans la théorie physique fondée par Leucippe et Démocrite, développée par Épicure et mise en poème par Lucrèce dans le *De natura rerum*, seuls existent originellement les atomes et le vide. Ils constituent le Tout, l'Univers. Les atomes sont des indivisibles, des simples, c'est-à-dire des particuliers non composés, des entités sans parties. Ils existent en nombre illimité et sont immuables. Toutes les choses qui ne sont pas des atomes, comme les charançons, les hortensias, les nappes, le chocolat, les personnes, les sons, les accès de mauvaise humeur et les courants d'air sont des agrégats éphémères, des composés d'atomes, des choses ayant des parties, des divisibles. Quant à l'univers, Épicure prétend qu'il est infini de deux manières, d'une part par le nombre de corps qu'il contient, et d'autre part par l'immensité du vide qu'il renferme [16]. La démonstration qu'il donne dans la *Lettre à Hérodote* (LAËRCE, 1999, Livre X, 41) repose sur l'idée que l'absence de limite

[13] C'est sans doute cette idée qui est exploitée par Gilles Deleuze dans *Différence et Répétition* (DELEUZE, 1968), où l'auteur distingue deux ordres, celui de la répétition et celui de la généralité, le premier relevant du qualitatif, des ressemblances, et le second, du quantitatif, des équivalences. La généralité concerne la substitution des éléments particuliers, la répétition bute en revanche devant une singularité sans remplaçant possible, car la répétition concerne toujours un événement unique. Deleuze emploi ici la conception du temps que Charles Péguy attribue aux événements, induisant une certaine intuition renversée de l'originalité d'un événement : « La prise de la Bastille, dit l'histoire, ce fut proprement une fête, ce fut la première célébration, la première commémoration et pour ainsi dire déjà le premier anniversaire de la prise de la Bastille. Ou enfin le zéroième anniversaire. On s'est trompé, dit l'histoire. On a vu dans un sens, il fallait voir dans l'autre. On a vu. Ce n'est pas la Fête de la Fédération qui fut la première commémoration, le premier anniversaire de la prise de la Bastille. C'est la prise de la Bastille qui fut la première Fête de la Fédération, une Fédération avant la lettre. » (PÉGUY, 1916, p. 142). Deleuze résume ceci en inversant la préséance temporelle : c'est la prise de la Bastille du 14 juillet 1789 qui répète, déjà, la fête de la Fédération du 14 juillet 1790.

[14] C'est peut-être ici la raison de l'attention particulière que les stoïciens réservent au traitement logique des expressions modales et des paradoxes qui en découlent comme celui de Dion et Théon élaboré par Chrysippe.

[15] Il est donc à remarquer que l'atomisme présuppose l'existence d'indiscernables ontologiques.

[16] Rappelons que l'introduction de l'illimité (*apeiron*) en philosophie est généralement attribuée à Anaximandre. Par ailleurs, Simplicius affirme qu'Anaximandre avait lui-même tiré la conséquence d'une pluralité de mondes à partir du principe d'illimité : « [Anaximandre] posait, semble-t-il, qu'il existe des mondes illimités et que chacun de ces mondes naît de cet élément illimité » (COLLECTIF, 2001, p. 32).

d'une chose implique que cette chose est infinie. Or une limite, c'est par définition ce qui sépare une chose de quelque chose de différent. Mais au-delà d'une limite au vide, il y aurait *encore* le vide, ce qui est contradictoire. Donc le vide est sans limite, ce qui, pour Épicure, revient à dire qu'il est infini. En conséquence, les atomes sont eux-mêmes en nombre infini, parce que s'ils ne l'étaient pas, ils se disperseraient dans l'infini du vide, et nul agrégat ne pourrait jamais se former. En termes plus contemporains, nous pourrions dire que pour un nombre d'atomes fini évoluant au sein d'un univers infini, la probabilité de rencontre serait quasi-nulle, et il n'y aurait donc pas de monde constitué [17].

Toujours dans la *Lettre à Hérodote*, Épicure tire de l'hypothèse des deux infinis que sont espace et matière, la conséquence de la pluralité des mondes (LAËRCE, 1999, Livre X, 45). Puisque les atomes sont en nombre infini, dit Épicure, c'est que le nombre de mondes qu'ils constituent est lui-même sans limite. Notre monde n'a donc aucune exclusivité ontologique. L'abondance de matière et la portée infinie de l'espace font que les agrégats se forment en nombre infini. Or Épicure déduit de la diversité des mondes un nombre infini de mondes semblables au nôtre (LAËRCE, 1999, Livre X, 45). Car s'il existe un nombre infini de mondes produits par une quantité inépuisable d'atomes, cela implique nécessairement un nombre infini de mondes *reproduits*.

On retrouve la même thèse chez Lucrèce qui écrit que dans la mesure où la théorie atomiste est valable dans toutes les régions de l'espace, il faut alors admettre l'existence d'autres mondes similaires au nôtre, quoique diversifiés par les espèces : « si la même force, la même nature subsistent pour pouvoir rassembler en tous lieux ces éléments dans le même ordre qu'ils ont été rassemblés sur notre monde, il [...] faut avouer qu'il y a dans d'autres régions de l'espaces d'autres terres que la nôtre, et des races d'hommes différentes, et d'autres espèces sauvages » (LUCRÈCE, 1978, p. 81) Mais il faut aussi en déduire que les individus qui appartient à la même espèce existent eux-mêmes en nombre illimité.

> [...] dans la somme des choses, il n'y en a pas qui soit isolée, qui naisse unique, qui grandisse unique et seule en son genre : mais chacune appartient à quelque famille, et très nombreuses sont celles de la même espèce. Et tout d'abord considère les êtres vivants : tu trouveras que c'est ainsi qu'ont été créés et la race des fauves errants sur les montagnes, et la descendance des hommes, et enfin les troupes muettes des poissons écailleux et toutes les espèces qui volent. Aussi, d'après le même principe, faut-il reconnaître que le ciel, et la terre, et le soleil, la lune, la mer, et tout ce qui est, loin d'être unique, existe, au contraire, en nombre innombrable ; puisque leur existence a son terme inébranlablement fixé, et qu'ils sont d'essence mortelle, tout autant que toutes les sortes de corps que chaque espèce terrestre présente en abondance. (LUCRÈCE, 1978, p. 81)

On voit ainsi que le principe d'identité des indiscernables qui prévaut dans la conception des cycles spatio-temporels des philosophes stoïciens n'est pas du tout retenu par les épicuriens, qui admettent bien des répétitions à l'identique [18].

[17] La démonstration de ces deux infinis est reprise de manière détaillée par Lucrèce dans le *De natura rerum* (LUCRÈCE, 1978, p. 36).

[18] Rappelons aussi la remarque mordante de Karl Marx, dont la thèse fut consacrée à Épicure, soulignant le retour des événements historiques à une différence de mode près : « Hegel fait quelque part cette remarque que tous les grands événements et personnages historiques se répètent pour ainsi dire deux fois. Il a oublié d'ajouter : la première fois comme tragédie, la seconde fois comme farce. (MARX, 2007).

✤

Comme l'indique Clément Rosset dans *L'anti-nature* il faut toutefois remarquer le caractère extrêmement ambigu du concept de nature (*physis*). Il est souvent considéré qu'il s'oppose à l'idée religieuse de *sur-nature*. Prenons toutefois garde à ce que, non seulement il « ne fasse que substituer une religion à une autre, mais encore en revienne toujours, sous couvert de lutter contre la superstition, à une source à la fois très générale et très vive de toutes les représentations religieuses, accomplissant ainsi deux grands pas en arrière pour un petit pas en avant » (ROSSET, 1973, p. 32). Or, pour ne pas marcher en écrevisse, nos principes philosophiques devraient retenir que l'idée de surnature repose justement sur l'idée de nature en tant que *raison* des choses, idée qu'elle présuppose. Ce qui fait dire à Rosset, après le jeune Marx de la *Différence entre la philosophie de la nature de Démocrite et celle d'Épicure*, que l'idée de nature ne s'oppose pas à l'idée religieuse, mais au matérialisme :

> Ce qui est contraire à l'idée de nature n'est pas l'idéologie religieuse mais, à l'opposé, la pensée matérialiste, qui refuse de voir dans l'existence tant l'effet de forces que le résultat de principes : pour rendre compte de ce que les hommes appellent la nature, le matérialisme se contente d'invoquer deux « refus de principe » qui sont l'*inertie* (refus d'introduire l'idée de force dans l'existence) et le *hasard* (seul apte à rendre compte de la possibilité des productions sans entorses au principe d'inertie). (ROSSET, 1973, p. 33)

Or toute l'originalité du point de vue atomiste dans sa version mature, c'est d'affirmer qu'il n'y a pas de place pour un monde authentique ou nécessaire et encore moins pour l'idée d'un monde idéal. Les mondes sont le fruit du hasard, et l'un d'entre eux ne peut avoir plus de valeur que les autres. Autrement dit, à côté de l'*innocence du devenir* (*Unschuld des Werdens)* que célèbre Nietzsche à travers le crible de l'éternel retour, il faut également reconnaître l'*innocence des origines* que célèbrent les atomistes abdéritains. Considérant cette libération de la pure contingence qui sourd des théories physiques par l'effet du synchrétisme matérialiste d'Empédocle, Clément Rosset constate que :

> la nature a été arrachée à un réseau d'interprétations magiques et mystiques, mais elle n'a pas encore été réintégrée dans un système philosophique qui lui restituera son ordre et sa finalité : elle existe à l'état libre (la liberté du hasard), constitue un domaine ouvert à toutes les possibilités d'artifice, lesquelles se savent, en l'absence d'organisation naturelle, permises et innocentes. (ROSSET, 1973, p. 129)

En effet, chez Empédocle, l'idée d'une *nature*, entendue comme création authentique, originale et surtout, nécessaire, est complètement disqualifiée, favorisant l'idée d'une pensée que Rosset appelle artificialiste. Ainsi, dans son poème *De la nature*, Empédocle soutient qu' « il n'y a de nature pour aucune des choses mortelles, ni aucune fin par la mort funeste, mais seulement mélange et dissociation des mélanges, que les hommes ont appelés nature » (COLLECTIF, 2001, Fgt 8, trad. cor., p. 376).

Avec la quadruple racine matérielle de l'existence de toute chose comme composé d'eau, de feu, d'air et de terre, Empédocle fait émerger en philosophie l'idée que les choses sont *par mélanges*, c'est-à-dire par hasard, et non *par essence*, c'est-à-dire par nécessité :

> Refus donc, chez Empédocle, de célébrer les essences ; refus de considérer que des essences soient à la source des productions naturelles. Ce qui existe trouve son chemin sans le demander à la forme, ni à l'essence, ni à la nature ; c'est seulement après coup que l'éphémère fixité de leur forme donne aux espèces l'apparence d'un destin. (ROSSET, 1973, p. 139)

C'est ici que s'attarderont les critiques d'Aristote qui n'a pas vu que le but d'Empédocle était précisément de renoncer à l'idée de nature (en tant que liée à l'idée de nécessité) plutôt qu'à l'instituer.

Mais si ce n'est par nécessité, comment naît alors un monde (*kosmoï*) ? D'après l'atomisme d'Épicure, les particules parcourent d'abord l'univers selon des trajectoires quasi-rectilignes à ceci près qu'elles vibrent d'un infime frémissement. Cette légère déclinaison, appelée *clinamen*, cause la rencontre des atomes entre eux. Ils s'agrègent en raison de corrélations géométriques, leurs jeux de relations constituant alors un monde. La déclinaison est donc non seulement la cause qui fait que le monde est ce qu'il est, mais elle conditionne également les diverses configurations possibles que le monde *aurait pu prendre*. La rencontre aléatoire des atomes autorise différents mondes possibles mais n'en actualise qu'un seul. C'est donc le hasard qui est au fondement des agrégations d'atomes. Or présupposer que le hasard est seul responsable de la création du monde, c'est admettre une vision non finaliste de l'univers prédisposée à conférer une place à l'idée de mondes répétés indéfiniment.

On retrouve l'hypothèse physique de mondes répétés, au XIXe siècle, dans un contexte laplacien. Elle peut être résumée par la description qu'en donne Mill :

> The state of the whole universe at any instant, we believe to be the consequent of its state at the previous instant... And if any particular state of the entire universe could ever reccur a second time, all subsequent states would return too, and history would, like a circulating decimal of many figures, periodically repeat itself. (J. S. MILL, 1904, p. 400-401)

Elle sert notamment de point d'appui à différentes entreprises utopistes dont l'argument général procède dans l'idée que si des mondes différents du nôtre existent, c'est donc que l'histoire de notre propre monde n'est peut-être elle-même pas figée de manière rigide et qu'il ne faut pas céder au nécessitisme. La pluralité des mondes doit certainement beaucoup aux spéculations scocio-politiques qui ressortissent des mondes étrangers, ceux qui permettent aux philosophes du XVIIIe siècle de critiquer leur propre régime politique ou les mœurs de leur pays, comme le font Montesquieu dans les *Lettres Persanes* et Voltaire dans *Zadig* ou encore Fontenelle dans les *Entretiens sur la pluralité des mondes*.

De ce point de vue, le monde pourrait s'écarter du cours qu'il est présumé suivre. Le révolutionnaire français Auguste Blanqui reprend ainsi l'essentiel de l'argumentation atomiste pour défendre l'idée de mondes répétés, mais y ajoute l'idée de quasi-répétitions, c'est-à-dire de mondes répétés jusqu'à un certain embranchement temporel à partir duquel ils diffèrent en raison de circonstances légèrement différentes.

Blanqui considère que tout individu est la résultante de combinaisons de corps simples qui obéissent à des lois déterministes. Dans *L'éternité par les Astres*, il affirme que tout individu est inévitablement amené à se répéter un nombre infini de fois dans l'univers. Contre Leibniz, Blanqui

prétend ainsi que si le principe d'identité des indiscernables semble vraisemblable à l'échelle de la Terre, il est complètement démenti à l'échelle de l'univers :

> On prétend que la nature ne se répète jamais, et qu'il n'existe pas deux hommes, ni deux feuilles semblables. Cela est possible à la rigueur chez les hommes de notre terre, dont le chiffre total, assez restreint, est réparti entre plusieurs races. Mais il est, par milliers, des feuilles de chêne exactement pareilles, et des grains de sable, par milliards. (BLANQUI, 1872, p. 51)

Le raisonnement qui mène Blanqui à admettre l'hypothèse d'individus parfaitement ressemblants suit de très près les démonstrations d'Épicure et de Lucrèce. Mais est également défendu à partir des découvertes de Lavoisier au XVIIIe siècle. Comme les atomistes abdéritains, Blanqui considère que tout individu résulte de la combinaison agrégative de corps simples dans le vide. L'espace est éternel, sans commencement ni fin, et il s'étend à l'infini. Les corps simples qu'il contient sont en nombre infini. L'univers est une inépuisable source de matière, et contient donc une infinité d'individus.

Mais il reste un pas à faire pour déduire la certitude d'individus répétés : il faut encore que le nombre de types d'atomes et le nombre de leurs combinaisons soient finis.

> Chacun des corps simples est sans doute une quantité infinie, puisqu'ils forment à eux seuls toute la matière. Mais ce qui ne l'est pas, infini, c'est la variété de ces éléments qui ne dépassent pas cent. Fussent-ils mille, et cela n'est pas, le nombre des combinaisons-types s'accroîtrait jusqu'au fabuleux, mais ne pouvant atteindre à l'infini, resterait insignifiant en sa présence (BLANQUI, 1872, p. 52)

Ici, Blanqui peut étayer ses arguments par les découvertes, en chimie, de Lavoisier, qui confirment effectivement qu'il n'existe dans l'univers qu'un nombre fini de types différents de corps simples tels l'oxygène et l'hydrogène, qui figurent parmi les découvertes du chimiste. En outre, la chimie s'accorde également à reconnaître que le nombre de combinaisons des atomes est lui-même fini.

Si le nombre d'atomes est infini, et qu'en revanche leur type et le nombre de combinaisons qu'ils peuvent constituer est fini, cela implique que les mêmes combinaisons sont inévitablement amenées à se répéter [19].

[19] On retrouve le développement de cette idée chez Jacques Delevsky qui écrit, dans un article de 1945, qu'« il s'agit de déterminer alors la durée totale de tous les arrangements spatiaux possibles. Si cette durée totale est finie, l'histoire du monde doit se répéter nécessairement, après l'épuisement de toutes les combinaisons possibles, et il y aura alors nécessairement des Retours éternels. Si, par contre, la durée totale de tous les arrangements spatiaux possibles est infinie, les variations cosmologiques peuvent se succéder dans l'histoire du monde, sans nécessiter de s'épuiser et de se répéter ; le cycle répétitionnel éternel ne s'impose pas alors nécessairement, et l'histoire du monde pourra changer infiniment, sans se répéter » (DELEVSKY, 1945). Dans le même article, Delevsky propose une formule mathématique pour évaluer la période d'un cycle, c'est-à-dire la durée totale de tous les arrangements spatio-temporels possibles. Delevsky considère que cette période est égale à $\mathrm{T} = \sqrt{2\pi} \times m^{\mathrm{N}} \times \left(\frac{m}{e}\right)^{n} \times n^{0,5} \times t$ où t est la moyenne arithmétique des différentes durées de tous les arrangements spatio-temporels possibles, où m désigne le nombre de cellules (emplacements) du volume supposé limité et isotrope de l'espace, où N désigne le nombre fini de simples (points-particules) et où n est le nombre des valeurs différentes que peut prendre la durée du temps de chaque arrangement entre le maximum et le minimum des durées.

> Tel est notre point de départ pour affirmer la limitation des combinaisons différenciées de la matière et, par conséquent, leur insuffisance à semer de corps célestes les champs de l'étendue. Ces combinaisons, malgré leur multitude, ont un terme et, dès lors, doivent se répéter, pour atteindre à l'infini. (BLANQUI, 1872, p. 55)

Il y a donc, pour toute chose, qu'il s'agisse d'un grain de sable ou d'une galaxie entière, une infinité de répliques de cette chose. Notons en outre que ces répétitions peuvent autant être spatiales que temporelles. Des mondes indiscernables peuvent se succéder dans le temps, mais il peuvent aussi co-exister dans l'espace.

Le nombre fini des combinaisons d'atomes dont le type est lui-même restreint implique une importante conséquence : dans l'univers, la répétition est l'ordinaire, et c'est la singularité qui s'avère exceptionnelle. Tout ce qui existe, existe en quantité infinie.

> La nature tire chacun de ses ouvrages à milliards d'exemplaires. Dans la texture des astres, la similitude et la répétition forment la règle, la dissemblance et la variété, l'exception. (BLANQUI, 1872, p. 55)

En particulier, à l'échelle des planètes, Blanqui constate que « les astres différenciés [...] sont dès lors réduits à un chiffre limité, et l'infinité des globes implique dès lors l'infinité des répétitions » (BLANQUI, 1872, p. 58). On peut alors classer les corps célestes selon deux catégories. D'un côté, les originaux, et d'un autre coté, les copies :

> Les originaux, c'est l'ensemble des globes qui forment chacun un type spécial. Les copies, ce sont les répétitions, exemplaires ou épreuves de ce type. Le nombre des types originaux est borné, celui des copies ou répétitions, infini. C'est par lui que l'infini se constitue. Chaque type a derrière lui une armée de sosie dont le nombre est sans limites. (BLANQUI, 1872, p. 56)

Comme on le voit, la notion d'original n'est pas entendue ici de manière axiologique. L'originalité n'est pas un privilège qu'aurait l'une des occurrences par rapport à ses répétitions. Il s'agit du regroupement abstrait de toutes les occurrences indiscernables sous la même nomination. Dans le langage de C.S. Peirce, on pourrait reformuler ceci en parlant de mondes-types (les originaux) et de mondes-tokens (les répétitions).

L'univers de Blanqui n'est pas seulement constitué d'une infinité d'individus répliqués. Il contient aussi une infinité de mondes présentant des événements identiques jusqu'à un certain instant, et qui divergent ensuite. Car supposons que nous considérions une terre-jumelle où « pas un caillou, pas un arbre, pas un ruisseau, pas un animal, pas un homme, pas un incident, [...] n'ait trouvé sa place et sa minute dans le duplicata » (BLANQUI, 1872, p. 56), bref, considérons une « véritable terre-sosie », une copie conforme de notre Terre à l'instant présent. Si nous laissons filer le cours du temps à partir de ce même instant, bientôt, il se pourrait que des divergences apparaissent entre les deux planètes. Contre le strict déterminisme physique, Blanqui admet une liberté humaine, la volonté des hommes induisant des divergences entre les mondes. Car en effet,

> quel homme ne se trouve parfois en présence de deux carrières ? Celle dont il se détourne lui ferait une vie bien différente, tout en le laissant la même individualité. L'une conduit

> à la misère, à la honte, à la servitude. L'autre menait à la gloire, à la liberté. Ici une femme charmante et le bonheur ; là une furie et la désolation. Je parle pour les deux sexes. On prend au hasard ou au choix, n'importe, on n'échappe pas à la fatalité. Mais la fatalité ne trouve pas pied dans l'infini, qui ne connaît point l'alternative et a place pour tout. Une terre existe où l'homme suit la route dédaignée dans l'autre par le sosie. Son existence se dédouble, un globe pour chacune, puis se bifurque une seconde, une troisième fois, des milliers de fois. Il possède ainsi des sosies complets et des variantes innombrables de sosies, qui multiplient et représentent toujours sa personne, mais ne prennent que des lambeaux de sa destinée. (BLANQUI, 1872, p. 61)

La théorie de la bifurcation des mondes due à la volonté humaine constitue sans doute la véritable originalité de la théorie de Blanqui de l'éternel retour. Car l'auteur met en avant un modèle « branchant » d'une pluralité de mondes concrets, là où l'atomisme antique semble toujours présenter la pluralité d'une manière « parallèle » [20]. On peut alors considérer que les mondes possèdent tous le même segment temporel commun, et qu'ils divergent ensuite à partir d'un moment t. Cela conduit à voir les mondes comme générés par des fissions successives selon le schéma de la figure 2.1 [21].

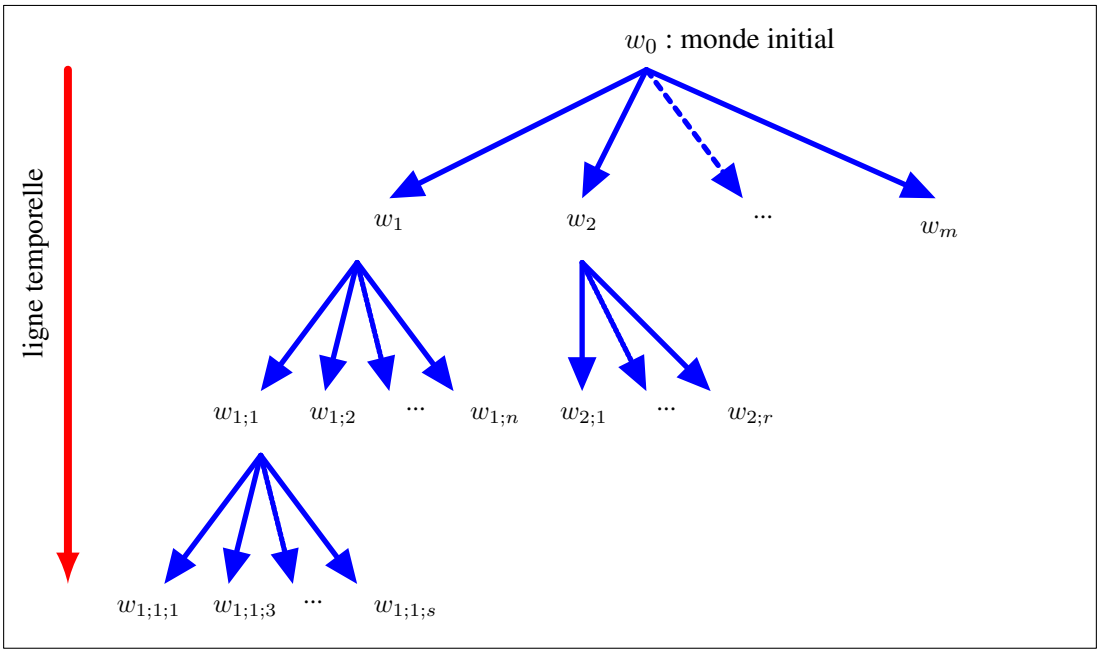


FIGURE 2.1 : Schéma de mondes branchants

Comme chez les stoïciens, la réflexion de Blanqui sur les mondes répétés se poursuit ensuite autour de la notion d'identité personnelle. La conclusion à laquelle aboutit Blanqui semble admettre une identité personnelle partagée entre un individu et ses doubles : « Tout ce qu'on aurait

[20] Il ne faut bien sûr pas prendre le mot au sens strictement géométrique et spatial, mais dans un sens plus large et imagé de mondes isolés les uns des autres, en tout cas, sans lien dynamique entre eux.

[21] Notons que nous retrouvons ce schéma dans différentes situations de duplications, comme par exemple dans le cas de reproductions successives d'une œuvre. Nous en reparlerons dans le chapitre 4 consacré aux œuvres textuelles.

pu être ici-bas, on l'est quelque part ailleurs. Outre son existence entière, de la naissance à la mort, que l'on vit sur une foule de terres, on en vit sur d'autres dix mille éditions différentes » (BLANQUI, 1872, p. 62). L'expression « on en vit » indique en effet que c'est nous-mêmes qui menons une carrière différente sur différents astres, et non un autre qui serait indiscernable mais numériquement différent. L'illusion d'une identité transmondaine agit-elle chez Blanqui comme une sorte de consolation stoïcienne face aux conditions défavorables de l'existence ? De fait, sa métaphysique frôle parfois la théologie lorsqu'il concède que par ses doubles, chaque individu est éternel :

> Ainsi, par la grâce de sa planète, chaque homme possède dans l'étendue un nombre sans fin de doublures qui vivent sa vie, absolument telle qu'il la vit lui-même. Il est infini et éternel dans la personne d'autres lui-même, non-seulement de son âge actuel, mais de tous ses âges. Il a simultanément, par milliards, à chaque seconde présente, des sosies qui naissent, d'autres qui meurent, d'autres dont l'âge s'échelonne, de seconde en seconde, depuis sa naissance jusqu'à sa mort. (BLANQUI, 1872, pp. 75–76)

Mais l'idée d'une identité entre les répliques issues de mondes différents paraît néanmoins absurde. Un individu et son sosie ne partagent que des ressemblances, ils sont numériquement différents. Aussi devrait-on au contraire reconnaître que tout ce qu'*on* est ici-bas, *on* ne l'est nulle autre part, justement, qu'ici bas. Un individu et son double ne sont tout simplement pas le même individu. D'un point de vue strictement physique, ce qui arrive à l'un peut n'avoir aucune sorte de conséquence sur ce qui arrive à l'autre. Nos contreparties nous restent inexorablement inaccessibles. Ce à quoi semble d'ailleurs consentir Blanqui lui-même dans le passage suivant :

> Pensez-vous qu'à l'autre bout de l'infini, dans quelque terre compatissante, le prince royal, arrivant trop tard sur Sadowa, ait permis au malheureux Benedeck de gagner sa bataille ?... Mais voici Pompée qui vient de perdre celle de Pharsale. Pauvre homme ! il s'en va chercher des consolations à Alexandrie, auprès de son bon ami le roi Ptolémée... César rira bien... Eh ! tout juste, il est en train de recevoir en plein sénat ses vingt-deux coups de poignard... Bah ! c'est sa ration quotidienne depuis le non-commencement du monde, et il les emmagasine avec une philosophie imperturbable. Il est vrai que ses sosies ne lui donnent pas l'alarme. Voilà le terrible ! on ne peut pas s'avertir. S'il était permis de faire passer l'histoire de sa vie, avec quelques bons conseils, aux doubles qu'on possède dans l'espace, on leur épargnerait bien des sottises et des chagrins … (BLANQUI, 1872, p. 64)

Remarquons dès lors que la théorie de Blanqui met en lumière la question de la caractéristique existentielle des doubles : quel intérêt, pour un habitant d'un certain monde de savoir que sa contrepartie mène une vie plus agréable dans un autre monde ? Cela ne semble même pas avoir valeur de consolation.

La répétition des mondes n'est pas contradictoire à l'idée d'une forme d'anti-nécessitarisme, voire de progressisme politique. Car il est vraisemblable qu'une idéologie se mêle aux spéculations physiques de Blanqui. Soulignons à ce titre que Blanqui est l'un des précurseurs de l'extrême gauche française. Ayant mené diverses activités révolutionnaires, il passa une grande partie de sa vie en détention - au point d'être surnommé « l'enfermé ». *L'éternité par les astres* fut d'ailleurs

rédigée à l'occasion de son dernier séjour en prison. De fait, la métaphysique des contreparties qu'il propose peut encourager la croyance qu'il n'y a pas de destinée inéluctable puisqu'il existe des variantes du monde. L'astronomie des mondes divergents sert ainsi de point d'appui pour l'engagement dans la lutte sociale. Les contradictions logiques qui émanent de l'affirmation d'une identité entre les contreparties de mondes différents sont négligées, parce que le but visé par Blanqui est peut-être moins de convaincre de la pertinence du modèle astronomique retenu que de persuader ses contemporains de la possibilité de changements politiques.

Les remarques précédentes indiquent ainsi l'importance de la pensée de mondes répétés dans un contexte moral, éthique ou progressiste. Même chez Nietzsche, l'un des penseurs les plus rétifs à l'idée de valeurs morales objectives, l'idée de la répétition temporelle du monde sert de crible pour sélectionner les actions à effectuer, en choisissant celles qu'on accepterait de revivre une infinité de fois. On peut mettre cette demande en rapport avec ce que prétend Giacomo Leopardi dans le *Zibaldone* :

> J'ai demandé à quelques personnes si elles seraient heureuses de revivre leur existence passée, à la condition de la revivre exactement telle qu'ils l'avaient vécue. [...] Pour ce qui est de vivre une seconde fois, tout le monde, y compris moi, s'en trouverait ravi ; mais avec ladite condition, plus personne. (LEOPARDI, 2000, p. 161)

Quelle est au juste la place de l'éternel retour chez Nietzsche ? Certains passages de son œuvre philosophique s'y réfèrent explicitement, par exemple celui-ci :

> Si, dans tout ce que tu veux faire, tu commences par te demander : "Est-il sûr que je veuille le faire un nombre infini de fois ?", ce sera pour toi le centre de gravité le plus solide. (NIETZSCHE, 1965, p. 242)

Nietzsche ne prétend pas que l'éternel retour soit un véritable principe métaphysique. Il s'agit d'abord d'une expérience de pensée visant un but pratique, sélectif. Le principe a pour objet d'opérer un choix. Comme l'écrit Gilles Deleuze [22], l'hypothèse de l'éternel retour donne une « loi pour l'autonomie de la volonté dégagée de toute morale : quoi que je veuille (ma paresse, ma gourmandise, ma lâcheté, mon vice comme ma vertu), je "dois" le vouloir de telle manière que j'en veuille aussi l'éternel Retour » (DELEUZE, 1969, p. 37). La conception rationnelle des mondes répétés à l'identique peut ainsi motiver notre manière d'agir.

Cosmogonies quantiques

Je terminerai ce tour d'horizon de la question physique des dupliques en abordant la question des multivers et des mondes parallèles en relativité générale et en physique quantique. En 1927, l'historien des sciences Abel Rey prétend ainsi que :

[22] Il faut préciser toutefois que Deleuze s'aventure ici un peu trop loin et qu'en plus de la pensée sélective, il confère également à l'éternel retour le rôle d'un « Être sélectif ». Ne se trouve reproduit dans l'éternel retour que ce qui a été sélectionné par la pensée : « parce que l'Être s'affirme du devenir, il expulse de soi tout ce qui contredit l'affirmation, toutes les formes du nihilisme et de la réaction : mauvaise conscience, ressentiment,... on ne les verra qu'une fois (DELEUZE, 1969, p. 38). Mais il ne semble pas indispensable d'ajouter au principe de pensée sélective une deuxième sorte qui serait, elle, métaphysique.

> L'Univers est une machine aveugle, construite de telle sorte qu'elle peut repasser une infinité de fois par les mêmes états. Son évolution est cyclique et l'orientation de cette évolution à un moment donné, ou pendant une période donnée, doit pouvoir se répéter semblablement de façon indéfinie. Autrement dit, la machine-univers est susceptible de restaurer son état initial. Mieux dit encore, il n'y a pas d'état initial, sauf celui que nous prenons arbitrairement pour origine, sur l'évolution d'un cycle. Il y a cycle et retour éternel. (REY, 1927, p. 14)

On trouve en effet l'hypothèse de mondes infiniment répétés dans les théories cosmologiques basées sur l'hypothèse d'une succession infinie de Big Bang et de Big Crunch. On retrouve ainsi

> [...] l'idée que, dans cet univers si vaste et immense, voire infini, se produisent régulièrement des faits identiques, et que l'on puisse établir des probabilités associées à la production de ces faits, plaide en faveur de la pluralité des mondes, et suggère même qu'existent très probablement, éloignées de nous et inaccessibles, répétées d'innombrable fois, des répliques de notre monde observable. (RABACHOU, 2016, p. 59)

Dans les thèses cosmogoniques interprétant les résultats de la physique contemporaine, l'inclusivité du monde est généralement clivée selon qu'elle joue un rôle épistémique ou ontologique. De fait, ces théories ont encouragé le développement de différentes hypothèses qui justifient l'idée de mondes multiples. Elles divergent cependant sur la nature de tels mondes, soit qu'il s'agisse d'univers causalement isolés les uns des autres au sein d'un même tout, soit qu'il s'agisse d'univers « superposés », soit qu'il s'agisse encore d'univers parallèles incommensurables.

Commençons par l'hypothèse de mondes isolés qui peut être tirée de la relativité générale. De fait, il est possible que certains objets, comme certains astres, nous restent absolument inaccessibles parce que nous sommes trop éloignés pour que leur lumière nous parvienne. Dans ce cas, nous serions épistémiquement exclus d'une partie de l'univers. Mais une chaîne causale s'étendrait sans doute encore entre nous et ces parties de l'espace. Notre isolement ne s'avérerait donc pas ontologique. Parler de « notre monde », ce serait donc évoquer un monde qui limite physiquement nos sphères d'action et d'observation.

L'une des interprétations de la théorie quantique affirme l'existence de mondes coïncidents quoique numériquement distincts et discernables. L'idée de coïncidents n'est d'ailleurs pas propre aux cosmogonies quantiques. En effet, rappelons-nous la remarque d'Arthur Eddington au sujet du dénombrement des objets ordinaires. Là où se tient une table ordinaire, il y aurait aussi une table physique, très différente de la première, quoiqu'elle coïnciderait avec elle :

> I have settled down to the task of writing these lectures and have drawn up my chairs to my two tables. Two tables! Yes; there are duplicates of every object about me - two tables, two chairs, two pens. (EDDINGTON, 1927, p. 1)

Ainsi, Eddington distingue la table du sens commun, qui est un objet étendu, coloré, substantiel, de la table scientifique, qui est incolore et qui contient plus de vide que de particules. Toute chose est ainsi dédoublée par autre chose. Aux objets ordinaires se superposent des objets physiques. Or, comme nous allons le voir, la superposition de coïncidents va occuper une place importante dans les théories quantiques.

Cette thèse de coïncidence trouve une continuation dans la physique quantique, car celle-ci prescrit que les photons se manifestent en adoptant deux attitudes antinomiques. En effet, l'expérience des fentes de Young montre qu'ils agissent à la fois comme des particules et comme des ondes. La théorie interprète ce résultat en stipulant que, tant qu'ils ne sont pas observés, les deux comportements sont simultanément présents. En revanche, lorsqu'ils sont observés, les photons agissent soit comme des interférences sans manifester de comportement particulaire, soit avec un comportement particulaire sans manifester d'interférences. L'expérience de pensée de Schrödinger illustre de manière spectaculaire ce résultat. Un chat est enfermé dans une boite contenant une particule radioactive et un dispositif d'empoisonnement déclenché par la désintégration de cette particule. Comme nous venons de le dire, en l'absence d'observateur, la théorie quantique prévoit que l'état de la particule est double, à la fois non-désintégrée et désintégrée. Nous devons alors en tirer la conséquence que le chat, dont la survie dépend de l'état de la particule, est lui-même double : à la fois vivant et non-vivant. Soulignons que la conclusion de l'expérience n'est pas que le chat est ou vivant ou mort, et qu'il est impossible de le savoir tant qu'on n'a pas ouvert la boite. La théorie quantique affirme précisément que le chat *est synchroniquement dans les deux états à la fois*.

On pourrait faire remarquer que le chat est le premier observateur de l'expérience qu'il subit [23], mais l'expérience de pensée que propose Schrödinger sert moins à convaincre de la pertinence de la théorie quantique que d'illustration destinée à mettre en exergue l'étrangeté de ses conséquences. Ce qui paraît particulièrement troublant dans l'expérience de pensée du chat de Schrödinger relève sans doute du changement d'échelle. Admettre une double carrière d'une entité aussi hypothétique que mal connue que l'est une particule quantique est une chose. Nous pouvons peut-être l'admettre s'il s'agit en fait de mettre en image une interprétation mathématique des fonctions probabilistes qui régissent le comportement d'entités physiques. Mais parler de la double vie d'un félin est tout autre chose. Nous acceptons cette conséquence bien plus difficilement dans le cas d'un individu aussi familier qu'un chat dont les conditions de persistance sont non-contradictoires.

La physique quantique repose essentiellement sur les probabilités, cela, non d'une manière simplement descriptive, mais d'une manière intrinsèque. L'indétermination joue un rôle au centre-même des objets physiques. En effet, leur comportement relève en partie de l'imprévisible, comme le stipule par exemple la thèse du mouvement brownien des particules. Toute la question est de savoir si le hasard qui règne en physique quantique est seulement épistémique (par exemple causé par la rencontre complexe de lignes causales), ou ontologique. Or l'hypothèse des multivers a pour vocation de réinterpréter les résultats de la théorie quantique de manière déterministe. La description quantique du comportement des particules repose sur une interprétation probabiliste de fonctions d'ondes comme amplitude de probabilités. Ces fonctions ont été découvertes par Schrödinger en 1925 puis développées dans les années 1925-1927 par Max Born et Werner Heisenberg. Elles indiquent la probabilité de la présence d'une particule et elles peuvent se trouver en combinaisons linéaires, donnant lieu à des états superposés. Cependant, lors d'une opération de mesure, l'objet quantique sera observé dans un état déterminé. C'est ce que les physiciens qualifient de phénomène de réduction du paquet d'ondes, dont la thèse prescrit l'effondrement de la fonction d'onde. Lors de l'intervention d'un observateur, le faisceau des états superposés est, *in*

[23] Einstein a d'ailleurs proposé une variante de l'expérience où le chat est remplacé par un baril de poudre explosive.

fine, projeté uniquement sur l'un d'entre eux. Selon l'interprétation classique de Bohr et d'Heisenberg, appelée interprétation de Copenhague, les équations mathématiques ne doivent cependant pas être interprétées comme des réalités, mais comme des représentations. Les tenants de cette interprétation insistent donc sur l'idée que les équations de la mécanique quantique ne sont que des descriptions de ce que nous pouvons connaître de la réalité, mais qu'elles ne décrivent pas la réalité en elle-même.

À l'opposé, certains physiciens comme David Bohm et Hugh Everett admettent que les équations mathématiques indiquent la nature même des choses. Parmi ces scientifiques, certains ont tenté de trouver des structures permettant de résoudre les paradoxes causés par l'interprétation quantique des observations. Une première manière de considérer les choses, la théorie de la décohérence élaborée par Dieter Zeh (ZEH, 1970) en 1970, affirme que l'état de superposition ne peut être maintenu qu'au prix d'une absence totale d'interactions avec l'environnement. L'idée est donc de mettre l'intentionnalité complètement hors de cause du phénomène d'effondrement de la fonction d'onde. C'est la notion d'interaction qui est ici centrale et l'on doit la substituer à celle d'observation. Pour un état de superposition, lorsqu'intervient une interaction avec une particule extérieure, celle-ci a pour conséquence la réalisation exclusive d'un état parmi ceux qui sont possibles. L'état de superposition est donc infiniment rare et ne concerne finalement que des objets extrêmement petits, à savoir, quelques particules isolées. La décohérence se produit donc même en l'absence de mesure ou d'observateur. Le chat de l'expérience de Schrödinger est donc dans un état parfaitement défini avant même que l'observateur n'ouvre la boite. Cela dit, le problème, pour la théorie de la décohérence, est d'expliquer comment survient une interaction extérieure quand on se situe à l'échelle de l'univers entier. La fonction d'onde ne montre pas le comportement réel des particules, mais seulement ce que nous pouvons connaître de ce comportement.

De nombreuses solutions alternatives à la théorie de la décohérence ont été données pour tenter de résoudre le paradoxe des états simultanés et contradictoires. Il n'est évidemment pas question d'en faire état ici, faute de compétence et de place. Nous pouvons toutefois mentionner schématiquement l'interprétation que Hugh Everett a donnée dans son mémoire de doctorat de 1957 (H. EVERETT, 1973). Everett considère que ce qui est mathématiquement déduit de la superposition des fonctions d'onde doit être littéralement interprété comme la réalité. Il y a donc bel et bien superposition de deux états. Plus encore, il y a superposition de deux *états de conscience*. Car selon Everett, ce qui doit être expliqué, ce n'est pas tant la simultanéité physique de deux événements contradictoires, mais le fait que l'observateur n'ait conscience que d'un seul état. Autrement dit, le mystère de la superposition passe de l'élément observé à l'observateur. Il s'agit en tout état de cause, pour Everett, de rejeter l'hypothèse d'effondrement de la fonction d'onde. Autrement dit, l'état de superposition de deux états contradictoires persiste même lors de l'observation. La théorie d'Everett propose donc de considérer une fonction d'onde dite « universelle » en ce qu'elle incorpore autant la conscience de observateur macroscopique que le phénomène étudié. La théorie prévoit que lors d'une mesure, la conscience de l'observateur se scinde en autant d'états que superpose la fonction d'onde. Deux états quantiques donneront ainsi lieu à deux états de conscience séparés. Une conscience ne sera donc jamais témoin que d'un seul état lors de l'observation. Bien sûr, cette explication laisse une part importante d'ombre sur la possibilité de coexistence des différents états de conscience des observateurs. Techniquement parlant, la théorie que développe Everett permet de rendre compte du principe physique d'unitarité qui permet aux vecteurs d'état

de conserver une norme constante. En cela, les physiciens reconnaissent qu'elle est pertinente d'un point de vue strictement formel, mais il reste plus difficile de définir ce que de telles bifurcations de conscience représentent vraiment dans la réalité.

L'interprétation d'Everett a incité des physiciens à souscrire complètement à l'idée de mondes multiples et coïncidents. En avançant sa théorie des embranchements de conscience, Everett défend la thèse de superposition en évitant l'hypothèse d'un effondrement de la fonction d'onde lors d'une mesure. Mais, avant 1970, Everett ne recourt jamais à la comparaison aux mondes parallèles. Ce sont, chez lui, les points de vue qui sont multipliés, et non les réalités observées. La théorie des mondes multiples sera essentiellement développée par Bryce Dewitt (DEWITT, 1973) qui, dans un article publié en 1973 et intitulé *The Many-Universes Interpretation of Quantum Mechanics*, réactualise la thèse d'Everett en l'adaptant à l'idée de mondes parallèles multiples. La démultiplication de ce qui relevait de branches de conscience chez Everett devient une démultiplication de mondes chez DeWitt.

Dans cette théorie des multiples univers, ou *multivers*, lorsqu'un atome change d'état d'énergie, il donne naissance à (au moins) deux mondes parallèles, l'un où il a changé d'état, et l'autre où il n'a pas encore changé d'état. Pour un nombre relativement faible de changements d'état, les mondes produits ne s'avéreront pas très différents les uns des autres. Or, étant donné que de tels changements se produisent à tout instant, le nombre de mondes quasi-indiscernables qui sont produits est conséquemment immense. Accepter la thèse de DeWitt, c'est donc considérer que vous, qui êtes en train de lire ce mémoire, avez de multiples contreparties qui font exactement la même chose dans des mondes quasi-similaires. Ce processus d'embranchements mondains pourrait fournir une explication à la flèche du temps, puisque les mondes deviennent de plus en plus nombreux au fur et à mesure de leur fission, et en outre, ils sont de plus en plus dissemblables en proportion de la distance temporelle qui nous sépare du moment de leur bifurcation. L'univers initial du Big-Bang aurait ainsi donné lieu à une ramification de mondes dont le nombre s'accroît au fur et à mesure du temps.

La théorie de DeWitt, qui a été plutôt bien acceptée par les cosmologues, mais plutôt mal par les physiciens, va finalement être amendée par David Deutsch (DEUTSCH, 1997). Selon Deutsch, il n'y a pas véritablement création d'univers à chaque changement d'états. Tous les univers, en nombre fini ou infini, existent depuis toujours, et sont tous indiscernables à l'origine. Les changements d'états ne sont pas des créations de mondes nouveaux, mais une distribution des modifications quantiques sur différents mondes. Il n'y a donc pas de fissions de monde dont le nombre augmenterait au fur et à mesure du temps, mais une différenciation de plus en plus importante de mondes originellement indiscernables les uns des autres. Cette théorie permet en outre de conférer une signification à la probabilité quantique d'un changement d'état, en la liant ainsi à la proportion des mondes qui réalisent ce changement. En particulier, une fois dans sa boite, le chat de Schrödinger est vivant dans la moitié des mondes où il est placé dans la boite et le chat est mort dans l'autre moitié des mondes où il est placé dans la boite.

Un argument que l'on peut mettre en avant pour rejeter les théories physiques qui proposent des mondes doubles inobservables est qu'elles ne répondent pas au critère poppérien des sciences. Ces théories sont constituées de thèses qui ne sont tout simplement pas réfutables. Rien ne permet donc de tester ces théories. En cela, elles sont de pures spéculations cosmogoniques ou métaphysiques, et finalement, ne sont pas très éloignées d'affirmations religieuses. Les partisans des

multivers répondent généralement que l'existence de ces derniers ne n'est pas une thèse mais une conséquence de la mécanique quantique, autrement dit, d'une théorie scientifique. Ainsi, même si les multivers ne sont pas observables et même si l'affirmation de leur existence n'est pas testable, cela n'est pas une raison suffisante pour les rejeter *a priori*, parce que leur existence est la conséquence logique d'une théorie scientifique qui, elle, est bien reconnue en tant que telle. D'autre part, la non testabilité du modèle des plurivers n'est pas logique, mais pratique. Il se pourrait donc que ce qui n'est pas testable aujourd'hui le soit demain, comme ce fut le cas avec la déformation gravitationnelle de la lumière émise par une étoile qu'Einstein avait prédit en 1936 mais qui ne fut observée que dans les années 80.

Après de nombreux détours, on voit donc ré-émerger, dans les théories physiques les plus audacieuses, l'hypothèse de Blanqui. En comparaison avec la théorie laplacienne, l'interprétation de la théorie quantique selon DeWitt met en jeu des mondes multiples qui ne sont pas juxtaposés, éparpillés dans l'univers, mais qui coïncident, qui sont comme superposés. Les différentes branches des fonctions d'onde ne pouvant interférer, en retour, les mondes qu'elles décrivent ne peuvent interagir. Et tout comme l'avait fait Blanqui à partir de sa propre théorie, Everett pronostique lui aussi que la théorie des mondes multiples implique l'éternité des individus : car pour un monde où une partie du corps vieillirait, il en existerait tout un tas dans lequel ce vieillissement n'aurait pas lieu. Il y aurait donc, en plus de mondes au vieillissement plus ou moins accéléré ou ralenti, des mondes, moins nombreux mais tout aussi réels, dans lequel un individu serait éternel. On retrouve donc, une fois encore, l'idée que le double exprime un attachement psychologique à l'idée consolatrice d'un corps de substitution. C'est peut-être l'une des raisons qui ont fait qu'une théorie aussi fantasque soit parvenue à persévérer en physique.

À la fin de *Univers parallèles*, son essai de synthèse sur les mondes multiples de la physique, Thomas Lepeltier fait remarquer que les spéculations sur ces mondes insistent souvent sur la question de l'identité personnelle,

> [...] sur l'idée que les autres « nous » sont nous, ou, pour le dire autrement, qu'il y a identité entre nous et tous les êtres qui sont comme nous et qui vivent dans ces univers parallèles ou multiples. [...] En somme, la thèse des univers parallèles ou multiples a le grand mérite de révéler le problème de l'identité personnelle : qu'est-ce qui fait que je sui moi, et que vous êtes vous ? (LEPELTIER, 2010, p. 252-255)

Mais évidemment, affirmer que l'individu d'un monde est identique à l'individu d'un autre au motif de leur indiscernabilité est très discutable, sinon complètement faux. Même en réduisant l'identité d'une personne à la configuration de ses atomes, il serait difficile de reconnaître comme identiques des individus de mondes distincts. Car dans la perspective des mondes branchants décrite ici, les configurations ne seraient les mêmes que de manière provisoire, elles changeraient à un certain moment du temps. Un individu étant constitué par l'ensemble de ses phases temporelles, nous aurions donc toujours, entre deux mondes distincts, des individus numériquement différents. Ainsi, les mondes des multivers ne peuvent pas vraiment servir de consolation à l'individu devant la mort. En revanche, ils peuvent favoriser l'idée de comparaisons, de possibilités.



La question des mondes répétés, que ceux-ci soient reproduits de manière spatiale ou temporelle, nous a confrontés aux notions d'authenticité et d'originalité. La tradition mimétique, dont relève la théorie platonicienne de la participation, consacre une place privilégiée à l'un des mondes en question, monde qualifié d'authentique en ce qu'il lui est réservé une importance métaphysique plus élevée que les autres (que le monde sensible dans le cas de Platon). Cette valeur ajoutée implique que ce qui est répété ne l'est jamais que de manière imparfaite. De ce point de vue, le double est toujours interlope : il lui est reproché d'exercer une forme d'usurpation, de manifestation corrompue d'une réalité à travers une représentation qui n'est jamais que partiellement imitative. L'authenticité est donc une notion essentiellement axiologique.

L'imitation est donc une relation engagée dans un rapport à l'authenticité, ce qui n'est pas nécessairement le cas de la représentation. L'originalité ne désigne que la précédence temporelle d'un monde, sans nécessairement ajouter un surcroît ontologique à l'entité qu'elle qualifie. Nous avons vu que dans le système physique des atomistes abdéritains, les mondes peuvent être *parfaitement* répétés, parce qu'un monde ne possède pas d'authenticité en rapport aux autres. Si un monde existe, dans le temps, avant sa répétition par un autre, ce dernier n'en est pas moins authentique, il n'est pas une imitation de l'original bien qu'il puisse en constituer une représentation.

A travers la discussion qui précède, nous avons également vu que la question des mondes répétés implique celle d'individus répétés. La question d'une éventuelle identité d'individus entre mondes physiques redondants est d'emblée réglée par la notion même de répétition : ne peut se répéter que ce qui est numériquement différent. Aussi, des individus provenant de deux mondes distincts sont inévitablement numériquement différents. Si c'est là notre conclusion en ce qui concerne l'éventualité d'une répétition de mondes physiques, il convient néanmoins, de reposer cette question dans le cas de mondes simplement possibles dans lequel intervient, comme nous allons le voir maintenant le problème des identités croisées.

2.1.2 L'espace logique et le double

Pensées contrefactuelles

Les théories pronostiquant un dédoublement de la réalité peuvent autant relever de la pure logique que de la physique. Dans le premier chapitre de *L'objet singulier*, Clément Rosset commente un passage de l'Iliade dans lequel Pandoros déplore avoir manqué sa cible :

> Un immortel sans doute à ses côtés se tient, les épaules dans un nuage enveloppées, et c'est lui qui tantôt aura fait dévier du but mon trait rapide. Car j'ai déjà tiré sur lui : je l'ai blessé près de l'épaule droite, au creux de la cuirasse ; je croyais bien alors l'envoyer dans l'Hadès. Je ne l'ai pas dompté pourtant : un dieu m'en veut ! (HOMÈRE, 1955a, p. V, 185-191)

Alors qu'il est certain d'avoir atteint mortellement son adversaire, Pandoros est bien obligé de constater que sa croyance n'est pas fidèle aux événements eux-mêmes. Cet écart entre les faits et leur représentation, entre l'ontologique et le doxastique, entre ce qui est arrivé et ce qui est cru, témoigne d'une forme psychologique de duplication. Pandoros croit ainsi assister à un « détournement du réel » (ROSSET, 1979, p. 11) dont seul un dieu pourrait être l'acteur. La vérité de droit, celle qui devait *légitimement* arriver, celle qui est attendue par Pandoros, est relayée par une

vérité de fait, ce qui arrive *vraiment*. À moins qu'il ne s'agisse là d'une forme psychologique de détournement due à Pandoros lui-même plutôt qu'à une quelconque divinité, détournement visant à remplacer un fait déplaisant par une alternative plus acceptable. La connaissance d'un fait - un but manqué -, est ainsi usurpée par une représentation alternative - un but atteint, voire, un but qui *aurait dû* être atteint.

À l'instar de Pandoros, il nous arrive d'imaginer en pensée des déviations du cours des événements, des scénarios contrefactuels plus ou moins vraisemblables. Nous fomentons des doubles du monde réel à partir de représentations s'écartant de ce qui a été, de ce qui est ou de ce qui sera. C'est notamment le cas lorsqu'on se demande si le monde aurait pu être radicalement différent. Mais comme le constate encore Rosset, nos transactions modales ordinaires préservent généralement la prérogative d'existence du monde réel :

> Habituellement, le double ne remet pas en question la réalité du monde en général [...]. Le double se contente le plus souvent d'évoquer une multiplication insolite de telle ou telle partie qui ne met pas pour autant en cause l'identité de l'ensemble, pas plus que la multiplication anarchique des cellules cancéreuses ne remet en cause l'identité du malade. (ROSSET, 1979, pp. 12-13)

Nos altérations du réel se cantonnent généralement à l'idée que tel événement particulier aurait *pu*, ou plus exactement, aurait *dû*, prendre une tournure différente. Le double relève alors de cet écart déploré entre une possibilité de droit, celle qu'on s'accorde et qu'on accepte comme légitime, et une réalité de fait, qui est souvent, une réalité « de trop ».

Qu'est ce qui incite à de telles entorses faites à la réalité ? Mais d'où tirons-nous de telles marges d'altérations ? Qu'est-ce qui permet cette libéralité des possibles ? Qu'est-ce qui nous laisse croire que les faits auraient pu s'écarter, avec une certaine marge de liberté, de leur réalisation effective ? Car si notre expérience nous montre que les choses sont ce qu'elles sont, elle ne semble pas nous prescrire directement ce qu'elles auraient pu être. Dès lors, qu'est-ce qui autorise de tels affranchissements de la réalité ?

Certains philosophes ont préconisé de considérables méfiances à l'égard de ce type de transactions modales. Hume a largement critiqué, sous l'angle particulier de la relation de cause à effet, le fait de conférer aux choses des possibilités qui excèdent leur seule actualité. De fait, il semble n'y avoir ici que deux alternatives : soit admettre que les choses eussent pu être différentes, dans les faits et selon une certaine marge, soit nier cette hypothèse. Dans ce dernier cas, on pourrait rejeter de manière globale les raisonnements modaux en disant qu'il ne s'agit finalement que de scénarios de fiction. Le contrefactuel, ce serait de l'illusoire.

À l'opposé de la libéralisation des possibles, la tragédie grecque opte pour une perspective fataliste des états de choses du monde qui finissent toujours par se réaliser, quels que soient les efforts entrepris pour y contrevenir et quelle que soit la diversité des configurations de situations élaborées. De ce point de vue, tout ce qui peut arriver, arrive, et tout ce qui est arrivé, devait arriver. Œdipe, informé de son destin, non seulement échoue à en modifier le cours, mais plus encore, provoque justement les circonstances propices au déroulement des événements qu'il cherche justement à détourner. Selon une pensée tragique, il n'y a donc pas lieu de supputer des doubles mondains, des alternatives modales, puisque ceux-ci ne sont finalement que des illusions. Au sens

métaphysique, le monde n'aurait pas pu être autre qu'il n'est. Dire que les choses auraient pu être autres, c'est faire le jeu d'un fantasme entretenu par méconnaissance, par mépris, ou par dégoût du réel.

Le fatalisme se rencontre aussi dans l'histoire des idées philosophiques. En témoigne notamment la philosophie de Spinoza qui soutient que la possibilité d'une chose implique *ipso facto* sa nécessité. Cette thèse conduit inévitablement à un effondrement modal, puisque admettre que la possibilité implique la nécessité, c'est rendre factuel tout de ce qui relève du possible. En outre, si nous rejetons l'hypothèse de vies alternatives, nous optons alors pour une forme de nécessité généralisée, ou *nécessitarisme*, dont la thèse principale pourrait être résumée par le principe fataliste suivant : rien de ce qui existe n'aurait pu être autrement, seul l'actuel est possible. Admettre ce principe, c'est affirmer que chaque façon dont une chose est, est une façon dont cette chose est nécessairement. Admettre ceci, c'est donner entièrement raison au fatalisme tragique. Mais cette thèse, radicale, est certainement très éloignée de nos intuitions ordinaires. Le nécessitarisme tragique est sans doute une philosophie morale visant à nous aguerrir quand nous sommes confrontés à des situations particulièrement difficiles. Mais malheureusement, il contrevient aussi à un certain nombre de nos manières d'amender nos actions. La croyance au possible nous permet en effet d'infléchir nos actions en les ajustant par essais-erreurs aux problèmes que nous rencontrons. Si l'intuition de possible s'avérait fausse, il faudrait néanmoins saluer la grande utilité de cette erreur...

Dans la *Confessio Philosophi*, Leibniz pose la question de la nécessité des choses qui adviendront : « tout ce qui sera ne sera-t-il pas absolument nécessaire, comme tout ce qui a été a nécessairement été et tout ce qui est présentement est nécessairement ? » (Gottfried Wilhelm LEIBNIZ, 1993, p. 57) À cette question, Leibniz répond :

> Au contraire ! Cela est faux, sauf s'il est entendu qu'il y a une réduplication et que votre proposition contient une ellipse dont les hommes sont coutumiers pour ne pas dire deux fois la même chose. Car votre proposition signifie : « tout ce qui est, il est nécessaire, s'il est, qu'il soit », ou, en substituant à nécessaire sa définition, « tout ce qui est futur, on ne peut pas concevoir, s'il est futur, qu'il ne soit pas futur ». Si l'on omet la réduplication, la proposition est fausse. (Gottfried Wilhelm LEIBNIZ, 1993, p. 57)

Ce sur quoi insiste ici Leibniz, c'est qu'il faut se garder de confondre l'irrévocable et le nécessaire. Cette confusion repose sur l'inférence spécieuse consistant à passer de l'idée que l'existence d'une chose implique sa nécessité à l'idée qu'il était nécessaire qu'elle le fut. Il est vrai (et trivial) de tout ce qui s'est déjà réalisé, que nécessairement, il est réalisé. En revanche, il n'est pas vrai de tout ce qui se réalise, qu'il se réalise d'une façon nécessaire.

Ainsi, nos projections modales proviennent sans doute de nos expériences vécues, parce que celles-ci nous incitent à croire que dans le futur, une chose pourra être qualitativement différente de ce qu'elle est actuellement. Nous savons bien que la fleur de lilas qui est éclose aujourd'hui sera fanée demain. L'observation que les choses changent et le souvenir de ce constat sont certainement les principales raisons qui incitent à accréditer l'hypothèse qu'une chose aurait pu être différente de ce qu'elle est actuellement. Autrement dit, les changements temporels plaident en faveur de l'idée modale de contrefaits : l'hypothèse qu'une chose deviendra différente conduit à l'hypothèse qu'une chose aurait pu être différente. Cela nous incite à croire que les particuliers concrets to-

lèrent non seulement des changements temporels, mais également à croire en des changements seulement possibles, qui *pourraient ou auraient pu les affecter* dans une situation différente. Nos intuitions nous incitent donc à concevoir parfois des perspectives différentes sur la réalité. Nous déplaçons notre point de vue, comme en faisant un pas sur le coté, et outrepassons en imagination les frontières du réel, élaborant des contrefaçons du monde actuel, des duplications alternatives de la réalité.

Logique modale

Comme on le voit, la modalité entretien un rapport étroit avec le thème du double. Afin de mieux dégager la particularité des expressions modales, commençons par examiner une simple proposition factuelle comme celle exprimée par la phrase :

(1) Emmanuel Macron a gagné les élections présidentielles françaises de 2017.

Cette phrase exprime un état de fait dont la vérité peut être éprouvée par une vérification de sa correspondance à un état de choses. Comme Pandoros placé devant le constat que son adversaire est toujours sur pied, force est de constater que Macron est l'actuel président de la République [24]. En outre, cette proposition relève d'une pensée singulière, c'est-à-dire d'une pensée à propos d'un individu dont la référence peut être donnée par ce que Bertrand Russell appelle une « relation d'accointance » (*acquaintance*) (RUSSELL, 1910) autrement dit, une relation qui est définie par la désignation directe d'une entité matérielle. Le recours au monde nous permet de vérifier que l'entité visée possède bien les prédicats qui lui sont attribués dans l'énoncé. Comparativement, il apparaît en revanche que nous ne disposons plus du même recours pour évaluer la valeur de vérité des propositions exprimées par les phrases suivantes :

(2) Emmanuel Macron aurait pu perdre les élections présidentielles de 2017.

(3) Emmanuel Macron aurait pu être comédien.

(4) Emmanuel Macron aurait pu être plus petit que Nicolas Sarkozy.

(5) Emmanuel Macron aurait pu être plus petit (qu'il n'est).

(6) Emmanuel Macron aurait pu avoir un frère jumeau.

(7) Emmanuel Macron aurait pu ne pas exister.

Puisque de tels énoncés concernent des possibles non actualisés (nous dirons dorénavant des « simplement possibles »), nous devons déterminer au juste sur quels vérifacteurs faire reposer les valeurs de vérité que nous attribuons à ces énoncés. Comment traduire formellement de telles intuitions concernant la possibilité, la nécessité et l'impossibilité ? Nos intuitions plaidant en leur faveur, ces énoncés semblent ainsi exprimer des alternatives contrefactuelles. Remarquons qu'en comparaison de l'exemple **(1)**, ces nouvelles phrases substituent à l'usage de l'indicatif celui du conditionnel. Ce mode grammatical prévient du statut particulier des propositions exprimées, à savoir, la catégorie des propositions *modales*.

[24] Ce passage du mémoire a été écrit en France en 2020.

Afin d'interpréter formellement des expressions simplement possibles, les logiciens ont ajouté deux nouveaux opérateurs aux opérateurs de la logique classique. D'une part le losange $\Diamond$, qui traduit la possibilité, et d'autre part le carré $\Box$, qui traduit la nécessité. Ainsi, l'expression modale « il est possible que ϕ » s'écrit formellement « $\Diamond\phi$ » et l'expression « il est nécessaire que ϕ » se note « $\Box\phi$ ». « $\Box\neg\phi$ » signifie qu'il est nécessaire que non-ϕ, et « $\Diamond\neg\phi$ » signifie qu'il est possible que non-ϕ. Ce qui n'est pas nécessaire est dit contingent ($\neg\Box\phi$), et ce qui n'est pas possible est dit impossible ($\neg\Diamond\phi$). Il est à remarquer que les deux opérateurs de modalité sont inter-définissables. La nécessité peut s'exprimer par la possibilité : « ϕ est nécessaire » s'interprète aussi :

$$\Box\phi = \neg\Diamond\neg\phi$$

Autrement dit, « ϕ est nécessaire » signifie que $\neg\phi$ est impossible . D'autre part, « ϕ n'est pas nécessaire », qui se traduit par :

$$\neg\Box\phi \iff \Diamond\neg\phi$$

est équivalent à « il est possible que $\neg\phi$ ». « Il n'est pas nécessaire que $\neg\phi$ », soit :

$$\neg\Box\neg\phi \iff \Diamond\phi$$

est équivalente à « il est possible que ϕ ». Et enfin « nécessairement $\neg\phi$ » soit encore :

$$\Box\neg\phi \iff \neg\Diamond\phi$$

est équivalente à « il est impossible que ϕ ». Les relations qui précèdent peuvent être représentées sous la forme du carré logique d'Apulée (voir la figure 2.2) : la nécessité est contradictoire à la contingence, la possibilité est contradictoire à l'impossibilité, la nécessité est contraire à l'impossibilité, la nécessité implique la possibilité et l'impossibilité implique la contingence.

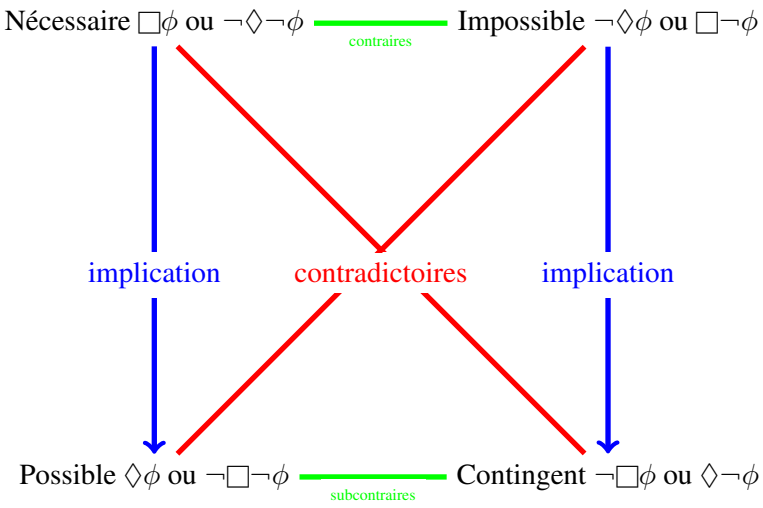


FIGURE 2.2 : Carré logique d'Apulée des opérateurs de possibilité et de nécessité

Mais revenons aux exemples précédents. Comme on peut le constater, parmi eux certains révisent les états de choses du monde (comme par exemple les énoncés **(2)** et **(3)**), d'autres augmentent **(6)** ou réduisent **(7)** le réservoir ontologique de ce qui existe dans le monde. L'exemple **(5)** implique un croisement entre un individu simplement possible et un individu actuel, une comparaison entre le Emmanuel Macron actuel et un autre-même, différent par la taille. Contrairement à l'exemple **(1)**, il nous est difficile d'attribuer directement une valeur de vérité aux propositions exprimées par de tels énoncés. Cependant, encore une fois, nos intuitions modales plaident en faveur de leur vérité. Mais comment évaluer celle-ci ? Sur quoi fonder la croyance en la véracité de telles allégations ? Par contraste, considérons quelques exemples supplémentaires :

(8) Emmanuel Macron aurait pu être un calamar.

(9) Emmanuel Macron aurait pu être un presse-purée.

(10) Emmanuel Macron aurait pu être Nicolas Sarkozy.

(11) Emmanuel Macron aurait pu se déplacer plus rapidement que la lumière.

(12) Emmanuel Macron aurait pu ne pas être identique à lui-même.

(13) Emmanuel Macron aurait pu être un personnage de fiction.

(14) Batman aurait pu être Emmanuel Macron.

À la différence des cas précédents, ces nouveaux énoncés semblent engager des contrefaits pesant bien plus lourdement dans l'ordre de nos croyances. S'ils paraissent également impossibles, il appert néanmoins que la nécessité qui les contraint pèse différemment d'un exemple à l'autre. Un adepte de la théorie de la réincarnation pourrait éventuellement admettre l'énoncé **(8)**, mais refuserait probablement l'énoncé suivant. Les énoncés **(8)** et **(9)** demandent l'abandon de l'appartenance spécifique de Emmanuel Macron, ce qui pourrait s'avérer fatal à sa persistance en tant qu'individu. Quant à l'énoncé **(10)**, il parait ambivalent, semblant signifier ou bien que deux individus auraient pu n'en faire qu'un seul, ou bien que Macron aurait pu prendre le rôle de Sarkozy. Ce dernier cas peut d'ailleurs être interprété de façon non paradoxale en considérant que l'occurrence « Sarkozy » n'est pas un nom propre logique dénotant son porteur mais un prédicat indiquant seulement le rôle joué [25].

Les énoncés **(11)** et **(12)** expriment des impossibilités de différentes natures, la première relevant de la physique, la deuxième de la logique. Il n'est pas logiquement impossible que les lois physiques eussent pu être différentes de ce qu'elles sont actuellement. Peut-être que d'autres configurations du monde auraient pu rendre possibles les voyages à des vitesses supraluminiques. L'impossibilité décrite par l'énoncé **(11)** est donc dérivée des lois de la physique. Il s'agit d'une nécessité causale, nomologique. Mais cela n'implique pas pour autant que les lois dont découle cette nécessité soient elles-mêmes nécessaires. En revanche, l'impossibilité dont relève l'exemple **(12)** ne semble pas relative à la manière dont le monde est actuellement ; elle semble au contraire absolue car d'aucune configuration possible du monde nous ne pourrions envisager l'éventualité d'une entité qui ne serait pas identique à elle-même. Nous avons là une impossibilité logique et aucune révision des données initiales du monde ne pourrait rendre l'énoncé **(12)** possible. Peut-être n'est-elle pas même concevable.

[25] Merci à Manuel Rebuschi pour avoir attiré mon attention sur ce point.

En résumé, les impossibilités peuvent donc être de différents ordres : logiques si elles enfreignent les axiomes de la logique, métaphysiques si elles se rapportent à l'essence ou à l'identité des choses, nomologiques si elles concernent les lois physiques. Elles peuvent également avoir un rapport doxastique, épistémique ou déontique. Remarquons aussi que l'impossibilité logique implique toutes les autres, mais qu'en revanche l'impossibilité épistémique ou nomologique n'implique pas l'impossibilité logique. De tels énoncés modaux, dont la valeur de vérité est identique dans toutes les configurations possibles du monde, expriment des propositions *nécessaires*. En revanche, des propositions qui admettent des valeurs fluctuantes, sont qualifiées de *contingentes*.

Dans un article (KRIPKE, 1963) séminal de 1963, Saul Kripke a introduit la structure de mondes possibles, pour conférer une sémantique aux expressions modales. L'idée est d'interpréter la vérité d'un énoncé modal par une structure de mondes possibles vérifiant cet énoncé, mondes qui sont accessibles selon certaines relations. On peut faire remonter la notion de monde possible à différentes sources. Elle apparaît bien sûr dans la philosophie de Leibniz, mais on la trouve aussi dans la conception cosmogonique des atomistes abdéritains, et dans la philosophie morale de Fontenelle (FONTENELLE, 1852). On pourra également noter une probable influence linguistique propre à l'élocution anglaise « *ways things could have been* », que, par comparaison, le français tend à formuler « façon dont les choses auraient pu être », ou « tournures que les événements auraient pu prendre ». Ainsi, la langue de Shakespeare utilise une image géographique de *chemins possibles* entre mondes vus comme des lieux, des emplacements. En outre, soulignons également que l'expression anglaise incite à penser que ce sont les choses elles-mêmes qui empruntent ces différents chemins. Elle donne l'image d'une sorte d'onto-géographie dont les frontières sont franchies par nos assertions modales. Ce sont les choses toutes entières qui sont conviées à cheminer selon des directions distinctes. En comparaison, l'expression française favorise l'idée que ce ne sont pas les choses elles-mêmes qui sont transportées, mais que ce sont leurs propriétés, leurs *manières d'être* qui peuvent prendre ou non un tour différent. Ces divergences de métaphores produites par l'anglais et le français se retrouvent dans les conceptions différentes des mondes possibles que nous exposerons plus loin.

Pour être formalisées dans ce langage, les expressions modales écrites en langue naturelle, mettons les expressions françaises au conditionnel, sont transformées en expressions à l'indicatif au moyen de la structure de mondes décrite précédemment. Par exemple, la phrase « Emmanuel Macron aurait pu ne pas être élu président » est retraduite par l'expression :

[P1] « Il y a au moins un monde logiquement possible accessible depuis le monde actuel dans lequel Emmanuel Macron n'a pas été élu président. »

Reconsidérons les exemples précédents en les retraduisant dans l'interprétation sémantique des mondes possibles :

(1*) Il y a un monde possible (le monde actuel) où Emmanuel Macron a gagné les élections.
(2*) Il y a un monde possible où Emmanuel Macron a perdu les élections.
(3*) Il y a un monde possible où Emmanuel Macron est comédien.
(4*) Il y a un monde possible où Emmanuel Macron est plus petit que Nicolas Sarkozy.
(5*) Il y a un monde possible où Emmanuel Macron (d'ici) est plus petit là-bas.

(6*) Il y a un monde possible où Emmanuel Macron a un frère jumeau.

(7*) Il y a un monde possible où Emmanuel Macron n'existe pas.

Et en considérant une possibilité seulement logique, les expressions précédentes seraient interprétées ainsi (quitte à s'avérer fausses) :

(8*) Il y a un monde possible où Emmanuel Macron est un calamar.

(9*) Il y a un monde possible où Emmanuel Macron est un presse-purée.

(10*) Il y a un monde possible où Emmanuel Macron est Nicolas Sarkozy.

(11*) Il y a un monde possible où Emmanuel Macron se déplace plus rapidement que la lumière.

(12*) Il y a un monde possible où Emmanuel Macron n'est pas identique à lui-même.

(13*) Il y a un monde possible où Emmanuel Macron est un personnage de fiction.

(14*) Il y a un monde possible où Batman est Emmanuel Macron.

La notion de monde possible

La sémantique des mondes possibles repose donc sur l'idée qu'une phrase de la forme « il est possible que ϕ » est vraie dans un monde w si et seulement si ϕ est le cas dans un monde w' accessible depuis w. Cette manière de procéder, en particularisant la valeur des énoncés à des mondes, peut finalement être rapprochée de la manière de conférer la vérité aux propositions qui sont exprimées dans un discours qui traitent de choses passées ou non encore advenues, en les considérant comme des faits advenus à *d'autres moments particularisés*.

Il reste néanmoins à préciser alors ce que l'on entend par la notion de « monde possible ». De quoi s'agit-il au juste ? Existent-ils vraiment ? Et si c'est le cas, existent-ils de manière concrète, comme le monde que nous habitons, ou au contraire, existent-ils de manière purement abstraite ? Et en quel sens s'agirait-il d'entités abstraites ? Seraient-elles seulement linguistiques, ou s'agirait-il de propriétés métaphysiques ?

Historiquement parlant l'introduction de la notion de monde possible en philosophie est due à Leibniz, qu'il décrit ainsi :

> J'appelle monde toute la suite et toute la collection des choses existantes, afin qu'on ne dise point que plusieurs pouvaient exister en différents temps et différents lieux. Car il faudrait les compter tous ensemble pour un monde, ou si vous voulez, pour un univers. (Gottfried Wilhelm LEIBNIZ, 1839, p. 108)

Quand Leibniz introduit cette notion, c'est pour rendre gloire au monde ayant la prérogative de l'actualité. Car d'après Leibniz, seul notre monde est réel, les autres ne sont que de simples émanations tirées de l'esprit d'un dieu :

> Even if it certain that what is more perfect will exist, still, the less perfect is none the less possible. Propositions of fact involve existence. But the notion of existence is such, that the

> existence is the sort of state of the universe which GOD chooses. But GOD freely chooses what is more perfect. Thus finaly a free action is involved [26]. (G. W. LEIBNIZ, 1903, p. 405)

S'il existe des mondes possibles, ils s'avèrent cependant *incompossibles* entre eux. Autrement dit, plusieurs mondes ne peuvent être réalisés en même temps. Il n'y en a en fait qu'un seul qui advient. L'exclusivité du monde est réalisée dans le système métaphysique de Leibniz par le truchement d'un processus sélectif. La métaphysique des mondes possibles est alors complètement rabattue sur une théologie : c'est un démiurge qui sélectionne le monde qui est instancié. Mais il reste à expliquer ce qui justifie un choix plutôt qu'un autre. Ici, Leibniz utilise le principe de raison suffisante : « jamais rien n'arrive sans qu'il y ait une cause ou du moins une raison déterminante, c'est-à-dire qui puisse servir à rendre raison *a priori* pourquoi cela est existant plutôt que non existant et pourquoi cela est ainsi plutôt que de toute autre façon » (Gottfried Wilhelm LEIBNIZ, 1969, p. I, 44) [27] qui implique qu'il doit y avoir une raison à ce choix. Selon Leibniz, la raison du choix d'un monde plutôt qu'un autre, c'est sa perfection. C'est parce qu'il est le plus parfait que notre monde a été sélectionné. Le critère de sélection est donc axiologique : le monde choisi est le meilleur des mondes possibles, entendu comme le monde qui offre la plus grande continuité et plus grand ordre entre ses composantes.

Une objection a parfois été adressée à la conception théologico-optimiste de Leibniz. Est-ce parce qu'il est le meilleur monde possible que nous l'habitons, ou est-ce parce que nous l'habitons que nous croyons qu'il est le meilleur possible ? Comment pourrions-nous alors être certains que nous sommes bien les *résidents* du meilleur des mondes, et non les « résidus » d'un autre monde possible non réalisé ? Comment savons-nous que notre existence dans ce monde est l'assurance de l'actualité du monde ? On pourra répondre que c'est justement parce que l'existence est la prérogative de l'actualité. Mais pour de nombreux philosophes, la question de l'existence d'un individu dans un monde possible doit être désolidarisée de la question de l'actualité d'un monde [28].

Dans *La pluralité des mondes*, David Lewis constate que la notion de monde possible est mal définie, en particulier chez Leibniz : « Il pourra paraître surprenant que ce livre sur les mondes possibles ne contienne également aucune discussion des conceptions de Leibniz. Est-ce parce qu'elles ne me semblent pas dignes de sérieuse attention ? Loin de là. Mais quand je lis ce que les historiens sérieux de la philosophie ont à dire, je me persuade que le fait de savoir en quoi consistaient exacte-

[26] La traduction du texte latin de Leibniz est due à Robert Merrihew Addam dans (ADAMS, 1974).

[27] Rappelons que Schopenhauer adresse de manifestes réserves à l'égard de l'attribution auto-proclamée du principe de raison suffisante par Leibniz : « Il le proclame très pompeusement dans plusieurs passages de ses œuvres, s'en fait accroître énormément à cet égard, et se pose comme s'il venait de le découvrir ; mais il n'en sait rien dire de plus, si ce n'est toujours que chaque chose, en général et en particulier doit avoir une raison suffisante d'être telle et non autre ; mais le monde savait ça parfaitement avant qu'il vînt le dire » (SCHOPENHAUER, 1966). Soulignons aussi que Platon a donné, dans le *Timée*, une formulation naïve du principe : « tout ce qui naît, naît nécessairement par l'action d'une cause, car il est impossible que quoi que ce soit puisse naître sans cause ». Cette formulation a le mérite de permettre de distinguer deux types de philosophies : d'un côté celles, empêtrées d'une façon ou d'une autre dans une quelconque idéologie religieuse (théiste, naturaliste, etc.) et qui sont par nature des philosophie du double en ce qu'elles cherchent une idée de nécessité en dehors des choses elles-mêmes, en interrogeant avec urgence leur *raison d'être*, et d'un autre côté, celles, artificialistes, qui rejettent complètement le principe de raison suffisante au profit du hasard : si les choses sont engendrées les unes par les autres, il n'est pourtant aucune raison de leur existence, aucune nécessité dans leur émergence. Comme on le voit, ces deux voies philosophiques attachent des importances disparates aux notions de nécessité et de possibilité.

[28] C'est notamment ce que défend Robert Merrihew Adams dans *Theories of Actuality* (ADAMS, 1974, p. 212).

ment les conceptions leibniziennes n'est pas une question facile » (LEWIS, 2007, p. 12). Pourtant, si nous parlons d'identité entre deux individus résidant dans des mondes différents, il faut commencer par éclaircir plus exactement ce que nous entendons au juste par « monde possible », ou plus précisément, il faut commencer par déterminer la teneur ontologique de telles choses. Comme le remarque Gideon Rosen, « accepter l'analyse des mondes possibles, c'est s'engager à croire en d'autres mondes possibles » (ROSEN, 1990, p. 327). Néanmoins, peu de philosophes sont prêts à accepter cet engagement. Dès lors, le philosophe qui analyse les discours modaux en employant la structure des mondes possibles, ce philosophe court « dans de telles circonstances, un grand risque de double-pensée : asserter l'existence de mondes à un moment, et la rejeter à d'autres » (ROSEN, 1990, p. 328).

La notion de mondes possibles est analysée de différentes manières. Le concrétisme, représenté principalement par David Lewis, soutient que les mondes existent concrètement, de manière matérielle. Inversement, l'abstractionnisme, représenté par Plantinga, Kripke et Stalnaker, soutient que les mondes possibles ne sont que des abstractions. Selon ce point de vue, ils peuvent alors relever d'états de choses non réalisés (PLANTINGA, 1974), d'ensembles maximaux de propositions (ADAMS, 1974), ou d'ensemble de propriétés (STALNAKER, 1976). Selon le point de vue combinatoire, les mondes possibles sont des recombinaisons de simples ((WITTGENSTEIN, 1993),(W. V. O. QUINE, 1977), (M. J. CRESSWELL, 1979)). Il peuvent aussi être considérés comme des parties du monde actuel. C'est notamment la conception de Nelson Goodman qui écrit, dans *Fact, Fiction and Forecast*, que : « All possible world lie in the within of the actual one. » (GOODMAN, 1973, p. 57). Enfin, selon un dernier point de vue, les mondes possibles ne sont que des fictions (ROSEN & DORR, 2002). De ce point de vue, les mondes simplement possibles n'ont aucun sorte d'existence ou d'être. Il ne sont que des dispositifs utiles, médiatisés par des jeux de croyance feinte. Il ne sont que des fictions utiles imaginées à partir du monde actuel.

On peut donc constater que les deux premières conceptions des mondes possibles confèrent une sorte d'existence à de multiples mondes, qu'ils soient vus comme des entités concrètes ou abstraites. Les deux conceptions suivantes n'admettent en revanche d'existence qu'au sens d'entités mentales ou d'actes de croyances feintes.

Les mondes possibles comme entités concrètes

Considérons alors deux mondes possibles quelconques w et w'. Étant donné que le monde possible w' est par définition complet, la proposition ϕ_w affirmant que « le monde w est actuel » est, dans le monde w', soit vraie, soit fausse. Comme le constate Robert Adams (ADAMS, 1974, p. 201) nous avons alors deux possibilités. Soit nous considérons que, pour tout monde possible w, la proposition ϕ_w est vraie dans w et fausse partout ailleurs. Soit nous considérons qu'il existe un monde possible w dans lequel cette proposition est vraie dans tous les mondes possibles, et que, quel que soit $w' \neq w$, $\phi_{w'}$ n'est vraie dans aucun monde possible.

Examinons d'abord la première possibilité. là une position relativiste et contingentiste : tous les mondes sont actuels de leur point de vue, et l'actualité leur est contingente, puisqu'il y a toujours un monde accessible dans lequel ϕ_w est fausse.

Cette conception implique ainsi une thèse indexicale de l'actualité. L'actualité du monde ne concernerait pas une propriété particulière qu'il serait le seul à exemplifier, comme l'existence,

mais une façon d'en parler pour ceux qui y résident, comme l'affirme le réalisme modal de Lewis. Cela résout, en quelque sorte, le problème métaphysique de ce qui confère l'actualité d'un monde. Mais la solution consiste à évacuer le problème, puisque l'actualité ne relève plus d'une prérogative ontologique, mais d'un fait linguistique. Comme l'écrit Lewis :

> the strongest evidence for the indexical analysis of actuality is that it explains why skepticism about our own actuality is absurd. How do we know that we are not the unactualized possible inhabitants of some unactualized possible world? We can give no evidence: what ever feature of our world we may mention, it is shared by other worlds that are not actual. Some unactualized grass is no less green, some unactualized dollars buy no less (unactualized) bread, some unactualized philosophers are no less sure they are actual. Either we know in some utterly mysterious way that we are actual; or we do not know it at all. But of course we do know it. The indexical analysis of actuality explains how we know it: in the same way I know that I am me, that this time is the present, or that I am here. All such sentences as "This is the actual world", "I am actual", "I actually exist", and the like are true on any possible occasion of utterance in any possible world. That is why skepticism about our own actuality is absurd. (LEWIS, 1970, p. 185)

La théorie indexicale écarte donc l'idée d'une prérogative ontologique et favorise l'idée qu'il y a plusieurs mondes concrets plutôt qu'un seul, autrement dit, penche en faveur des doubles ontologiques du monde. Pourtant cela semble assez contre-intuitif. Nous ne pensons généralement pas que la différence entre Emmanuel Macron et Sherlock Holmes soit seulement liée à ce que nous habitons le même monde que le premier à l'exclusion du second. Nous ne pensons pas que l'actualité soit seulement une question de leur relation à notre égard. Nous pensons au contraire que le premier est doté d'une prérogative que ne possède pas l'autre. L'un existe, l'autre pas. Cette intuition joue un rôle majeur en faveur de l'idée que l'actualité est une notion absolue, et non relative.

Le point de départ du réalisme modal de David Lewis est ancré sur l'idée que la sémantique des mondes possibles a pour objet d'expliquer nos attitudes modales. Ce ne sont pas ces dernières qui doivent fournir un sens aux mondes possibles, mais bien les mondes possibles qui doivent expliquer nos attitudes modales. Selon Lewis, cela implique de prendre au sérieux les mondes possibles :

> Il est incontestablement vrai que les choses auraient pu être autrement que la manière dont elles sont. Je crois comme vous que les choses auraient pu être, de bien des manières, différentes. Mais qu'est-ce que cela signifie ? Le langage ordinaire permet cette paraphrase : il y a bien des façons différentes dont les choses auraient pu être, en plus de la manière dont elles sont actuellement. À première vue, cette phrase exprime une quantification existentielle. Elle énonce qu'il existe différentes entités répondant à une certaine description, à savoir « les manières dont les choses auraient pu être » [...] Prenant la paraphrase littéralement, je crois donc en l'existence d'entités qu'on pourrait appeler "les manières dont les choses auraient pu être". Je préfère les appeler des "mondes possibles". (LEWIS, 1973, p. 84)

Selon Lewis, les entités que sont les mondes possibles répondent à quatre réquisits. Le premier, c'est qu'ils sont de même nature que le monde actuel. Ils ne diffèrent pas par le genre de choses

qu'ils sont. Les mondes possibles ne sont pas des *ersatz* de mondes, des abstractions. Ils ne sont pas des entités linguistiques, ni des ensembles consistants et maximaux de phrases ou de propositions, ni même des ensembles d'états de choses :

> J'insiste énergiquement pour qu'on n'identifie pas les mondes possibles, de quelque manière que ce soit, avec de respectables entités linguistiques. Je les considère comme de respectables entités par elles-mêmes. Quand je professe un réalisme des mondes possibles, cela doit être entendu littéralement. Les mondes possibles sont ce qu'ils sont, et ils ne sont pas autre chose que cela. (LEWIS, 1973)

Les mondes possibles sont des *réalités concrètes*. Tout ce qui existe dans un monde possible existe de la même manière qu'existe ce qui réside dans notre monde.

Une conséquence de ce premier réquisit, c'est que l'actualité n'est pas la prérogative de notre monde. « Actuel » est un terme indexical, il sert à désigner le monde habité par celui qui l'utilise. Quand nous parlons du temps présent pour faire référence à l'existence d'une chose, cela ne présume pas de considérer que cette chose n'existe qu'à ce moment-là. Par comparaison, Lewis soutient que l'existence ne doit pas être réduite à l'exclusivité d'un monde pas plus qu'à l'exclusivité d'un moment : les individus habitent non seulement un lieu et un temps, mais aussi un monde.

Le second réquisit est que les mondes possibles sont extrêmement inclusifs. À l'intérieur d'un monde, l'existence de toute chose est toujours reliée causalement à l'existence de toute autre chose. Tous les individus d'un monde nouent des relations spatio-temporelles entre eux, cela, quelles que soient leurs dispersions spatiale et temporelle.

> Le monde dans lequel nous vivons est une chose très inclusive. Il n'est pas un morceau de bois ni une pierre visible qui n'en fasse partie. Il en va de même pour vous comme pour moi. De même pour la Terre, pour le système solaire, la totalité de la voie lactée, les lointaines galaxies observées au télescope, et (en supposant que de telles choses existent) tous les fragments d'espace vide entre les étoiles et les galaxies. Rien n'est si éloigné de nous qui ne puisse faire partie de notre monde. Il n'est pas une seule chose, à quelque distance qu'elle se trouve, qui ne doive y être incluse. Le monde est également inclusif dans le temps. Ni la Rome antique, ni les ptérodactyles, ni les nuages primitifs de plasma depuis longtemps disparus ne nous sont si éloignés dans le passé, ni les étoiles éteintes si éloignées dans le futur, qu'ils ne fassent partie du même monde. (LEWIS, 2007, p. 15)

Le troisième réquisit est l'isolement entre mondes. Les mondes possibles sont causalement isolés les uns des autres [29]. Si deux choses entretiennent des relations spatio-temporelles, alors elles appartiennent nécessairement au même monde.

[29] Lewis n'est d'ailleurs pas le premier auteur à admettre une multitude de mondes possibles qui soient sans relation causale avec le nôtre. On doit au *Eureka* d'Edgard Poe une description assez proche de celle de Lewis : « Qu'il me soit permis seulement de déclarer que je me sens, pour mon compte personnel, porté à imaginer (je n'ose pas me servir d'un terme plus affirmatif) qu'il existe réellement une succession illimitée d'Univers, plus ou moins semblables à celui dont nous avons connaissance, à celui-là seul dont nous aurons jamais connaissance, — du moins jusqu'au moment où notre Univers particulier rentrera dans l'Unité. Cependant, si de tels groupes de groupes existent, — et ils existent, — il est suffisamment clair que, n'ayant pas de participation dans notre origine, ils ne participent pas à nos lois. Ils ne

> Aucune distance spatiale ne les sépare d'ici, et ils ne sont ni éloignés ni proches dans passé ou le futur. Aucune distance temporelle ne les sépare du temps présent. Ils sont isolés : il n'y a absolument aucune relation spatio-temporelle entre les choses qui appartiennent à différents mondes. Ce qui arrive dans un monde n'est pas plus la cause de ce qui arrive dans un autre. (LEWIS, 2007, p. 16)

En outre, il n'y a pas non plus de recouvrements possibles entre différents mondes, ni parties communes entre ces mondes. L'isolement des mondes a pour conséquence que nous ne faisons pas les mondes nous-mêmes, puisqu'il n'est rien d'extérieur à un monde qui puisse jamais agir de manière causale sur lui. Ce que nous fabriquons, ce sont des « langages, des concepts, des descriptions et des représentations imaginaires qui s'appliquent aux mondes » (LEWIS, 2007, p. 18), mais pas des mondes, qui sont toujours primitifs et déjà donnés.

Le quatrième réquisit est le rattachement de la conception modale à l'existence des mondes, ou thèse de plénitude. Concevoir qu'il y a des façons différentes dont les choses pourraient être implique pour Lewis la réalisation de ces choses dans des mondes possibles. Ainsi, pour toute chose et pour toute manière possible d'être de cette chose, il y a une infinité de mondes possibles qui réalisent, pour cette chose, cette manière d'être. Concevoir que les poules auraient pu avoir des dents, c'est accorder l'existence de poules dentées dans différents mondes possibles. En ce sens, les possibilités constituent en quelque sorte des témoignages d'existence : chaque énoncé modal affirme en fait l'existence de multiples mondes possibles qui le réalise. Toutes les carrières alternatives d'une chose de ce monde-ci sont ainsi développées dans d'autres mondes. N'importe quelle manière contrefactuelle dont une chose aurait pu être est la manière dont cette chose est dans un monde. Il existe, *littéralement*, des mondes possibles dans lesquels Emmanuel Macron n'a pas été élu, où les poules ont des dents, où l'on peut voyager plus vite que la lumière, où l'auteur de ces lignes n'existe pas, où Batman chevauche des licornes, où personne n'existe.

En raison de son apparent caractère fantasmagorique, la théorie de Lewis est globalement rejetée, principalement en raison du premier réquisit concernant la concrétude des mondes. La plupart des philosophes qui accordent l'existence aux mondes possibles considèrent au contraire qu'ils sont des entités abstraites.

Les mondes possibles comme entités abstraites

Passons maintenant à la seconde possibilité exposée par Adams, à savoir, qu'il existe un monde possible w dans lequel la proposition ϕ_w signifiant « le monde w est actuel » est vraie dans tous les mondes possibles, et que, quel que soit $w' \neq w$, $\phi_{w'}$ n'est vraie dans aucun monde possible. Cela induit que le monde actuel l'est de manière absolue, mais aussi, de manière nécessaire. On perd alors l'intuition que les choses auraient pu être différentes.

Robert Stalnaker objecte au réalisme modal de Lewis que les différentes manières dont les choses pourraient être ne sont pas des mondes possibles mais des propriétés du monde, dont certaines ne sont pas instanciées. Si, contrairement à ce qu'indique Lewis, les mondes possibles

nous attirent pas et nous ne les attirons pas. Leur matière, leur esprit ne sont pas les nôtres, ne sont pas ce qui agit, influe dans une partie quelconque de notre Univers. Ils ne pourraient impressionner ni nos sens ni nos âmes. Entre eux et nous, les considérant tous pour un moment collectivement, il n'y a pas d'influences communes. Chacun existe, à part et indépendant, dans le sein de son Dieu propre et particulier (E. A. POE, 1932, p. 126) ».

ne sont pas également actuels, comment caractériser alors l'actualité de notre monde ? Pour certains philosophes, c'est sa complexité qui lui confère l'actualité. C'est notamment ce que soutient Charles Hartshorne quand celui-ci écrit :

> [...] that a possible world is as definite and complex as the corresponding actual one [...] reduces the distinction between possible and actual to nullity. Value is in definiteness and definiteness is the 'soul of actuality'. Were possibility equally definite it would be redundant to actualize it. (HARTSHORNE, 1965, p. 189)

Il demeure pourtant vraisemblable de considérer qu'un monde simplement possible est bien quelque chose d'infiniment complexe. De manière purement formelle, un monde simplement possible est un ensemble tel que pour tout couple de propositions antithétiques, l'une d'entre elles est réalisée et l'autre pas. On peut alors voir un monde possible comme un livre d'un monde, un ensemble de propositions consistant et complet décrivant un monde. Parler de « mondes possibles » ne relève alors que d'un abus de langage, comme ceux qui sont courants en mathématiques. Comme l'écrit Richard Jeffrey :

> Let us call a non empty set of sentences a *novel*. The novel is *consistent* if the sentences that constitute it do not logically contradict each other; and a novel is *complete* if non sentence of the langage could be added to it without rendering it inconsistent. Equivalently, a complete, consistent novel is a set of sentences which contains exactly one of each contradictory pair of sentences in the agent's language. A complete, consistent novel describes a *possible world* [...] but if talk of possible worlds seels dangerously metaphysical, we can focus attention on the novels themselves and speak of a complete, consistent novel as actually being a possible world. [...] Then possible worlds are complete consistent novels, which in turn are certain sets of sentences. (Jeffrey, 1990, page 208)

Dans « Actualism and Possible Worlds », (PLANTINGA, 1976), Alvin Plantinga décrit les mondes possibles comme des états de choses (*state of affairs*) logiquement possibles (c'est-à-dire consistants) et maximaux (c'est-à-dire complets). Tous les mondes possibles existent, mais seul l'un d'entre eux est réalisé. Dans l'explication que donne Plantinga, chaque monde a sa propre collection d'individus actualisés, mais le domaine des individus est le même pour tous les mondes. Simplement, dans certains mondes, il y a des individus qui sont *exemplifiés*, comme Napoléon dans notre monde, et d'autres individus qui existent mais ne sont pas exemplifiés, comme Sherlock Holmes dans notre monde.

Pour traiter les mondes possibles, la stratégie de Plantinga consiste à réduire toutes les propositions modales *de re* à des propositions *de dicto*. Pour ce faire, il sépare le contenu actuel du contenu non-actuel de la proposition : dire que « Emmanuel Macron n'aurait pas pu être un presse-purée » revient à affirmer que « Macron est de manière essentielle un non-presse-purée » et que la proposition exprimée par la phrase « Macron aurait pu être un presse-purée » est nécessairement fausse. Accepter la modalité *de re*, c'est admettre que certains objets possèdent leurs propriétés de manière essentielle. Cependant, s'il est essentiel à Macron qu'il ne soit pas un presse-purée, il est en revanche accidentel qu'il ait été élu. L'essentialiste affirme que les objets possèdent à la fois des propriétés accidentelles et des propriétés essentielles et que certaines propriétés sont détenues essentiellement par certains objets, mais pas par tous.

Selon Plantinga, un monde possible est un possible état de choses, mais en revanche, tous les possibles états de choses ne sont pas des mondes possibles. Il n'y a pas de correspondance un-à-un entre les mondes possibles et les états de choses. Avoir été élu est par exemple un état de chose possible pour Macron, mais cet état de choses n'est ni assez complet ni assez inclusif pour constituer à lui seul un monde possible. Un état de choses doit être *maximal* pour se voir attribuer le rôle de monde possible. On dit qu'un état de chose S inclut un état de choses S' s'il n'est pas possible que S s'actualise sans que S' s'actualise également. Autrement dit, S inclut S' si et seulement si (S $\wedge$ $\neg$S') est impossible. De plus, un état de chose S exclut un état de choses S' s'il n'est pas possible que les deux se réalisent conjointement, autrement dit si (S $\wedge$ S') est impossible. On dira alors qu'un état de choses S est complet ou maximal si pour tout état de choses S' donné, S inclut ou exclut S'. Plantinga définit alors un monde possible comme un état de choses possible maximal, un tel monde étant réalisé (monde actuel) ou non (monde simplement possible).

Il y a bien au moins un monde possible qui se produit, autrement dit, qui correspond à un état de choses maximal, à savoir le monde actuel. Ensuite, affirme Plantinga, il ne peut y avoir qu'un seul monde se produisant. Car de fait, si deux mondes w et w' se produisaient, alors il y aurait un certain état de chose S maximal qui serait à la fois inclus dans le monde w et exclu dans le monde w'. En conséquence, si w et w' étaient tous deux actualisés, l'état de choses S serait alors à la fois actualisé et non-actualisé. Il s'en suit, déduit Plantinga, qu'il ne peut y avoir deux mondes se produisant. Un seul monde possible se produit (*obtain*) : c'est le monde actuel. Il n'y a donc pas de double existant du monde actuel, le seul monde qui possède la propriété d'être instancié, de se produire, est le monde actuel.

Il s'agit cependant sans doute ici plus d'un accord de principe que d'une véritable démonstration. Car il paraît évident que l'argument présuppose d'emblée que l'actualité est la prérogative d'un monde plutôt qu'elle ne le prouve, sans quoi l'on pourrait éviter la contradiction en affirmant que l'état de chose maximal S est actualisé dans w et non-actualisé dans w'.

Comme on le voit, les interprétations possibles qu'Adams confère à l'expression « le monde w est actuel » impliquent soit d'abandonner le caractère absolu de l'actualité, soit d'abandonner son caractère contingent. Si l'on refuse ceci, on peut par exemple dénier aux mondes possibles leur caractère complet. Une autre solution, défendue par Adams, consiste à opposer l'actualisme au possibilisme. L'actualisme considère que :

> [Possibilists] begin with the whole system of possible worlds and see the actual world first of all as a possible world, a member of that system. I propose to begin, instead, with the actual world, to treat talk about the system of possible worlds as a way of talking about a proper part of the actual world, and thus to gain, so to speak, a standpoint outside the system of possible worlds from which judgments of actuality which are not world-relative may be made. *Actualism*, with respect to possible worlds, is the view that if there are any true statements in which there are said to be nonactual possible worlds, they must be reductible to statements in which the only things there are said to be are things which there are in the actual world and which are not identical with nonactual possibles. (ADAMS, 1974, p. 202)

Le possibilisme, quand il concerne les mondes possibles, est la thèse qu'il y a des mondes simplement possibles et que la notion de monde possible ne nécessite pas d'être analysée en termes de choses actuelles.

Les mondes possibles comme fiction

Le concrétisme et l'abstractionnisme sont tous les deux d'accord sur le fait de reconnaître une forme d'existence aux mondes possibles, soit comme entités concrètes, soit comme entités abstraites. Mais si l'on considère que les mondes possibles sont des entités abstraites, il vient alors qu'ils n'ont rien de commun avec le monde actuel. Les mondes possibles ne contiennent pas des individus, ils ne sont pas régis par des lois physiques, etc. Dès lors, il devient difficile d'affirmer que le monde actuel est un monde possible parmi les autres. La solution consiste alors à proposer que les mondes possibles, et en particulier le monde dit « actuel », sont tous des constructions culturelles, sans pour autant céder à un idéalisme récusant l'idée d'un monde réel. Comme l'écrit Umberto Eco,

> Selon Hintikka, les mondes possibles se divisent en deux : ceux qui sont en accord avec nos attitudes propositionnelles et ceux qui ne le sont pas. En ce sens, notre engagement envers un monde possible est, comme le dit Hintikka, un fait "idéologique". (ECO, 1979, p. 170)

De fait, admettre que ce qui est nommé « monde actuel » dans la sémantique modale relève d'une construction sociale est peut-être un sacrifice, mais il s'avère méthodologiquement nécessaire pour ne pas réduire le possible au concevable *hic et nunc*. Mais comme le souligne Umberto Eco, c'est pourtant ce que font aussi Montesquieu, Voltaire et Fontenelle, et avec eux, toute la philosophie moderne, « quand elle a essayé de comparer "nos" coutumes avec celles des peuples sauvages, en échappant aux préjugés axiologiques de l'ethnocentrisme » (ECO, 1979, p. 173). Avec la précaution méthodologique de traiter le monde de référence comme un monde possible parmi d'autres, on retourne alors à l'idée de Lewis que le qualificatif d'actuel devrait être utilisé à la manière d'un embrayeur indexical comme le sont aussi les expressions « ici » et « maintenant », mais sans recourir nonobstant à une pluralité de mondes concrets :

> Une expression comme "le monde actuel de référence" indique n'importe quel monde *d'où* un habitant jugerait et évaluerait les autres mondes (alternatifs et seulement possibles). Bref, pour le Petit Chaperon rouge qui jugerait un monde possible où les loups ne parlent pas, le monde "actuel" serait le sien, celui où les loups parlent. (ECO, 1979, p. 173)

On devrait ainsi réfuter toute forme d'existence aux mondes possibles. Comme le souligne John Nolt (NOLT, 1986), la possibilité d'un monde n'implique en rien l'existence (concrète ou abstraite) de ce monde. L'une des stratégies anti-réalistes d'explication des mondes possibles, défendue par Cresswell, consiste à tout simplement rejeter tout contenu de sens à la notion de monde possible. Cette notion est vue alors comme un pur dispositif structurel. Comme l'écrit Cresswell,

> Given a notion of the consistency of a set of sentences, and given classical propositional logic, one can easily prove that every consistent set of sentences can be embedded in a 'maximal consistent' set in which for every sentence α either α is in the set or not α is in the set. This carries no commitment to the nature of worlds, and in the standard completness proofs they are maximal consistent sets of wff of a formal language. From this it follows that provided the truth of possibly α (meaning that α is possible) requires that α be consistent then α is true in at least one possible world. (J. CRESSWELL, 2006, p. 309)

Pour parler de mondes possibles, nous avons donc besoin d'une condition de maximalité qui assure que pour chaque monde w et chaque jugement P, P ou sa négation est vrai dans w, mais pas les deux. Selon Nolt, cette précaution permet alors d'individuer les mondes possibles. Mais Nolt refuse de réduire les mondes possibles à des ensembles de propositions :

> The things that interest me have properties (being a universe in which people live and breathe, for example) that sets of statements and similar abstract entities lack. There is no identity here. (NOLT, 1986, p. 440)

Les mondes possibles sont en effet plus que cela, ils doivent présenter des propriétés qui manqueraient à de simples ensembles de propriétés. Comme l'écrit Nolt, les mondes possibles sont des univers dans lesquels des gens « vivent et respirent ». Les mondes possibles sont relatifs à des croyances, ils ne sont pas de simples constructions linguistiques passives, mais se rapportent à des intentions. Ils nous engagent dans des jeux de croyance (*make-believe*), que celles-ci soient feintes ou réelles :

> This is not to say that possible worlds are nothing more than language. They are worlds, as I have argued all along. But in saying that they are worlds, I am already using the possibilistic language. That is, I am already engaged in the sort of "make-believe" that we are trying to understand. (NOLT, 1986, p. 442)

L'approche que défend Nolt est donc une forme modale de fictionnalisme où les mondes possibles sont considérés comme des fictions utiles. Leur statut ontologique est donc purement hypothétique. Nolt offre ainsi un point de vue intermédiaire entre d'une part le concrétisme qui affirme que les mondes simplement possibles sont tout aussi concrets que le monde actuel, et d'autre part l'abstractionnisme qui ne voit dans ces mondes qu'un ensemble maximal de propositions.

De fait, si nous souhaitons parler de mondes possibles, nous devrions les considérer comme aussi concrets et réels que le nôtre. C'est justement la croyance en cet aspect concret des mondes possibles qui permet par exemple d'éclipser la réalité en faveur d'une autre lorsqu'elle s'avère déplaisante. En outre, un individu fictionnel comme Sherlock Holmes devrait être imaginé comme aussi concret qu'Emmanuel Macron, et non pas comme une entité abstraite. Nous aurons l'occasion d'insister d'ailleurs plus amplement sur ce problème dans le prochain chapitre. C'est l'intérêt de la théorie fictionnelle des mondes possibles, qui reprend en grande partie la théorie de Lewis, mais en n'accordant de réalité aux mondes possibles que sous le rapport de l'imagination. Il s'agit de faire *comme si* les mondes possibles existaient, comme s'il s'agissait du monde actuel.

Le fictionalisme modal consiste ainsi à interpréter les mondes possibles comme des fictions, à traiter n'importe quel individu d'un monde possible de la même manière qu'est traité Sherlock Holmes. Cela engage à considérer que les traductions d'expressions modales en termes de mondes possibles sont littéralement fausses. Il n'y a ni mondes simplement possibles, ni individus simplement possibles. Il n'y a pas, par exemple, de musique simplement possible que j'aurais pu enregistrer ce matin si je m'étais levé un peu plus tôt. Ce qui est vrai, c'est que d'après la fiction des mondes possibles, il y a une musique simplement possible que j'aurais pu enregistrer ce matin.

%enditemize

✤

Les mondes possibles sont donc certainement moins des doubles métaphysiques du monde réel que des rapports doxastiques et épistémiques à un monde unique et infiniment complexe, des représentations tirées d'un monde exclusif.

Cette première section aura été l'occasion de considérer la question d'une pluralité de mondes et l'hypothèse de mondes indiscernables quoique numériquement différents. Nous allons maintenant aborder deux questions relatives aux individus doubles : l'une concerne l'existence actuelle ou simplement possible d'indiscernables dans un monde donné, et l'autre, la question des mêmes, qui est en partie indépendante de la précédente, concerne la question des doubles vies d'individus transmondains. Nous allons successivement consacrer les deux sections qui suivent à ces questions.

2.2 La question des « mêmes »

2.2.1 Introduction

C'est un fait que parmi les contributions philosophiques importantes de la seconde la moitié du XXe siècle, deux d'entre elles, à savoir *On the plurality of Worlds* de David Lewis et *Naming and Necessity* de Saul Kripke, confèrent une place éminente au problème du double. Dans l'œuvre de Lewis, le double est celui du réalisme modal et des contreparties concrètes d'individus. Dans le texte de Kripke, la notion de double est sans doute plus subtile, mais tout aussi présente. Les illustrations de duplications pullulent. Il est partout question de modèles, d'originaux, de copies. Ces dupliques portent sur diverses catégories de choses : des personnes (Aristote, Nixon, Napoléon, Walter Scott) ; des individus inanimés (Vénus) ; des espèces naturelles (l'or, les tigres, l'eau) ; des personnages de fiction (Sherlock Holmes, Ulysse) et des espèces fictionnelles (les licornes). En outre, dans une note de bas de page, alors qu'il considère le problème du vague de nos identifications (comme par exemple avec le bateau de Thésée) qui induit la non-transitivité, Kripke esquisse une conception d'une philosophie du double, voire d'une logique du double à partir de la notion de réplique. Cette notion doit pourtant être distinguée de celle de Lewis en ce qu'elle ne ressortit ni à la notion de monde possible au sens du réalisme modal ni à la théorie des contreparties. Kripke propose ainsi de « dire que l'identité stricte ne s'applique qu'aux particuliers (les molécules), la relation de réplique s'appliquant aux particuliers « composés » de ces particuliers de base (les tables). On réputerait alors la relation de réplique vague et non transitive. » (KRIPKE, 1982, p. 39 note 18). Mais Kripke demeure sceptique sur le fait de parvenir à un niveau ultime de particuliers de base pour lesquels la relation d'identité ne serait jamais vague. De toute façon, dans les cas pratiques, l'identité, au sens commun, est bien suffisante.

Le problème du double doit alors être mis en rapport avec le domaine d'application de la sémantique des mondes possibles, qui a montré son efficacité à l'égard de divers types de questions, et principalement autour des problèmes liés :

1. à l'attribution de propriétés modales *de re* dans des situations contrefactuelles, qu'il s'agisse de propriétés aléthiques comme dans l'expression « Macron aurait pu perdre les élections »

ou de propriétés dispositionnelles, comme dans « ce morceau de sucre aurait fondu s'il avait été plongé dans ce café » ;

2. aux attributions *de re* d'attitudes propositionnelles focalisées sur des entités singulières comme par exemple « Il croit que Jean Racine était mathématicien » ;

3. et à la possibilité de propositions vraies concernant des fictions, comme par exemple « Sherlock Holmes est un détective ».

Dans les sections qui suivent, je m'occuperai de la question du double dans le premier type d'usages. Le deuxième et troisième types feront l'objet des chapitres suivants.

2.2.2 La question de l'identification transmondaine

Deux points de vue

Comme nous l'avons constaté dès le premier chapitre, le double relève du jugement paradoxal d'un individu x qui à la fois serait identique à y et différent de y. Dans le cas particulier d'expressions modales, de tels énoncés semblent proliférer. Lorsque l'on interprète par exemple l'énoncé « Emmanuel Macron aurait pu perdre les élections » par l'expression « Il y a un monde possible dans lequel Emmanuel Macron perd les élections en 2017 », on peut se demander comment concilier ces doubles vies d'un même individu dont les carrières diffèrent selon le monde considéré.

Car c'est un fait que nos attitudes modales font parfois référence à une même chose considérée dans des situations diverses, et parfois même, dans des situations contradictoires. En effet, quand nous référons à une chose tout en énonçant une proposition contrefactuelle, nous pensons à elle comme si elle était présente dans différents mondes à la fois. Cela revient à concevoir une multiplicité de doubles d'une chose qui serait comme cela dans tel monde, et comme ceci dans tel autre. Lorsque nous affirmons parler d'une même chose, et non de choses numériquement différentes, nous abordons alors un problème à l'intersection de la logique et de la métaphysique, appelé communément la « question de l'identité croisée » (*cross-world identity*), ou encore le « problème de l'identité transmondaine » (*transmodal identity*).

Il est notoire que de nombreux logiciens considèrent que, dans un contexte intensionnel, affirmer de telles « mêmetés » ne constitue pas un vrai problème. En revanche, les philosophes divergent sur le genre de relations qui lient un même individu entre différents mondes. Car attribuer de manière modale un prédicat à un certain individu implique que ce prédicat soit également attribuable, dans un monde possible, à un individu ayant un lien privilégié avec son antécédent. Se pose alors la question du genre de lien que les deux occurrences entretiennent entre elles. Pour la plupart des partisans réalistes des mondes possibles, il s'agit banalement de la relation d'identité. Ainsi, dire que « Emmanuel Macron aurait pu ne pas être élu », c'est supposer qu'il y a un monde possible dans lequel Emmanuel Macron lui-même, et pas un autre, n'est pas élu. Plus précisément, la proposition exprimée par l'énoncé est vraie si et seulement s'il y a un monde possible accessible depuis le monde actuel dans lequel il est le cas qu'Emmanuel Macron (lui-même) perd les élections présidentielles de 2017. C'est ainsi admettre, comme le soutient Kripke parmi d'autres, que l'identité se prolonge entre différents mondes possibles. Les conditions de vérité des énoncés modaux *de re* sont alors ramenées à des questions d'identité transmondaine.

À l'inverse, David Lewis prétend que l'identité n'a rien à voir avec l'existence d'individus représentés différemment dans des mondes possibles, ces mondes devant être considérés comme des choses *concrètes*. La relation cherchée n'est pas l'identité, mais la ressemblance.

Dans ce qui suit, je souhaite faire ressortir plus particulièrement les querelles métaphysiques et épistémiques autour des doubles, querelles qui se sont particulièrement cristallisées dans les discussions des sémantiques modales, en particulier dans les réalismes modaux, qu'ils soient d'obédience actualiste (Kripke) ou possibiliste (Lewis). Il appert que dans les conceptions modales défendues, le double est toujours celui de la théorie adverse, il est toujours rejeté sur la conception opposée. D'un côté, pour les partisans de l'identité transmondaine, le double provient de nos usages des relations de ressemblance. En revanche, les doubles échappent complètement à l'identité, ils n'entretiennent avec elle, *in fine*, qu'une sorte d'illusion stimulée par des ressemblances trop nombreuses. Partant, le double d'une chose est donc un « étrangement ressemblant », mais sa différence numérique doit être soulignée. De l'autre côté, du point de vue des partisans de la ressemblance, le double est caractérisé par des entités qui semblent *a priori* discernables, n'ayant pas les mêmes propriétés d'un monde possible à l'autre, mais qui s'avèrent, réflexion faite, identiques. C'est donc la discernabilité qui est considérée ici comme illusoire, et l'accent est mis sur l'identité numérique stricte entre individus présents dans différents mondes. D'une théorie à l'autre, on peut donc avoir affaire à des doubles menant des vies similaires ou à des *mêmes* menant des vies distinctes.

La question est de déterminer la pertinence de telles entités dès lors qu'elles sont introduites dans les théories. Doivent-elles être rejetées ? Apportent-elles au contraire un appui pour la réflexion ? Comme nous l'avons souligné dans ce qui précède, à en croire la thèse de Clément Rosset, nous devrions uniformément les rejeter, puisque le double est l'indice sûr de l'irréel. Mais voyons cela par nous-mêmes.

Avant d'aller plus loin, prenons toutefois le soin de constater que la question problématique du lien supposé exister par exemple entre le vainqueur des élections présidentielles françaises de 2017 et le même individu, mais qui les aurait perdues, ressortit en fait à deux types de questions. L'une est épistémique, l'autre, métaphysique. Pour ne pas confondre les deux questions, je nommerai la première la question de l'identification transmondaine et la seconde la question de l'individuation transmondaine. Commençons par examiner la question de l'identification transmondaine.

La question épistémique de l'identification transmondaine

Le premier type de questions concerne le problème de l'identification d'un même individu suivi à travers différents mondes. En effet, puisqu'un individu peut s'avérer, d'un monde possible à un autre, qualitativement très différent, on peut se demander comment procéder pour le reconnaître. Comment peut-on être assuré que, dans un monde possible, l'individu que nous visons est bien la même personne que Emmanuel Macron dans le monde actuel ? Comment savoir que l'individu qui perd les élections de 2017 dans cet autre monde n'est pas plutôt Nicolas Sarkozy ou Harrison Ford ? Il faut avoir un crible d'identité pour désigner des individus identiques entre mondes possibles.

La question de l'identification transmondaine peut être illustrée par un épisode de la série télévisée *Twin Peaks* de David Lynch (LYNCH & FROST, 2017). Au long de cette série, l'inspec-

teur Dale Cooper enquête sur le meurtre d'une jeune femme nommée Laura Palmer, qu'il n'a pas connu personnellement. Au gré des interrogatoires qu'il mène, Cooper parvient à collecter des informations concernant la personnalité et les relations de Palmer. Du point de vue de Cooper, Palmer est identifiée par un ensemble de propriétés, comme être une adolescente blonde aux yeux bleus, ayant tels parents et tels amis, fréquentant tels lieux, etc. La fiction mélangeant le genre policier au genre fantastique, les investigations de Cooper l'amènent progressivement à visiter divers mondes parallèles. En particulier, dans l'épisode 18 de la troisième saison, Cooper accède à un monde possible dans lequel Laura Palmer n'a probablement pas été tuée. Mais quel genre de relation l'inspecteur est-il en droit d'établir entre la Palmer qui a survécu, et celle qui est morte dans son monde ? S'agit-il de la même personne ? Comment répondre à cette question ? Quels critères Cooper a-t-il à sa disposition ? Les informations collectées dans son monde semble en effet inefficaces dans l'autre monde : car bien des différences qualitatives distinguent la présumée Palmer non-assassinée de celle qui est morte. L'une a atteint l'âge de maturité et porte un patronyme différent, et elles n'ont pas le même lieu de résidence. D'un côté, il se pourrait qu'il s'agisse de la « même » personne, parce qu'un certain nombre de faits concordent. Mais d'un autre côté, il y a tout autant de faits qui indiquent que cela pourrait être quelqu'un d'autre. Comment Cooper parviendrait-il, dans ces conditions, à reconnaître Palmer ?

Selon David Kaplan, la question de l'identification à travers les mondes possibles est « le problème central d'intérêt philosophique dans le développement de la logique intensionnelle » (KAPLAN, 1979, p. 94). Dans « Transworld Heir Line », Kaplan définit le problème de l'identification transmondaine ainsi : comment pouvons-nous dire, parmi les individus qui constituent un monde possible, celui qui est « l'héritier » (*heir*), autrement dit qui représente un individu d'un autre monde ? Pour y parvenir, nous devrions avoir une sorte de lunette donnant accès aux mondes simplement possibles, un « Jules-Verne-o-scope ».

Comme nous l'avons vu avec l'illustration précédente, on aboutit cependant à une impasse. Puisque nos lignes transmondaines ne sont basées que sur l'observation de propriétés purement qualitatives, donc sur la ressemblance, il vient alors que différents individus d'un même monde peuvent suffisamment ressembler, sous diverses descriptions à un même individu d'un autre monde. La Laura Palmer qui a survécu ressemble sans doute physiquement à celle qui est morte. La première est l'« héritière » de la seconde sous la description « ressembler physiquement ». Mais d'autres individus pourraient également s'avérer les héritiers de Palmer sous d'autres descriptions, comme par exemple, sous la description « avoir tels parents ». Comment choisir alors le bon candidat ? Le point sur lequel insiste Kaplan est que nous n'avons pas de bijection entre les domaines d'individus de deux mondes possibles, et qu'il est difficile de considérer que les lignes de mondes reliant les domaines de chaque monde sont des fonctions, puisque cette notion exige que chaque élément d'un domaine soit en relation avec un unique élément d'un autre.

Selon Kaplan, ce problème demande de prendre position parmi les trois points de vue suivants : sceptique, relativiste et métaphysique. La première alternative consiste à affirmer que l'identification n'est tout bonnement pas possible, ce qui peut être un motif de rejeter globalement la sémantique modale des mondes possibles comme moyen d'expliquer nos attitudes intensionnelles *de re*. En l'absence d'un critère d'identification, nous pourrions douter de l'efficacité de la théorie sémantique des mondes possibles pour expliquer nos attitudes modales. Le point de vue relativiste consiste à admettre une pluralité d'héritiers pour un même individu selon l'aspect qui est visé,

ce qui constitue le cœur de la théorie lewisienne des contreparties. Enfin, la dernière alternative, métaphysique, est typiquement le point de vue de certains logiciens comme Kripke. Elle consiste à adopter une théorie basée sur l'idée que l'identité n'est pas fondée sur les faits qualitatifs. L'identité paraît alors être « un fait nu (*bare identity*) » comme l'écrit Kaplan :

> In fact, the use of models as representatives of possibles worlds has become so natural to logicians that they sometimes take seriously what are really only artifacts of the model. In particular, they are led almost unconsciously to adopt a *bare particular metaphysics*. (KAPLAN, 1975, p. 97)

Réponse actualiste au problème de l'identification

Le problème de l'identification transmondaine, tel qu'il est souligné par Kaplan, a globalement été rejeté par Saul Kripke comme étant un pseudo-problème philosophique. Certes, il semble effectivement qu'il n'y ait aucun moyen de répondre de manière absolue à la question de l'identification. C'est ce que Kripke concède à Kaplan tout au long de la première conférence de *Naming and Necessity* où l'auteur rejette l'hypothèse que la compréhension de nos attitudes modales *de re* passe nécessairement par la connaissance d'un critère purement qualitatif d'individuation, autrement dit, par la connaissance de conditions nécessaires et suffisantes permettant de référer à un individu. Mais selon Kripke cela n'est pas véritablement un obstacle pour parvenir à conférer une signification aux expressions modales *de re*, qui sont toujours énoncées à coup de stipulations, et non par la rencontre plus ou moins fortuite avec des mondes possibles :

> [...] Même s'il y avait un ensemble purement qualitatif de conditions nécessaires et suffisantes pour être Nixon, la conception que je défends n'impliquerait pas qu'il faille trouver ces conditions *avant* d'être en mesure de demander si Nixon aurait pu [perdre] les élections, ni qu'il faille reformuler la question en termes de conditions. Nous pouvons considérer simplement Nixon, et demander ce qui aurait pu lui arriver *à lui* si diverses circonstances avaient été différentes. (KRIPKE, 1982, pp. 34-35)

Comme le préconisent Penelope Mackie et Mark Jago dans l'article *Transworld Identity* (MACKIE & JAGO, 2017) de la Standford Encyclopedia of Philosophy, la question du critère d'identification doit cependant être affinée par trois suppositions supplémentaires :

1. La première supposition, relative à la reconnaissance, consiste à dire que nous devons posséder un critère d'identité transmondaine pour être capable de dire d'un individu d'un autre monde dont on connaît les propriétés s'il est ou non identique à Nixon.
2. La deuxième supposition est relative à l'assurance de faire bonne référence (*security of reference*). Elle prescrit que nous devons posséder un critère d'identité transmondaine pour Nixon afin de savoir que, quand nous disons par exemple qu'il y a un monde possible dans lequel Nixon n'est pas élu, c'est bien de lui dont il s'agit, et pas d'un autre.
3. Enfin, la troisième supposition, dite d'intelligibilité, stipule que nous devons posséder un critère d'identité transmondaine afin de comprendre l'affirmation qu'il y a un monde possible dans lequel Nixon n'a pas été élu.

Kripke commente ainsi la première supposition :

> Ce qu'il y a de plus contestable, dans cette histoire d'identité, c'est la conception des mondes possibles qui la sous-tend. Un monde possible y est assimilé à un pays étranger, vis-à-vis duquel on est dans la position de l'observateur. Peut-être Nixon est-il parti pour ce pays, peut-être ne l'a-t-il pas fait, en tout cas on dispose seulement de qualités observables. Par l'observation, on peut voir toutes les qualités de quelqu'un, mais on ne peut pas voir qu'il est « Nixon ». On voit que quelqu'un a les cheveux roux [...], mais on ne voit pas qu'il est Nixon. C'est pourquoi il faut avoir un moyen de déterminer en termes de propriétés si une chose qu'on rencontre est la même qu'on a vue précédemment ; il faut avoir un moyen de dire, quand on rencontre un de ces mondes possibles, qui est Nixon dans ce monde. (KRIPKE, 1982, p. 31)

Il ne faut donc pas concevoir nos stipulations modales à la façon dont l'inspecteur Cooper, dans *Twin Peaks*, se promène d'un monde à l'autre. Nous ancrons toujours nos modalités à partir des individus du monde actuel, et la référence est donc donnée d'emblée, comme si Cooper transportait le cadavre de Palmer dans ses pérégrinations. C'est donc, soutient Kripke, une conception erronée des mondes possibles qui implique la croyance en la nécessité d'un critère d'identité pour répondre à la question de la reconnaissance. Mais nous n'avons pas du tout besoin de voir les choses de cette manière là. Nous n'avons pas besoin de parcourir les mondes possibles comme des ensembles purement qualitatifs dans lesquels nous chercherions l'élément adéquat.

Dans *Naming and Necessity*, Kripke répond également à la deuxième supposition, la question de la référence. Sa réponse consiste à dire que, dans les attitudes modales *de re*, nous ne devons pas voir les mondes possibles comme s'ils s'étalaient indifféremment devant nous, un peu comme les cartes d'un jeu que nous pourrions choisir arbitrairement. L'idée que les mondes possibles doivent être donnés de manière purement qualitative permet certes de conférer une signification aux hypothèses contrefactuelles. Mais s'il n'est pas possible de reconnaître une entité entre mondes possibles, il n'est pas utile non plus d'exiger de poser ces questions d'identification :

> « supposons qu'il y ait, dans un certain monde possible, un homme dont le chien s'appelle Checkers, qui ressemble à un personnage du répertoire de David Frye, et que cet homme, dans ce monde possible, gagne les élections ». Cet homme ressemble-t-il assez à Nixon pour qu'on puisse l'identifier à Nixon ? À quoi bon exiger cela ? Ce n'est pas la façon dont, ordinairement, nous envisageons des situations contrefactuelles. Nous disons simplement « supposons que cet homme ait été le perdant ». (KRIPKE, 1982, p. 33-34)

Macron aurait pu ne pas gagner les élections. Constatons que formuler cette hypothèse contrefactuelle n'engage rien d'autre. Macron aurait-il repris ses études de philosophie ? Aurait-il plu ce jour-là ? Aurait-il eu un nombre pair de cheveux sur la tête ? Les mondes possibles sont des mondes complets. Autrement dit, chacune de ces questions doit y trouver une réponse. Dans ce cas, devons-nous décider de toutes choses afin d'avoir une interprétation satisfaisante ? Kripke constate que ce que la théorie induit ici est préjudiciable à nos intuitions. La proposition modale exprimée par « Nixon aurait pu perdre les élections » ne concerne pas l'ensemble des choses qui auraient pu être vraies, mais le fait qu'un seul individu, Nixon, ne gagne pas les élections.

Kripke souligne ainsi que l'identification transmondaine ne doit en aucun cas être caractérisée de la manière suivante : choisir un monde possible, et dans ce monde, se demander si un individu observé est identique ou non à un individu du monde actuel. Car c'est demander bien plus que ne l'exigent nos intuitions modales. Ce qu'il revient de constater, insiste Kripke, c'est que nos propositions modales consistent à sélectionner un individu dont on conserve rigidement la référence dans un autre monde possible. C'est-à-dire qu'elles focalisent notre attention sur une entité du monde actuel, puis lui appliquent des altérations possibles. Si on admet ceci, alors l'identification transmondaine n'est *jamais* problématique, puisque la référence est toujours donnée d'emblée. La théorie sémantique des désignateurs rigides affirme ainsi qu'un nom propre comme « Macron » désigne le même individu que celui du monde actuel à travers les mondes possibles, ceci, par une sorte de visée initiale. Les noms propres logiques « attrapent » des individus actuels, et conservent la référence dans les alternatives contrefactuelles. La focale est rigidifiée à travers les transactions modales. La théorie des désignateurs rigides permet ainsi de répondre à la question épistémique de l'identification transmondaine, car, comme l'écrit Jean-François Recanati,

> quand ce nom propre figure dans un énoncé modal décrivant une situation contrefactuelle, il ne cesse pas de désigner cet individu : au moyen d'un tel énoncé, nous transportons imaginativement cet individu dans un autre monde possible. (FRANÇOIS, 1983, p. 110)

Comme on le voit, l'objection principale de Kripke à l'égard de l'identification transmondaine consiste à dire qu'on prend le problème à l'envers. Le problème de l'identification transmondaine n'est pas celui consistant à chercher, dans des mondes possibles donnés d'avance, des individus se correspondant un-à-un car nos intuitions modales ne requièrent pas de voyager d'un monde à un autre en tentant d'y apparier des entités identiques. C'est donc nos croyances à propos de la nature des mondes possibles qui engendrent ici des illusions. Or, « un monde possible n'est pas un pays lointain qu'on rencontre sur son chemin ou qu'on regarde au téléscope » (KRIPKE, 1982, p. 32). Nos attitudes modales *de re* visent un objet donné dans le monde, et constitue un monde possible autour de cet objet :

> Nous ne commençons donc pas avec les mondes possibles [...] pour, ensuite, nous enquérir des critères d'identification à travers les mondes ; au contraire, nous commençons avec les objets, que nous avons et que nous pouvons identifier dans le monde réel. Nous pouvons ensuite demander si certaines choses auraient pu être vraies de ces objets. (KRIPKE, 1982, p. 41)

Ainsi, les mondes possibles sont toujours stipulés à partir d'une visée intentionnelle sur l'individu qui est l'objet de l'attitude modale. Lorsque nous disons que « Emmanuel Macron aurait pu ne pas être élu », nous savons d'emblée qu'il s'agit de Macron, et non d'un autre. Il n'y a donc pas de problème de savoir si c'est de la bonne personne dont on parle, la référence est donnée de manière sûre par l'acte même de penser à tel individu plutôt qu'à tel autre.

La troisième supposition affirme que comprendre une expression comme « Macron aurait pu perdre les élection » demande de répondre à la question « qu'est-ce que cela impliquerait, pour un individu d'un monde possible, de perdre les élections ? » Mais à nouveau, il ne semble pas légitime de devoir répondre à cette question par un critère d'identification afin d'expliquer nos attitudes modales, pas plus qu'il n'est légitime d'ailleurs de devoir répondre à la question temporelle

« qu'est-ce que cela fait à tel enfant d'être identique à Macron ? » En outre, la compréhension de l'affirmation « Macron aurait pu perdre les élections » est sans doute rendue possible par le fait que nous savons, dans ce monde, ce que signifie « perdre les élections ». Selon Kripke, il n'y a donc guère plus de problèmes concernant l'intelligibilité des hypothèses contrefactuelles.

La thèse de la désignation rigide implique la nécessité de l'identité : pour tout énoncé d'identité ne contenant que des désignateurs rigides, si cet énoncé est vrai, alors il est nécessairement vrai. Plus formellement :

Nécessité de l'identité (NI)

Si « a » et « b » sont des désignateurs rigides alors $(a = b) \rightarrow \Box(a = b)$ [30].

Les enjeux sémantiques et métaphysiques

La réponse de Kripke à la question de l'identification transmondaine passe par une théorie sémantique, précisément, par la révision de la conception descriptiviste de la signification des noms propres et le retour à la théorie de Mill. Une conséquence de la théorie des désignateurs rigides est en effet, selon Kripke, que les noms propres sont solidaires des objets et sont indépendants de la permanence de tel ou tel prédicat.

La conception de Kripke s'oppose directement à la théorie sémantique de Russell, qui considère que les noms propres sont des abréviations de descriptions définies. Pour Russell, un nom propre, au sens grammatical du terme, comme par exemple « Walter Scott », n'est pas un nom propre au sens logique. Ce nom abrège en réalité une description définie comme « l'auteur de *Waverley* ». Chez Russell, la classe des authentiques noms propres est rabattue sur les seules formulations déictiques et indexicales. Seules les formulations faisant référence à des *sense data*, à des objets connus par relation directe (*acquantance*) peuvent être véritablement qualifiés de noms propres logiques.

De nombreux exemples sont utilisés par Kripke dans *Naming and Necessity* pour souligner l'incohérence de la théorie descriptiviste. Ceux-ci ressortissent à des altérations fictives de choses, de personnes ou de substances *actuelles*. Il s'agit en particulier de prouver qu'une description définie d'un individu dont on modifie en pensée la carrière peut venir à manquer sa cible. Considérons par exemple que le nom propre « Aristote » ait pour signification « le plus illustre des élèves de Platon et des précepteurs d'Alexandre ». Si cette information se révélait fausse, si Aristote n'avait par exemple jamais connu Alexandre, il faudrait alors admettre son inexistence. Mais cette déduction n'est évidemment pas légitime, puisque Aristote aurait pu exister même sans correspondre à la description de précepteur d'Alexandre.

Dans l'article intitulé « Proper Name », John Searle (J. SEARLE, 1958) affirme que ce que montrent véritablement de tels arguments, c'est qu'il ne faut pas restreindre la signification d'un nom à une unique description définie. Nous devrions caractériser les noms propres comme des

[30] On doit à Ruth Marcus Barcan (BARCAN, 1947) une démonstration logique de la nécessité de l'identité. Supposons que $a = b$ où a et b sont deux désignateurs rigides. Considérons la propriété $\mathrm{P}_a(z) : \Box(z = a)$, c'est-à-dire, la propriété, pour un individu quelconque z, d'être nécessairement identique à a. Par la nécessité de la réflexivité, il vient que Pa_a. D'autre part, puisque a est identique à b, par application de la règle de substituabilité des identiques, il vient que b possède toutes les propriétés de a. En particulier, on a $\mathrm{P}_b(a)$. Autrement dit, $\Box(b = a)$

abréviations pour des faisceaux de propriétés qui particularisent les individus. Néanmoins, l'objection de Kripke reste valable même dans le cas d'un faisceau individualisant de propriétés, parce que la totalité des propriétés pourrait bien s'avérer fausse.

C'est déjà ce que soulignait Wittgenstein : bien des descriptions différentes peuvent être attribuées à un personnage comme Moïse. Aussi, concevoir la signification de ce nom comme un ensemble de descriptions conduit à dissimuler la chose qui se tient derrière ces propriétés, voire à l'éclater en références multiples sans que l'on sache finalement de quel individu il s'agit au juste :

> Considérez cet exemple-ci : si l'on dit « Moïse n'a pas existé », cela peut signifier différentes choses. Cela peut signifier : les Israélites n'avaient aucun chef lorsqu'ils sortirent d'Egypte - ou bien : leur chef ne s'appelait pas Moïse - ou bien : il n'a pas pu exister d'homme qui ait pu accomplir tout ce que la bible relate de Moïse - ou bien, etc. (WITTGENSTEIN, 1961, p. 153)

Dans la théorie descriptiviste, il peut être attribué au nom « Moïse » différentes descriptions, selon qu'on considère par exemple qu'il est une figure politique ayant milité pour l'indépendance des juifs, selon qu'on l'identifie comme tel individu de telle époque ayant vécu à tel lieu, ou encore selon qu'on le considère comme l'homme responsable des dix plaies d'Egypte, etc. Mais selon la définition que nous adoptons, une proposition comme « Moïse a existé » ou comme « Moïse n'a pas existé » prend alors des significations très différentes. Car comme le constate Wittgenstein,

> si maintenant je fais un énoncé au sujet de Moïse, suis-je toujours prêt à mettre l'une quelconque de ces descriptions pour « Moïse » ? Je dirai pas exemple : « par 'Moïse' j'entends l'homme qui a accompli ce que la bible relate de Moïse, ou du moins beaucoup de ce qu'elle en dit ». Mais combien de choses ? Me suis-je décidé à reconnaître pour fausses un certain nombre de choses pour que j'abandonne comme fausse ma propre proposition ? Le nom de Moïse est-il alors pour moi d'un usage solide et déterminé sans équivoque dans tous les cas possibles ? - N'est-ce pas comme si je disposais de béquilles et que je fusse prêt à m'appuyer sur l'une quand on voudrait me soustraire l'autre, et inversement ? (WITTGENSTEIN, 1961, p. 153)

La théorie des descriptions définies aboutit à dupliquer les références. Le nom de Moïse désigne tantôt un personnage fictionnel produisant des miracles et tantôt un personnage historique qui fut une figure politique israélite. La théorie descriptiviste est donc référentiellement instable, parce que nos propos au sujet de Moïse concernent constamment des individus différents sans que l'on sache au juste lequel est visé dans les expressions qui en mentionnent le nom.

L'instabilité référentielle suscitée par la théorie descriptiviste est d'ailleurs également indiquée par Frege, qui souligne que les conceptions du nom « Aristote » peuvent notablement diverger d'un individu à un autre : celui qui sait qu'Aristote était le précepteur d'Alexandre n'interprétera pas l'énoncé « Aristote est né à Stagire » de la même façon que celui qui ne le sait pas. Dans le langage ordinaire, il y a des fluctuations du sens qui sont possibles et acceptables du moment que la référence demeure la même. Mais justement qu'est-ce qui permet à la référence de rester identique en dépit des modifications du sens ?

En outre, Kripke souligne aussi, et inversement, qu'il est possible de faire référence à des individus même lorsque ce que l'on sait d'eux s'avère faux. Comme l'a montré Keith Donnellan

dans l'article *Reference and Definite Descriptions* (DONNELLAN, 1966), il y a deux usages des descriptions définies, l'un est attributif, l'autre est référentiel. On peut désigner un convive en le décrivant comme « la personne qui boit un martini ». Or, il se pourrait que la personne en question ne boive que de l'eau. La description serait bien entendue fausse, et donc ne correspondrait pas à un usage attributif, mais cela ne gênerait le moins du monde de faire référence à la bonne personne. Et cela pourrait s'avérer le cas même s'il y avait, à côté de cette personne, une autre buvant réellement du martini. Plus généralement, on peut d'ailleurs reconnaître une personne tout en étant incapable d'en donner une description, tout comme l'on peut aussi désigner une substance naturelle sans en connaître les propriétés chimiques.

Les incohérences et la fragilité référentielle que Kripke relève dans la théorie des descriptions définies le conduisent à considérer, après John Stuart Mill, que les noms propres réfèrent de manière directe. Les noms propres comme « Macron » réfèrent directement aux entités qu'ils nomment. Le fonctionnement des noms propres n'est donc pas satisfactionnel - il ne consiste pas à remplir un certain nombre d'attributs caractéristiques, mais relationnel. La désignation d'un nom propre est ancrée sur un acte de baptême, puis véhiculée par des chaînes de communication de ce nom. Bien évidemment, la fixation initiale de la référence peut s'établir par la médiation d'une description définie. On peut par exemple baptiser la planète Vénus par la description « la première étoile visible à l'aube ». On peut également fixer la référence par un acte de désignation directe, de manière indexicale. On pourrait désigner « cette étoile-là » en pointant la planète du doigt. L'usage d'un nom relève d'une transmission sociale liée à un apprentissage. Le nom propre véhicule directement la référence en ramenant à l'acte inchoatif de désignation de l'objet visé.

Le fait que les noms propres soient des désignateurs rigides privilégiés, qu'ils le soient, disons, de manière « intrinsèque », n'exclut pas pour autant que des descriptions définies puissent être employées également de manière rigide. Lorsque nous disons que « Le président de la république aurait pu être estropié d'une main », nous pouvons vouloir dire deux choses : soit qu'Emmanuel Macron lui-même aurait pu être estropié, soit qu'un tout autre président aurait été élu, qui lui, aurait été estropié. Il y a donc deux emplois possibles de la description définie « le président de la république », dans l'un, c'est au même individu qu'est attribuée la propriété « être estropié », alors que dans l'autre, il s'agit d'un individu différent. Formellement, on a les deux interprétations suivantes :

1. $\exists x\Big(\mathrm{P}x \wedge \forall y(\mathrm{P}y \rightarrow y = x) \wedge \Diamond \mathrm{E}x\Big)$
2. $\Diamond\exists x\Big(\mathrm{P}x \wedge \forall y(\mathrm{P}y \rightarrow y = x) \wedge \mathrm{E}x\Big)$

Dans le premier cas, la description n'entre pas dans le champ de l'opérateur modal, on dit que le contexte d'usage est large. Dans le second cas, la description entre dans le champ de l'opérateur modal, on dit que le contexte est étroit. Aussi, en contexte large, lorsque je dis que l'actuel président aurait pu perdre les élections, je fixe d'emblée l'attention sur un objet du monde actuel, Emmanuel Macron, auquel je soumets hypothétiquement des attributs différents. On peut donc désigner un objet de manière rigide sans devoir stipuler pour ce faire que la désignation est indépendante des propriétés de cet objet. La visée désignatrice peut en effet passer dans une description définie exemplifiée dans le monde actuel, mais que l'individu visé peut éventuellement perdre dans

un autre. Ainsi, de ce point de vue, la différence entre un nom propre et une description définie n'est pas que le premier, contrairement au second, ne connoterait pas, mais que le premier ne peut jamais relever que d'une portée large, contrairement au second qui peut relever ou d'un contexte large, ou d'un contexte étroit.

On pourrait sans doute voir la théorie de la référence directe à la manière de Strawson (STRAWSON, 1959b) comme une sorte de pragmatique : une désignation nécessite la prise en considération d'une *intention* de référence, elle ne pas être réduite à la seule adéquation entre une description particularisante et un état de choses censé y référer. Mais derrière la dimension sémantique de la réponse de Kripke au problème de l'identification transmondaine apparaît aussi un second enjeu, qui est, lui, métaphysique. Le fait est que Kripke change la discussion de terrain, en défendant des thèses sémantiques, ne doit pas occulter cet intérêt métaphysique. Il s'agit en effet pour lui de contester la théorie de l'individuation par faisceau de propriétés (*Bundle theory*). Cette théorie, initiée par David Hume, propose une alternative à la théorie du *substratum* postulant que les individus instancient des propriétés, des universaux, qui sont supportées par une substance fine, elle-même sans qualité, diversement désignée sous les appellations de « réceptacle » (PLATON, 1950c, p. 48c53c) ou de « particulier nu » (*bare particular*). Ainsi, le partisan de la théorie du *substratum* concède qu'il y a à la fois des particuliers et des universaux portés par ces particuliers. Parce qu'il est exclu, par définition, qu'ils soient eux-mêmes des universaux, les particuliers sont donc seulement des particuliers nus distincts des propriétés qu'ils soutiennent. Puisqu'ils ne peuvent posséder, par eux-mêmes, aucune qualité, ils relèvent ainsi d'une catégorie plutôt étrange de choses, que John Locke qualifie de « je ne sais quoi » (LOCKE, 1998a, II, xxiii, §2). À l'inverse, la théorie du faisceau propose une explication purement qualitative de l'individuation. Elle est donc plus simple en ce qu'elle élimine la stipulation de porteurs nus sous les propriétés. Elle affirme que les particuliers surviennent sur les universaux. Un particulier n'est donc rien de plus que la somme des propriétés qu'il exemplifie.

Dans la première conférence de *Naming and Necessity*, Saul Kripke rejette la théorie du faisceau, mais affirme dans le même temps ne pas admettre pour autant la théorie des particuliers nus [31],

> On dit souvent que, si une situation contrefactuelle est décrite comme une situation où *Nixon* joue un rôle, et si l'on ne suppose pas que cette description est réductible à une description purement qualitative, alors on suppose l'existence de mystérieuses entités, les « particuliers nus », qui sont dénués de propriétés et servent de substrat aux qualités. Ce n'est pas le cas [...]. Si une qualité est un objet abstrait, un faisceau de qualités est un objet encore plus abstrait, abstrait à un niveau encore plus élevé, et non un particulier. (KRIPKE, 1982, p. 40)

Ainsi, pour Kripke, une table n'est ni un faisceau de propriétés purement qualitatives, ni un particulier nu, « elle n'est pas une chose sans propriété qui serait derrière les propriétés » (KRIPKE, 1982, p. 40). Il s'agit, précise Kripke, d'un faux dilemme, la question n'ayant, selon Kripke, aucune importance. En effet, nous pouvons parler, entre mondes possibles, d'une et même seule table sans même connaître ses critères d'identité. Et de fait, la réponse de Kripke à l'objection d'échec d'identification transmondaine passe par deux thèses :

[31] Dans la traduction qui suit, j'ai modifié l'appellation « simples particuliers » qui me semble trompeuse et peu adaptée par celle de « particuliers nus ».

1. Les questions d'identité d'individus représentés peuvent être fixées par des actes de stipulation.

2. La stipulation fournit un moyen épistémique d'accéder aux mondes possibles.

John Divers constate que la première thèse paraît acceptable, mais il conteste la validité de la seconde, qu'il considère au mieux comme induisant de sérieuses confusions (DIVERS, 2002, p. 273). En effet, nous pourrions stipuler que Macron aurait pu être un œuf-poché. Mais cela ne doit pas nous faire croire que nous accédons pour autant à un monde *possible* :

> It is one question how we know which objects are the objects of our de re modal thought and talk, and perhaps there stipulation has a legitimate role. It is another question altogether how we know what is modally true of those objects, and there stipulation has no legitimate role to play. (DIVERS, 2002, p. 273)

En outre, on voit que la réponse de Kripke se fait au profit d'une sorte de confusion entre la question épistémique et la question métaphysique. À cet égard, Tero Tulenhemo écrit que :

> The problem with the idea of rigid designation is not the suggestion that proper names refer directly without ascribing properties to their bearers. The problem is that the very notion of referring to the same object in distinct worlds presupposes the notion of cross-world identity. Yet, Kripke wishes to use the notion of rigid designation to clarify the meaning of statements involving individuals that appear in several possible worlds. That is, his strategy is question-begging: he attempts to employ a notion that presupposes the possibility of speaking of cross-world sameness to clarify the meaning of statements about cross-world sameness. The notion of reference is linguistic and world-relative—it is about interpreting non-logical symbols world by world. The notion of cross-world identity, again, does not presuppose recourse to language, and it is about world lines determining which objects in which domains realize the same individual. (TULENHEIMO, 2017, p. 45)

La notion de référence est relative aux mondes, elle n'est pas absolue comme l'est l'identité. La façon dont Kripke rend compte de la signification des propositions modales *de re* consiste néanmoins à procéder à une sorte de fusion entre identité et référence, en conférant à cette dernière un caractère absolu. Cela nous amène au deuxième problème : celui de l'identité croisée, ou de identité transmondaine.

2.2.3 La question de l'individuation intermondaine

Exister dans plusieurs mondes

Dans la conception des mondes possibles de Leibniz, le César d'un monde possible qui n'aurait pas franchi le Rubicon ne serait pas le même individu que le César qui franchi le fleuve dans notre monde. Non pas au sens que les deux individus ne se ressembleraient pas, mais au sens qu'il n'y aurait pas d'identité *numérique* entre eux. De fait, dans la conception leibnizienne des mondes possibles, un individu du monde actuel ne peut jamais être numériquement identique à un individu

d'un monde simplement possible. Contre Leibniz, Saul Kripke défend au contraire l'idée que les individus des mondes possibles peuvent bien être les mêmes que ceux du monde actuel. Autrement dit, lorsque nous parlons d'un César qui n'aurait pas franchit le Rubicon, nous voulons parler de César lui-même, comme s'il avait pu vivre *une double vie* dans un monde différent.

Si les remarques de Kripke paraissent répondre au problème de l'identification, il n'en reste pas moins qu'elles ne visent pas à infirmer l'existence de critères d'identité transmondains. En effet, il reste possible de prétendre qu'il doit bien exister un critère d'identité transmondain pour Macron, au sens d'un ensemble de propriétés telles que leur exemplification par tout individu et tout monde possible soit nécessaire et suffisante pour que cet individu soit Macron.

Mais dès lors, on glisse vers un autre type de question, qui n'est donc plus de nature épistémique, mais métaphysique. Il s'agit de déterminer un critère d'identité des individus dans différents mondes possibles. Il ne s'agit alors plus d'une question d'authentification mais d'individuation entre mondes possibles. Constatons que le problème d'identité transmondaine est certainement ambigu du fait de ce qu'on entend exactement par « identité ». Pour David Lewis, il n'y a jamais eu de problème d'identité, transmondaine ou autre, parce que l'identité est « excessivement simple et peu problématique » (LEWIS, 2007, p. 295), n'étant rien d'autre que la relation que toute chose entretient avec elle-même et exclusivement avec elle-même. S'il appert que nous formulons quantité de problèmes sous la forme de questions d'identité, nous pouvons aussi, remarque Lewis, les formuler autrement, sans même mentionner la relation d'identité.

De fait, il y a différentes façons de parler des doubles vies d'Emmanuel Macron. David Lewis souligne qu'en un sens, il n'y a pas de mystère à cette existence transmondaine :

> Arrive-t-il jamais que quelque chose existe selon deux mondes différents [...]. La réponse affirmative ne devrait pas prêter à discussion : cela arrive très souvent. (…) Cette question-là, (…), ne mérite pas d'être appelée un problème. (LEWIS, 2007, p. 297)

On peut alors tenter de caractériser l'existence transmondaine par l'idée de parties intermondaines communes à un même individu de la manière suivante :

Il y a un individu x tel qu'il y a deux mondes distincts M_1 et M_2 où x est à la fois un constituant de M_1 et de M_2.

Si cette condition semble bien acceptée par certains partisans actualistes, elle s'avère rejetée par Lewis (LEWIS, 2007) en raison que sa théorie interdit aux individus de faire *entièrement* partie de plus d'un monde. Précisons cependant que cela n'exclut pas, chez Lewis, l'hypothèse d'individus ayant des parties transmondaines. Lewis défend une méréologie non restreinte, que ce soit spatialement, temporellement, mais aussi de monde à monde :

> Le partage de parties est, somme toute, un lieu commun. En effet, n'importe quelle partie de n'importe quel monde fait partie d'innombrables sommes méréologiques qui s'étendent au-delà de ce monde. (LEWIS, 2007, p. 304)

Le problème ne tient donc pas au fait qu'un individu présent dans un monde soit un constituant d'un individu plus gros ayant d'autres parties dans d'autres mondes. En revanche, écrit Lewis,

« si une chose [...] semble problématique - inconsistante, ce qui n'est pas un mince mot - c'est la manière dont la partie commune de deux mondes est supposée avoir des propriétés différentes dans l'un et l'autre monde » (LEWIS, 2007, p. 304). Ce que rejette Lewis dans le recouvrement des mondes est donc l'idée qu'une partie constitutive d'un individu puisse exister dans deux mondes à la fois.

Le paradoxe de Chisholm

Roderick Chisholm a indiqué deux paradoxes dirigés contre la formulation la plus faible de l'identité transmondaines IM_1. Le premier paradoxe ressortit de la discernabilité. Soit un individu x d'un monde M_1. Soit un monde M_2 distinct de M_1. Or si l'un se distingue de l'autre, c'est que, dans ce monde, il y a quelque chose de différent du monde M_1 [32]. Dans ce cas, aucun individu de M_2 ne peut se dire indiscernable à x : car il s'en distingue dans le fait que dans son monde existe une différence qui le rend discernable de x dans M_1. Dans ce cas, par la contraposée d'indiscernabilité des identiques, on peut en déduire que x n'est pas identique à y. C'en est alors fini de l'identité transmondaine.

Le même genre d'argument est d'ailleurs mobilisé par Lewis sous le nom des propriétés intrinsèques accidentelles. Le problème de variation modale des qualités intrinsèques d'un individu sert de critique principale, chez Lewis, à la conception kripkéenne des mondes possibles :

> [Hubert Humphrey] aurait pu avoir six doigts à la main gauche. Un autre monde le représente ainsi. Supposons maintenant que la représentation *de re* fonctionne selon l'identité trans-mondaine. Ainsi, Humphrey, qui fait partie de ce monde et possède ici cinq doigts à la main gauche fait également partie d'un autre monde, et là, il a six doigts à la main gauche. [...] Lui-même, en personne - le seul et le même, entièrement identique à lui-même -, a cinq doigts à la main gauche et cependant il en a six. Comment cela se peut-il ? (LEWIS, 2007, p. 305)

La caractérisation méréologique de Lewis le mène à comparer l'identité transmondaine appliquée à l'interprétation de l'énoncé modal concernant Humphrey au cas de frères siamois, Ted et Ned, qui partageraient la même main gauche, mais dont on dirait de Ted qu'il possède cinq doigts à se main gauche alors qu'on dirait de celle de Ned qu'elle en compte six. Aussi Lewis en conclut-il que l'emploi de l'identité transmondaine conduit à « un double discours et à une contradiction ».

Il n'est sans doute pas très difficile de répondre à cette première objection. En effet, on peut rétorquer que le principe d'indiscernabilité des identiques ne s'applique pas entre des mondes possibles différents. C'est donc une réponse analogue, dans le cas modal, à celle utilisée dans le cas temporel. Le principe doit, nous l'avons vu, être relativisé temporellement, mais il doit également l'être de manière modale. Ce que les partisans de l'identité transmondaine disent, c'est que Humphrey a cinq-doigt-dans-ce-monde, et qu'il a six-dans-l'autre. Mais Lewis écarte cette possibilité par le fait que la théorie kripkéenne considère justement que c'est Humphrey tout entier qui est inclus dans chacun des deux mondes, et non certaines parties propres.

[32] Cette affirmation, que nous acceptons ici au bénéfice de l'argumentation, est néanmoins controversée. Nous y reviendrons plus loin, dans la section traitant de l'haeccéitisme.

> Si Humphrey, en effet - lui-même en personne, la totalité qu'il constitue - doit mener une double vie, en tant qu'il fait partie de deux mondes différents, ses propriétés intrinsèques ne peuvent en aucune manière intelligible différer d'un monde à l'autre. (LEWIS, 2007, p. 307)

Quoi qu'il en soit, passons maintenant au second paradoxe de Chislholm, devenu le carrefour des polémiques concernant l'identité dans les contextes modaux. À l'instar des paradoxes sorites, son mécanisme repose sur la succession de petits changements menant *in fine* à une contradiction avec les hypothèses de départ. Précisément, il s'agit d'échafauder une série de mondes possibles dans lesquels deux individus permutent successivement leurs propriétés qualitatives et extrinsèques respectives.

Chisholm a donné une première formulation de son paradoxe dans l'article « The Logic of Knowing » :

> Consider that possible world which differs from this one only in the fact that in that one General De Gaulle has the same weight that I have in this one and in that one I have the same weight that General De Gaulle has in this one; then consider a world in which we similarly exchange other properties; then worlds in which we change more and more properties; then the world in which I have all the properties that General De Gaulle has in this one except that of weight, and in which he has all the properties that I have in this world except that of weight; and finally that possible world (if there is one) in which I have all the properties that General De Gaulle has in this world and in which he has all the properties that I have in this one. Am I identical with General De Gaulle in some of these possible worlds and not in others ? If so, in which ones ? And how are we to decide ? (CHISHOLM, 1963, p. 174)

Nous avons deux mondes absolument indiscernables dans lesquels deux individus ont permuté l'ensemble de leur propriétés, deux mondes où ils ont échangé leurs rôles respectifs, sans voir modifiée la moindre propriété qualitative. Que devons-nous penser d'une telle possibilité ?

De fait, il y a ici deux problèmes. On retrouve bien sûr le problème épistémique de l'identification transmondaine que nous avons décrit précédemment, car nous pouvons déduire du paradoxe qu'il est tout à fait impossible de reconnaître un individu en examinant ses seules propriétés qualitatives. Comme nous avons déjà traité cette question ci-dessus, c'est essentiellement un second problème qui nous intéressera plus particulièrement ici. Ce problème est, lui, métaphysique : la conclusion du paradoxe semble impliquer l'admission d'individus numériquement différents mais qualitativement indiscernables. Mais également, à un autre niveau, la possibilité de mondes numériquement différents mais qualitativement indiscernables, ce que l'on qualifie généralement de mondes « haeccéitiques ».

Chisholm a développé les conséquences de ce paradoxe dans un second article intitulé « Identity Through Possible Worlds : Some Questions » (CHISHOLM, 1967). Le paradoxe met en jeu les contraintes modales de l'essentialisme. Rappelons alors les principes de base de cette théorie.

On dit qu'un particulier x *instancie un ensemble* $\mathcal{E}$ de propriétés $\{p_1, \ldots, p_n\}$ si x instancie chacune des propriétés p_i de l'ensemble $\mathcal{E}$. Dans la suite, afin de simplifier le propos, on considérera que toutes les propriétés p_i sont mutuellement modalement-indépendantes (par exemple, on ne considère pas à la fois des propriétés comme « être un mammifère » et « être un animal »).

On définit alors l'*essence sortale* (ou spécifique, etc.) d'un particulier x comme un ensemble $\mathcal{E}$ tel que :

1. le particulier x ne pourrait pas ne pas instancier l'ensemble $\mathcal{E}$;
2. tout particulier qui instancie $\mathcal{E}$ est de la même sorte que x.

On appelle *essentialisme* la thèse que :

1. les propriétés d'un individu sont des universaux (engagement ontologique envers les propriétés) ;
2. tout particulier x instancie l'ensemble complet de ses propriétés $\mathcal{P}$ de deux façons différentes :
   (a) par un ensemble $\mathcal{E}$ de propriétés essentielles ;
   (b) par un ensemble $\mathcal{A} = \mathcal{P} - \mathcal{E}$ de propriétés accidentelles (*id est*, non-essentielles) ;

Notons que du fait de l'introduction de modalités *de re* dans les définitions, l'essentialisme présuppose une approche intensionnaliste.

Passons maintenant au paradoxe, dont je change légèrement la présentation afin de la simplifier [33]. Considérons deux individus du monde actuel M_1, disons, Socrate et Platon. Si l'on accepte l'essentialisme sans céder à un hyper-essentialisme, et si l'on rejete l'idée d'essences individuelles non qualitatives, on doit alors reconnaître que ces individus possèdent des qualités essentielles qu'ils partagent, comme par exemple la propriété d'être un homme ou celle d'être issu d'un ovule fécondé, et qu'ils possèdent également, chacun, des qualités accidentelles qui les départagent. Par souci de simplification, nous illustrerons le paradoxe en limitant à deux les propriétés essentielles et à trois les propriétés accidentelles. Nous noterons alors $\{E_1, E_2\}$ les propriétés essentielles que partagent à la fois Socrate et Platon, nous noterons $\{S_1,S_2,S_3\}$ l'ensemble des qualités accidentelles de Socrate et $\{P_1,P_2,P_3\}$ l'ensemble des qualités accidentelles de Platon.

Puisque ces qualités sont contingentes, Socrate et Platon auraient pu voir échangée l'une de leurs qualités respectives accidentelles. Admettons qu'il s'agisse de S_1 et P_1. En traduisant en terme de sémantique modale, nous dirions qu'il y a un monde possible M_2 dans lequel Socrate posséderait les qualités $\{E_1, E_2, P_1, S_2, S_3\}$ et où Platon posséderait les qualités $\{E_1, E_2, S_1, P_2, P_3\}$. Arrêtons-nous ici un moment : d'un monde l'autre, s'agit-il encore du même Socrate ? Pouvons-nous affirmer d'un tel individu qu'il est encore identique au Socrate du monde M_1 ? Une première objection pourrait en effet être que le changement implique la rupture d'identité, autrement dit qu'il s'agit là d'une entorse au principe d'indiscernabilité des identiques. Mais nous avons déjà répondu plus haut à cette objection par la relativisation du principe aux mondes. Par précaution, nous devrions peut-être quand même utiliser un nom différent pour chaque entité mondaine. Nous pourrions par exemple noter l'entité qui correspond à Socrate dans le monde M_i par « Socrate-dans-M_i ».

[33] Je reprends ici la manière dont Filipe Drapeau Contim l'expose dans (DRAPEAU VIEIRA CONTIM, 2010).

Après cette précaution, continuons le déroulement du paradoxe. Il est le cas que Socrate-dans-M_2 aurait pu avoir la propriété P_2 que possède Platon-dans-M_2, et il est respectivement le cas que Platon-dans-M_2 aurait pu avoir la propriété S_2 que possède Socrate-dans-M_2. En termes de mondes possibles, il a un monde M_3 où les propriétés de Socrate-dans-M_3 sont $\{E_1, E_2, P_1, P_2, S_3\}$ et où celles de Platon-dans-M_3 sont $\{E_1, E_2, S_1, S_2, P_3\}$. Enfin, en réitérant une dernière fois les mêmes substitutions qualitatives entre les individus de M_3, on aboutit finalement à un monde possible M_4 dans lequel Socrate-dans-M_4 possède les qualités $\{E_1, E_2, P_1, P_2, P_3\}$ et où Platon-dans-M_4 possède les qualités $(E_1, E_2, S_1, S_2, S_3\}$.

Que signifie le paradoxe de Chisholm ?

Il n'est pas évident de comprendre exactement quelles sont les répercussions du paradoxe de Chisholm, et les auteurs en tirent d'ailleurs diverses conclusions plus ou moins éloignées. Il s'agit, pour ce qui nous concerne en particulier, de déterminer quelles conséquences du paradoxe ressortissent au problème du double. À quoi aboutit au juste ce paradoxe ? Je considère que trois conséquences importantes peuvent en être tirées.

La première conséquence du paradoxe est d'aboutir à une contradiction. C'est seulement en ce sens que le problème de Chisholm mérite d'être appelé « paradoxe ». Le problème est le suivant. Premièrement, l'application croisée du principe d'identité des indiscernables entre les deux mondes M_1 et M_4 implique que Platon-dans-M_1 est identique à Socrate-dans-M_4, puisque le premier partage exactement les mêmes propriétés, tant intrinsèques qu'extrinsèques, que le second. Mais par ailleurs, nous pouvons aussi déduire, par transitivité de l'identité transmondaine, que Socrate-dans-M_1 est identique à Socrate-dans-M_4. De ces deux résultats, on peut en déduire finalement, à nouveau par la transitivité de l'identité, que Platon-dans-M_1 est identique à Socrate-dans-M_1, ce qui est bien sûr contradictoire avec notre hypothèse de départ. Voilà donc une conclusion absurde.

Nous pouvons cependant éviter la conclusion absurde en choisissant de restreindre mondainement le PIdI, voire en le rejetant complètement. Or, si l'on conteste que le PIdI puisse s'appliquer entre des mondes distincts, Socrate-dans-M_1 pourrait alors être indiscernable de Platon-dans-M_4 sans lui être identique. Entre les mondes M_1 et M_4, Socrate et Platon ont en effet permuté la totalité de leurs propriétés qualitatives, sans pour autant céder à l'identité numérique. Le paradoxe aboutirait alors à la conclusion suivante : Socrate-dans-M_1 est numériquement distinct de Platon-dans-M_4 et Socrate-dans-M_1 est numériquement identique à Socrate-dans-M_4 par application de la transitivité de l'identité transmondaine. On peut alors parler de deux carrières alternatives de Socrate, l'une où Socrate joue le rôle de Socrate-dans-ce-monde, et l'autre où Socrate joue le rôle de Platon-dans-ce-monde. Et inversement pour Platon. La seconde conséquence que nous pouvons tirer du paradoxe de Chisholm est donc la possible existence d'indiscernables. En outre, puisque Socrate et Platon ont été choisis de manière arbitraire, il faudrait reconnaître que n'importe quel individu d'un monde donné aurait une infinité de répliques indiscernables dans d'autres mondes. Le couple Socrate-dans-M_1 et Platon-dans-M_4 déroge ainsi directement au principe modéré d'identité des indiscernables transmondains : ils sont indiscernables tant par les propriétés intrinsèques qu'extrinsèques, mais sont néanmoins numériquement distincts. Le paradoxe est certes évité puisque la conclusion absurde n'intervient plus. Mais nous aboutissons néanmoins à des

entités plutôt étranges. Soulignons encore une fois qu'un partisan du PIdI pourrait admettre des indiscernables isolés dans des mondes possibles différents. Dans ce cas, éviter la contradiction ne nécessiterait pas de rejeter le PIdI, mais seulement d'en restreindre l'application à un seul et même monde.

La troisième conséquence du paradoxe de Chisholm ne concerne pas les individus comme parties de monde, mais les mondes comme individus maximaux. Car l'indiscernabilité qui est la seconde conséquence établie ci-dessus semble également attribuable aux mondes possibles dans lesquels les carrières de Socrate et Platon alternent. De fait, les mondes M_1 et M_4 sont parfaitement indiscernables. Car si l'on suppose qu'entre les transformations successives de Socrate et Platon, les mondes ne sont pas affectés par d'autres changements que ceux qui visent spécifiquement ces deux individus, il vient alors que le monde M_4 ne diffère en rien, qualitativement parlant, du monde M_1. Il semble que nous ayons, entre ces deux mondes, deux individus qui ont permuté leur rôle *ceteris paribus*. Autrement dit, nous avons deux mondes complètement semblables si ce n'est que les rôles tenus par deux de leurs habitants sont échangés. Cela semble impliquer que l'identité de ces deux protagonistes ne soit pas complètement déterminée par les qualités qu'ils instancient. Autrement dit, nous avons là un problème d'identité nue entre deux mondes, qu'on appelle, pour cette raison, des mondes haeccéitiques. Le terme désigne ainsi deux mondes absolument identiques quant aux qualités extrinsèques et intrinsèques, mais néanmoins numériquement différents.

2.2.4 Réponses au paradoxe de Chisholm

Résumons-nous : le paradoxe de Chisholm conduit à une contradiction, à des indiscernables transmondains, à des mondes haeccéitiques qui ne diffèrent que par des identités nues. Devant les conclusions du paradoxe, on peut adopter trois attitudes :

1. accepter la première conclusion et donc aussi les trois autres (par effet d'explosion logique) ;
2. rejeter la première conclusion et aucune, certaines ou toutes les trois autres en rejetant les prémisses du paradoxe ;
3. rejeter la première conclusion et aucune, certaines ou toutes les trois autres en rejetant le raisonnement impliqué dans le paradoxe.

Je vais examiner tour à tour ces trois alternatives.

Accepter le paradoxe

Pour certains philosophes, la conclusion du paradoxe doit bien être acceptée. Ce qu'elle indique, c'est le caractère absurde d'une corrélation entre l'essentialisme et les modalités.

Les propriétés essentielles sont en effet rattachées, depuis Aristote, à l'idée de nécessité : dire qu'un individu possède une propriété essentielle, c'est dire que cet individu n'aurait pas pu ne pas avoir la dite propriété, autrement dit, qu'il la possède de manière nécessaire. Cela conduit à dire alors que toute propriété essentielle est aussi une propriété nécessaire, et réciproquement.

La condition semble certes suffisante : une propriété qui est essentielle à Socrate est aussi une propriété qui lui est nécessaire, c'est-à-dire une propriété qu'il a dans tous les mondes possibles où il existe. En revanche, il n'est pas sûr qu'elle soit aussi nécessaire.

En effet, Kit Fine (FINE, 1994) a montré que cette manière de concevoir l'essence d'une chose implique des conséquences assez étonnantes. Considérons l'exemple suivant : la propriété d'être tel qu'il y a une infinité de nombre premiers. C'est une propriété nécessaire à Socrate. Mais il parait étrange d'affirmer qu'il s'agit aussi d'une propriété essentielle. Considérons encore un autre exemple. Admettre une équivalence entre l'essence et la nécessité implique par exemple de ranger la propriété d'être le seul membre du singleton $\{Socrate\}$ parmi les propriétés essentielles de Socrate pour la raison qu'il est nécessaire que Socrate soit le seul membre de l'ensemble $\{Socrate\}$. Pourtant, il semble que cette propriété ne soit pas essentielle à Socrate. Ainsi, le fait d'avoir une propriété nécessaire n'implique pas que cette propriété soit essentielle à l'objet qui l'instancie.

Ces exemples indiquent que la caractérisation modale des propriétés essentielles n'est sans doute pas pertinente. Pour la remplacer, Fine propose de revenir à une caractérisation par les définitions « réelles », qu'on trouve également chez Aristote : les propriétés essentielles d'un objet sont celles qui font partie de la définition de l'objet. Évidement, toute la difficulté est alors repoussée sur la question de savoir ce qu'est au juste la définition d'un objet, car ce sont plutôt les mots ou les concepts qui ont des définitions, mais pas les choses elles-mêmes. L'idée de définitions « réelles », opposées aux définitions « nominales » semble ainsi assez illusoire.

On peut alors opter pour le rejet complet de l'essentialisme, comme y convie Quine. Penser qu'une chose aurait pu être différente de ce qu'elle est actuellement en usant de mondes possibles implique que la responsabilité de nos attributions modales soit bien trop délaissée à notre imagination. Comme l'écrit Quine, « you can change anything to anything by easy stages through some connecting series of possible worlds » (W. v. O. QUINE, 1976b, p. 861). Cela conduit Quine à rejeter la pertinence philosophique de la notion de monde possible, en ce que de tels mondes diffèrent du monde actuel par la permissivité qui semble sans limite dans les altérations qu'ils autorisent en pensées. Dans le monde réel, l'identification des individus à travers leur persistance temporelle nous est imposée par les faits. Le monde actuel nous *impose* les phases momentanées successives qui mènent aux transformations factuelles, cela, « pour le meilleur et pour le pire » (W. v. O. QUINE, 1976b, p. 861). Aussi, selon Quine, ce que montre le paradoxe, c'est, de manière globale, que l'essentialisme est absurde. Néanmoins, il est souvent objecté à la conclusion de Quine que des paradoxes similaires apparaissent dans des cas purement temporels ne recourant pas à la modalité, et qu'il n'est donc pas aussi simple de se débarrasser ainsi de l'essentialisme (ENGEL, 1985, ch. V).

La théorie des contreparties, développée par David Lewis, est en partie affranchie des conclusions du paradoxe de Chisholm parce qu'elle refuse l'identité transmondaine au profit d'une relation non transitive. En ce sens, elle ne réfute pas le paradoxe, mais au contraire en accepte toutes les conclusions.

Selon Lewis, et comme Quine, le paradoxe est un argument montrant que l'usage de l'identité entre des mondes différents mène bien à des contradictions. Il reste alors à voir comment la théorie des contreparties permet d'expliquer les assertions qui ressortissent de modalités *de re* comme « Emmanuel Macron aurait pu ne pas être élu » tout en évitant de souscrire à l'identité

d'un même individu entre deux mondes possibles. La stratégie consiste à aménager une place à l'idée d'existence transmondaine nonobstant le rejet de l'identité entre mondes possibles, par la médiation d'une relation non transitive de similarité ou de représentation. En un sens, Lewis ne rejette pas l'idée que Macron existe dans plusieurs mondes. Mais selon lui, nous devrions examiner les problèmes d'identité transmondaine plus finement en distinguant les questions suivantes :

1. Y a-t-il un recouvrement entre mondes distincts ? Y a-t-il des parties communes entre mondes distincts ?
2. Y a-t-il des individus recouvrant des mondes différents ?
3. Pour un individu actuel, exister dans un monde possible exige-t-il d'admettre l'identité numérique entre deux occurrences d'un même individu dans deux mondes ?

Selon Lewis, la première question pourrait être formulée de manière à savoir s'il y a identité entre des choses qui sont partiellement dans des mondes différents. En formulant ainsi la question, nous n'avons plus vraiment affaire à un problème d'individuation. En fait, il s'agit plutôt d'une question méréologique posée en terme d'identité. Admettons qu'un monde ait comme partie Emmanuel Macron lui-même. C'est d'ailleurs de cette façon que notre monde représente que Macron existe. Affirmer que Macron réside dans plusieurs mondes, cela revient à dire que la même partie nommée Macron est recouverte par des mondes différents. En outre, cette partie commune doit aussi posséder des propriétés différentes selon les mondes : dans l'un, Emmanuel Macron gagne les élections, dans l'autre il les perd. Il y a un élément x tel qu'il y a un monde M distinct d'un monde M' et tels que x est un constituant de M et x est un constituant de M'. Cette hypothèse, nous l'avons déjà vu, est complètement rejetée par Lewis. Les mondes sont sans partie commune, ils n'admettent pas de recouvrement.

La réponse de Lewis à la deuxième question est affirmative, et ne nécessite, cette fois encore, aucun recours à la notion d'identité : un individu peut être, de façon partielle, présent dans différents mondes, à l'image d'une route possédant divers tronçons dans différents états. La libre quantification méréologique pratiquée par Lewis implique inévitablement que les sommes méréologiques ne soient en aucun cas restreintes. La condition que les parties d'un individu quelconque soient toutes contenues dans le même monde est donc rejetée par Lewis. Il y a bien des individus transmondains. Lewis définit ainsi la notion de *stade* :

> si un individu possible X fait partie de l'individu Y transmondain, et si X n'est pas une partie propre d'un quelconque autre individu possible faisant partie de Y, alors X est appelé un *stade* de Y. (LEWIS, 2007, p. 327)

Pour Lewis, un tel individu transmondain, quoiqu'il existerait bien, ne représenterait pas une possibilité *de re*. Cela signifie que si toute possibilité représente un monde, en revanche, toute somme méréologique ne représente pas nécessairement une possibilité *de re*. Il y a bien des individus transmondains, mais ces individus ne sont pas des possibilités au sujet d'individus actuels :

> je conteste non pas l'existence d'individus transmondains - si je quantifie sans restriction - mais l'attention qu'on leur porte. Je ne refuse pas l'existence des individus transmondains,

> bien que je dise qu'en un certain sens qu'il ne leur est en aucune manière possible d'exister. Il est possible pour une chose d'exister si et seulement s'il est possible à la totalité de cette chose d'exister. C'est-à-dire si et seulement s'il y a un monde dans lequel la totalité de cette chose existe. (…) Les parties de monde sont des individus possibles ; les individus transmondains sont des individus impossibles. (LEWIS, 2007, p. 322)

La réponse de Lewis à la troisième question est négative. La principale critique de Lewis à l'égard de la théorie sémantique de Kripke est en effet l'objection des « doubles vies », autrement dit la stipulation d'individus transmondains menant des carrières alternatives ou, pour ainsi dire, des *mêmes* contrefactuels. C'est cette conséquence néfaste de la théorie kripkéenne de l'identité à travers les mondes possibles que récuse Lewis. De fait, si un grand nombre de problèmes philosophiques sont formulés avec la relation d'identité, il ne s'avère pas moins que ces problèmes ne nécessitent pas cette formulation. La façon dont nous énonçons nos problèmes ne doit pas abuser notre compréhension de la question. La relation d'identité ne doit pas restreindre les réponses possibles aux problèmes de la persistance temporelle ou transmondaine d'un individu : « La compagnie Trans-world Airlines est intercontinentale et non interplanétaire » (LEWIS, 2007, p. 295).

La relation entre l'individu qui représente une possibilité *de re* et son antécédent n'est pas l'identité numérique en ce sens qu'il n'y a pas d'individus transmondains *possibles*. Lewis nous doit cependant une explication pour la relation qui se tient entre Emmanuel Macron ici et l'individu de l'autre-monde qui perd les élections. Lewis substitue à l'identité numérique une relation de représentation qu'il appelle « relation de contrepartie » (*counterpart relation*). Emmanuel Macron a été élu dans ce monde, et certes, il aurait pu ne pas l'être. Cette possibilité est bien une possibilité *de re*, au sujet de Macron et de nul autre que lui. Un monde possible qui ressortit à l'hypothèse contrefactuelle que Macron n'ait pas été élu ne contient pas Macron, sans quoi nous aurions affaire à un *double-même* kripkéen. Il contient par contre un représentant de Macron, entendu au sens de l'individu qui ressemblerait le plus, dans un tel monde, à l'individu du monde actuel. Ce que contient un tel monde n'est pas Macron lui-même, mais un représentant de Macron.

Mais qu'est-ce au juste qu'une représentation *de re* ? Constatons d'abord qu'une représentation peut, par nature, se faire *in absentia*, c'est-à-dire, en l'absence de la chose représentée. Tout comme la statue de cire du musée Grévin représente, dans notre monde, Emmanuel Macron, elle pourrait aussi bien représenter Emmanuel Macron dans un monde où il n'existe pas. Une chose n'a pas besoin d'exister pour être représentée. D'ailleurs, au musée situé 221B Baker Street à Londres, il y a bien une statue de cire qui représente Sherlock Holmes.

Le personnage de cire qui représente Macron existe en tant qu'*artefact* de cire et en plus, il représente quelque chose d'autre, un autre individu qui existe lui aussi, ailleurs. Dans un monde possible, un individu peut représenter *de re* un autre individu sans que des parties soient recouvertes entre les deux mondes. Un monde possible représente le fait qu'Emmanuel Macron perde les élections, de la même manière que la statue de cire représente que Macron existe dans notre monde, qu'il bouge, qu'il fait tel geste, qu'il remporte les élections. Et c'est aussi comme cela que Macron perd les élections dans un autre monde. Dans un monde où Macron perd les élections, ce n'est donc non pas Macron lui-même qui les perd mais une contrepartie (*counterfeit*). Aussi, les modalités *de re* ne sont pas d'authentiques propriétés qui seraient instanciées par des individus, mais elles servent de descriptions pour des représentations de ces choses dans des configurations

différentes du monde.

La relation de contrepartie sert donc, chez Lewis de substitut pour la relation d'identité transmondaine :

> Là où certains diraient que vous existez dans différents mondes, dans lesquels vous avez des propriétés quelque peu différentes et dans lesquels il vous arrive des choses quelque peu différentes, je préfère dire que vous êtes dans le monde actuel et dans aucun autre, mais que vous avez des contreparties dans différents autres mondes. Vos contreparties vous ressemblent étroitement, à la fois dans le contenu et le contexte, par de nombreux aspects. Elles vous ressemblent plus que toute autre chose ne vous ressemble dans leur monde. Mais elles ne sont pas vraiment vous. Car chacune d'elles est dans son propre monde, et que vous seul êtes dans le monde actuel. En parlant de manière relâchée, on pourrait en effet dire que vos contreparties sont vous dans d'autres mondes, que vos contreparties et vous-mêmes êtes les mêmes, mais cette mêmeté n'est pas plus une identité littérale que la mêmeté qui se tient entre vous aujourd'hui et demain. Il serait préférable de dire que vos contreparties sont les personnes que vous auriez été, si le monde avait été autrement. (LEWIS, 1968, p. 27-28)

La force de la théorie des contreparties réside ainsi dans le fait qu'elle ne pose pas le problème en termes d'identité transmondaine, qui s'exposent, quant à eux, aux critiques de Quine liées aux contextes intensionnels, aux entités sans conditions d'individuation :

> Quine a protesté que les possibles inactualisés sont des éléments anarchiques, presque incorrigiblement impliqués dans les mystères de l'individuation. Cela devrait bien être vrai de tous les possibles inactualisés qui mènent des doubles vies, flânant entre les portes de deux mondes à la fois. Les possibles inactualisés auxquels je crois, confinés chacun dans son propre monde et reliés uniquement par des liens de ressemblance à leurs contreparties quelque part ailleurs, ne posent aucun problème d'individuation. Au moins, ils posent seulement de tels problèmes d'individuation autant qu'ils s'érigent à l'intérieur d'un simple monde [34] (LEWIS, 1973, p. 87)

Kripke objecte qu'une telle façon de concevoir l'existence dans un autre monde est contre-intuitive : si ce ne sont pas les objets visés eux-mêmes qui habitent les mondes possibles, les perspectives qu'ils ouvrent n'offrent aucun intérêt existentiel pour les individus eux-mêmes. Lorsqu'on admet la vérité de la proposition « Emmanuel Macron aurait pu ne pas être élu », c'est bien de Macron lui-même dont on veut parler, et non d'un autre qui le représenterait ailleurs : quel intérêt cela aurait-il pour lui ?

> [Dans la conception des contreparties] si nous disons « Humphrey aurait pu gagner les élections (si seulement il avait fait ceci ou cela) », nous ne parlons pas de quelque chose qui aurait pu arriver à *Humphrey*, mais à quelqu'un d'autre, une « contrepartie ». Il est toutefois probable qu'aux yeux de Humphrey le fait que quelqu'un d'*autre* aurait été victorieux dans un autre monde possible ne présente strictement aucun intérêt, quelle que soit la ressemblance entre cet autre et lui. (KRIPKE, 1965, p. 33 note 13)

[34] Ma traduction.

Lewis répond à cette objection qu'elle semble sans fondement chez quelqu'un qui rappelle « qu'un monde possible n'est pas un pays lointain qu'on rencontre sur son chemin ou qu'on regarde au téléscope » (KRIPKE, 1965, p. 32). Comment Kripke peut-il prétendre, objecte Lewis, que c'est Humphrey lui-même qui remporte les élections, s'il n'admet pas la théorie réaliste des mondes possibles ? Seul un réaliste authentique pourrait dire que c'est d'Humphrey dont il s'agit. Mais que serait, dans la théorie actualiste, Humphrey dans monde abstrait ? Lewis admet que la théorie des contreparties ne parvient sans doute pas à rendre complètement compte de l'intuition que c'est Humphrey lui-même dont il s'agit, et pas seulement d'un représentant très similaire. Mais il insiste sur le fait que, de toute façon, seul un authentique réaliste modal pourrait soutenir cette double existence :

> S'il y avait des réalistes modaux authentiques croyant au recouvrement des mondes, ils seraient en effet en droit de prétendre que Humphrey lui-même gagne les élections dans un autre monde. Ce serait effectivement un bon point pour eux, mais non pour toute autre conception que la leur, et en particulier, pour aucune forme d'ersatzisme que ce soit. (LEWIS, 2007, p. 300)

Notons aussi au passage que d'un point de vue sémantique, la théorie des contreparties nécessite de réviser la théorie de la désignation rigide des noms propres. Car selon Lewis, ce n'est plus directement à Emmanuel Macron que l'on réfère lorsqu'on utilise son nom dans une proposition contrefactuelle, mais à un représentant, autrement dit, à quelque chose se substituant à sa référence en l'absence de celle-ci. Avec la théorie des contreparties, la théorie de la référence directe est donc admise de manière intramondaine, mais réfutée de manière transmondaine au profit d'un descriptivisme. En revanche, et contrairement à ce qui est souvent affirmé dans la littérature consacrée à la question, Lewis ne rejette pas complètement la thèse de désignateurs rigides. Comme l'indique Filipe Drapeau Contim (DRAPEAU VIEIRA CONTIM, 2012), la théorie des contreparties a pour conséquence sémantique une thèse de quasi-rigidité.

La théorie des contreparties est une théorie extensionnelle donnant une explication à nos intuitions concernant les modalités *de re*. Il y a principalement deux versions de cette théorie dans les travaux de Lewis. Celle d'origine est publiée en 1968 dans l'article *Counterpart theory and Quantified Modal Logic* (LEWIS, 1968). Dans l'article *Counterparts of Persons and Their Bodies* de 1971 (LEWIS, 1971), cette version a été remaniée en profondeur. Sa version mûre est exposée en détail dans *De la pluralité des mondes possibles* (LEWIS, 2007). Plutôt que d'utiliser des opérateurs de modalité habituels, le carré et le losange, Lewis formalise les énoncés modaux avec des quantificateurs. L'idée est d'expliquer la modalité, donc d'analyser les opérateurs modaux. Pour cela, Lewis introduit les mondes possibles comme des individus. À cette fin, il étend le domaine de quantification en utilisant sa théorie réaliste des mondes possibles. Par exemple, l'énoncé « il n'y a pas de bière » doit être entendu comme une abréviation pour l'énoncé « il n'y a pas de bière *ici* » mais certainement pas comme « il n'y a de bière nulle part ». La même remarque devrait, selon Lewis, valoir pour tout individu. Dire « il n'y a pas d'anneau qui rende invisible » signifie plus exactement « dans notre monde, il n'y a pas d'anneau qui rende invisible », mais pas du tout qu'un tel anneau n'existe absolument pas ailleurs.

La logique des contreparties utilise quatre prédicats primitifs (LEWIS, 1968) :

- Wx pour « x est un monde possible » ;
- Ixy pour « x est dans le monde possible y » ;
- Ax pour « x est actuel » ;
- Cxy pour « x est une contrepartie de y ».

Pour donner un exemple, une proposition comme $\Box$Pa s'exprime dans le langage des contreparties de la manière suivante : toutes les contreparties de a sont P soit encore :

$$\forall y \forall x \left((\mathrm{W}y \wedge \mathrm{I}xy \wedge \mathrm{C}xa) \rightarrow \mathrm{P}x \right)$$

Le tour pratique de cette interprétation formelle est que la nature habituellement intensionnelle des propositions modales y trouve une traduction extensionnelle. De surcroît, la formalisation élégante de Lewis permet l'expression de modalités qui ne peuvent l'être dans le formalisme de la logique modale. Par exemple, les principes 1 et 2, affirmant que rien n'existe ailleurs que dans un monde et que rien n'est dans deux mondes à la fois, permettent de donner une signification à l'intuition modale que Emmanuel Macron aurait pu ne pas exister, en raison du fait que Emmanuel Macron existe dans le monde actuel, et donc dans aucun autre monde.

Dans sa première version, la relation de contrepartie peut être caractérisée ainsi :

Relation de contrepartie Cxy := L'individu x dans le monde M est la contrepartie de l'individu y dans le monde M' si et seulement si x ressemble en d'importants points à y et que rien ne ressemble plus, dans M, à y que x.

Les contreparties d'un individu x sont les individus de M' qui ressemblent à x plus qu'à toute autre chose du monde M. Avec la théorie des contreparties, on peut évaluer la vérité de la proposition exprimée par la phrase « Emmanuel Macron aurait pu perdre les élections » ainsi : la proposition est vraie si et seulement si il y a un monde possible M' et un individu x de ce monde possible tels que :

- rien, dans le monde M', n'est plus ressemblant sous l'aspect F, à Emmanuel Macron dans M que x dans M' ;
- x perd les élections dans M'.

Dégageons plus précisément les caractéristiques de la relation de contrepartie. Comme nous l'avons dit, la relation de contrepartie est une relation de ressemblance comparative :

> The counterpart relation is a relation of similarity. So it is problematic in the way all relations of similarity are: it is the resultant of similarities and dissimilarities in a multitude of respects, weighted by the importances of the various respects and by the degrees of the similarities. (LEWIS, 1968, p. 116)

Contrairement à l'identité, elle n'est donc pas absolue, mais toujours relative à un contexte de comparaison. Un individu peut toujours ressembler à un autre sous un nombre important d'aspects mais aussi s'en distinguer sous un nombre tout aussi important d'aspects. Emmanuel Macron ressemble certes à Nicolas Sarkozy sous la prédication « élu aux élections nationales », mais s'en écarte, entre autres aspects, par la physionomie et la personnalité. Remarquons que la relation de contrepartie n'est pas une relation transitive. Un individu a peut être extrêmement ressemblant à un individu b et b être extrêmement ressemblant à un troisième individu c sans que a ne ressemble *suffisamment* à c.

L'une des critiques récurrentes à l'égard de la théorie des contreparties, c'est qu'elle n'explique pas vraiment nos intuitions modales. Par exemple, il est souvent affirmé que la proposition exprimée par « Macron aurait pu être très différent (de ce qu'il est actuellement) » ne trouve pas d'interprétation en termes de contreparties. En effet, la relation de contreparties étant une relation de similarité, l'interprétation de la proposition précédente dans la théorie de Lewis implique l'existence, dans un monde, d'une contrepartie très dissemblable de son antécédent. Mais cela paraît contradictoire avec la définition de contrepartie. La réponse de Lewis est cependant qu'il y a toujours une infinité d'aspects sous lesquels la contrepartie ressemblerait à son antécédent, et aussi une infinité d'autres aspects pour lesquels elle s'en distinguerait. La relation étant circonstancielle, on peut traduire l'intuition modale précédente en disant par exemple que, du point de vue de la personnalité, on a bien affaire à une contrepartie, bien que celle-ci pourrait s'avérer par exemple très différente du point de vue de la physionomie. Il n'y a donc pas véritablement de contradiction ici, en raison du caractère toujours relatif de la relation de contrepartie utilisée.

Dans la première version de la théorie des contreparties, Lewis affirme qu'il existe des mondes tels que certains individus d'un monde donné ne sont les contreparties d'aucun individu d'un autre monde.

> It would not have been plausible to postulate that, for any two worlds, anything in one was a counterpart of something in the other. Suppose there is something x in world w - say, Batman - which does not much resemble anything actual. If so, x is not a counterpart of anything in the actual world. (LEWIS, 1968, p. 116)

De même, Lewis soutient qu'il existe des mondes pour lesquels certains individus d'un monde donné n'ont aucune contrepartie dans un autre monde :

> It would not have been plausible to postulate that, for any two worlds, anything in one had some counterpart in the other. Suppose whatever thing x in world w it is that resembles you more closely than anything else in w is nevertheless quite unlike you; nothing in w resembles you at all closely. If so, you have no counterpart in w. (LEWIS, 1968, p. 116)

La deuxième version de la théorie des contreparties, exposée à partir de (LEWIS, 1971) et développée dans (LEWIS, 2007), se démarque de la première par la prise en compte de la pluralité éventuelle des contreparties d'un même particulier à l'intérieur d'un même monde. Les propriétés possibles ou nécessaires d'un individu ne dépendent pas de sa nature mais du contexte de communication dans lequel ces propriétés lui sont attribuées. Ainsi, le cinquième principe, qui prescrivait qu'il ne peut y avoir deux contreparties d'un individu dans un même monde, n'est plus retenu.

Dans la seconde version, les liens de proximité deviennent plus lâches par l'abandon de la clause d'unicité. La relation de contreparties qui, apparaît initialement comme une relation de similarité, est remplacée dans la seconde version de la théorie par une relation plus générale de représentation. Puisque la relation de contrepartie est ancrée sur une relativité contextuelle et sur un degré de similarité, un individu peut avoir plusieurs contreparties différentes dans un même monde possible. La relation de contrepartie (en tant que relation de représentation) n'est donc pas une relation univoque (un-à-un) comme l'est l'identité. Il s'agit d'une relation plurivoque (plusieurs-plusieurs). Comme le souligne Filipe Drapeau Contim (DRAPEAU VIEIRA CONTIM, 2012), il est à remarquer que la relation de contrepartie peut non seulement varier de manière contextuelle (plusieurs contreparties différentes pour un même individu), mais également de manière synchronique (plusieurs relations en même temps pour un même individu), ce qui permet de déjouer les paradoxes de constitution matérielle en distinguant différents modes synchronique d'appréhension d'un individu plutôt que différents individus coïncidents. Une autre conséquence importante de cette révision est que toute chose d'un monde est toujours représentée dans tous les autres mondes possibles, parce que toute chose peut toujours être représentée d'une certaine manière par une autre.

Cependant, la relation de contrepartie n'est pas une relation de représentation classique. En effet, Emmanuel Macron dans le monde actuel représente lui aussi une infinité d'individus différents dans des mondes possibles. On a des représentations qui sont des choses bien réelles pour des individus simplement possibles, ce qui constitue un évident atout pour une théorie fictionaliste des intentions, comme nous le verrons dans le prochain chapitre.

En outre, il y a ici quelque chose qui me semble important, quoique cela demeure tacite dans les écrits de Lewis, à savoir, la question de déterminer si la relation de contrepartie est ou non réflexive (principe 6). En tant que relation de représentation, il faudrait répondre non puisqu'une relation de représentation n'est jamais réflexive. Mais en tant que relation de similarité, la réponse serait oui, en raison que, dans son monde, rien n'est plus ressemblant à un individu que lui-même, et ceci, sous tout rapport. Remarquons en effet le lien entre la réflexivité et le principe 5 dans la version première de la théorie. Si l'on accepte que tout individu d'un monde est sa propre contrepartie, cela implique que rien n'est la contrepartie d'autre chose dans son propre monde. En effet, admettons qu'il existe un individu y qui soit la contrepartie d'un individu x numériquement différent de y dans le même monde. Dans ce cas, rien ne ressemblerait plus à x que y dans le monde M. Mais cela est contradictoire avec le fait que rien ne ressemble plus, dans M, à x que x lui-même. La réflexivité (principe 6) implique donc le principe 5. Autrement dit, rejeter le principe 5 nécessite aussi de rejeter la réflexivité, c'est-à-dire, le principe 6. On devrait donc affirmer que la relation de contrepartie (comme représentation) n'est généralement pas réflexive.

Considérons maintenant la question de la symétrie. Lewis précise que si « la relation de contreparties n'est pas toujours symétrique, elle l'est certainement souvent » (LEWIS, 2007, p. 299). La démonstration qu'en donne Lewis dans (LEWIS, 1968, p. 116) est la suivante. Supposons qu'un individu x d'un monde M soit une sorte de mélange de deux frères y et y' habitant un monde M'. Admettons en particulier que x ressemble étroitement à chacun des deux individus y et y', plus que toute autre chose de M. Dans ce cas, x est la contrepartie de y (et de y'). Mais il se pourrait aussi que, dans M', y' ressemble plus à x que ne le fait y. Dans ce cas, y ne serait pas la contrepartie de x. On a donc mis en exergue un cas pour lequel on a Cxy mais pas Cyx.

Quant à la question de la transitivité, on peut remarquer que la relation de contreparties ne l'est

généralement pas. En effet, pour trois individus x, y et z logeant respectivement dans trois mondes M, M' et M", le fait d'avoir Cxy et Cyz n'implique pas celui d'avoir Cxz. En effet, dans M", il pourrait y avoir un individu t qui ressemblerait plus à x qu'à z. C'est justement cette caractéristique qui lui permet d'échapper à l'identification entre Socrate-dans-M_1 et Platon-dans-M_4.

La théorie des contreparties accepte les conséquences du paradoxe de Chisholm en s'affranchissant de la menace d'absurdité par le fait d'attribuer la contradiction à un emploi transmondain de l'identité là où la théorie des contreparties ancre les individus dans des mondes distincts. Si la théorie des contreparties semble consistante du point de vue du paradoxe de Chisholm, elle n'évite pas pour autant la possibilité d'indiscernables transmondains (Lewis est neutre sur la question de la validité du PIdI), ni celle des mondes possibles indiscernables (Lewis les rejette), ni celles des métamorphoses (Lewis semble parfois les accepter, parfois les rejeter).

Rejeter les prémisses du paradoxe

Le mécanisme de résolution du paradoxe de Chisholm par la théorie des contreparties est d'ordre logique, en ce qu'il s'agit de remplacer l'identité par une relation non transitive. D'autres stratégies sont possibles. Ceux qui rejettent la conclusion paradoxale de Chisholm peuvent aussi rejeter ses prémisses, qu'on rappelle ici :

1. Principe de tolérance constitutive : un individu aurait pu être constitué d'un amas matériel légèrement différent.
2. Principe de restriction constitutive : un individu n'aurait pas pu être constitué d'un amas matériel entièrement différent.

Le principe de tolérance constitutive qui sert aux différents étapes dans l'argument du paradoxe de Chisholm peut être formulé de la manière suivante. Pour Sxy signifiant « x est suffisamment coextensif à y » et pour Mxy signifiant « x est constitué par y », on a :

$$\text{(TCM)} \quad \forall s \forall b_1 \forall b_2 \; (Sb_1b_2 \rightarrow \Box(Msb_1 \rightarrow \Diamond Msb_2))$$

Pour éviter le paradoxe, on peut diversement rejeter cette implication. On peut par exemple baisser le curseur du principe de tolérance constitutive à zéro, et par conséquent, positionner celui du principe de restriction constitutive au maximum. C'est sans doute une autre manière de refuser l'essentialisme, mais cette fois, par une sorte d'hypertrophie. De ce point de vue, si ce n'est actuellement le cas, Socrate n'aurait pas pu avoir un seul de ses cheveux de la même longueur que l'un de ceux de Platon. On bloque ainsi le mécanisme du paradoxe en affirmant de toute propriété qu'elle est essentielle à l'identité de l'individu qui l'instancie. Dès lors, Socrate n'aurait plus une seule propriété accidentelle. Le premier pas dans le possible est d'emblée refusé puisque l'identité d'une chose serait complètement dépendante de l'identité stricte de la totalité de ses parties matérielles. Le moindre changement impliquerait alors une rupture définitive d'identité. Un vélo auquel on changerait un pneu ne serait plus le même vélo, un chat qui perdrait un poil ne serait plus le même chat et un galet qui perdrait une molécule par l'action de la marée ne serait plus le même galet. Un telle révision de notre conception de l'identité impliquerait d'innombrables difficultés pour parler des objets communs. A cet égard, elle ne semble donc pas bienvenue.

Alvin Plantinga souligne que ni l'hyper-essentialisme ni la théorie des contreparties ne sont aptes à expliquer nos attitudes modales *de re*. La critique de Plantinga consiste à dire que, dans la théorie des contreparties, quand on vise un objet concret en lui attribuant une propriété possible, on accède à des mondes dans lesquels on ne peut pas importer la propriété essentielle d'être identique à l'objet visé. L'objection de Plantinga tend ainsi à rapprocher la théorie de Lewis d'une forme d'hyper-essentialisme [35]. Les choses n'auraient pu être différentes pour aucun particulier concret de ce monde. Lewis ne considère pourtant pas que sa théorie relève de l'hyper-essentialisme, pour la raison qu'elle donne bien une explication à nos identifications modales. En effet, la théorie des contreparties explicite bien la manière dont une chose est représentée comme étant, dans un autre monde, différente de ce qu'elle est actuellement. Du fait de l'introduction de modalités *de re* dans les définitions, l'essentialisme présuppose une approche intensionnaliste. Est-ce compatible avec l'approche extensionnaliste de Lewis ? Le point général est que pour Lewis, les choses auraient certes pu tourner autrement, mais qu'en revanche, pour un individu donné dans un monde, ses parties constitutives n'auraient pas pu être différentes. Cette rigidité des propriétés instanciées (ou des parties incluses) par les individus pourrait effectivement relever d'une forme d'essentialisme maximal : toute chose a toutes ses propriétés de manière essentielle.

La question est finalement de savoir quel genre d'*essences* Lewis admet dans sa métaphysique. Sont-elles des essences nominales ou réelles ? Il semble que Lewis propose une vue intermédiaire, peut-être conceptualiste, entre le rejet de l'essentialisme (autrement dit la thèse que les essences sont exclusivement nominales) et la thèse qu'il y a des essences réelles sous-jacentes à nos classifications [36].

Mais on peut trouver dans la littérature consacrée à l'essentialisme des philosophes qui reconnaissent défendre un essentialisme maximal. Nous pourrions bien sûr nommer Leibniz. Plus près de nous, Roderick Chisholm affirme également souscrire à une telle perspective, quoiqu'il utilise plus exactement les termes d'essentialisme méréologique pour désigner sa propre conception. Chisholm défend ainsi l'idée que : « If x is an S-part [37] of y, then y is such that in every possible world in which y exists x is an S-part of y » (CHISHOLM, 1973b, p. 587). Néanmoins, il concède que le principe est faux dans le cas des objets ordinaires. On peut bien changer le pneu d'un vélo sans que sa persistance temporelle ne cesse. Chisholm distingue alors deux types d'usage de l'identité : l'un est strictement philosophique, et le second n'est qu'une forme lâche véhiculée par l'emploi du langage commun au sujet des objets ordinaires. Le principe exposé ci-dessus n'est en fait valable que pour les individus *primitifs*. Chisholm distingue ainsi d'un côté les objets ordinaires comme les vélos qui n'ont de conditions d'identité qu'en un sens vague et relâché, et d'un autre côté les objets primitifs qui demeurent strictement identiques si et seulement s'ils conservent strictement les mêmes parties. Or, comme nous l'avons déjà fait remarquer plus haut, une telle manière d'envisager les choses est certainement trop contre-intuitive pour que nous y cédions. Si nous acceptons

[35] C'est aussi ce que constate Nathan Salmon qui écrit que : « Counterpart theory is, at bottom, just a inflexible brand of essentialism... The counterpart theorist can mouth the words 'ship a might have been made from slightly different matter', but any such pronouncement by the counterpart theorist in modal operator discourse is a verbal camouflage that merely postpones the inevitable. What matters is what the counterpart theorist means by these words, and more importantly, what is not meant by these words » (SALMON, 1981, p. 236).

[36] La question de l'essentialisme défendu par Lewis semble toutefois assez complexe et dépasse le cadre de cet mémoire. À ce sujet, voir (DRAPEAU VIEIRA CONTIM, 2012).

[37] Le terme "S-part" désigne la notion proprement philosophique de partie pour la distinguer de la notion plus lâche du langage commun.

la signification d'une expression comme « cette tarte aurait pu être mieux cuite », cela présume que nous rejetions d'emblée la thèse de l'hyper-essentialisme, autant en un sens lâche qu'en un sens plus philosophique, s'il est possible de conférer une signification véritable à une telle distinction.

Réviser l'argument du paradoxe

Il reste maintenant à examiner les stratégies qui excluent la première conséquence du paradoxe en sollicitant le rejet de l'argument, autrement dit, en proposant une révision logique. De ce point de vue, la principale idée consiste à déterminer une manière d'éviter l'application de la transitivité de l'identité *entre* les mondes possibles.

Une première manière de procéder consiste à réviser la logique de l'identité en adhérant à la théorie de l'identité relative. Contrairement à la théorie des contreparties qui propose une nouvelle relation non-transitive en lieu et place de l'identité, il s'agit ici de considérer que la relation qui lie les individus transmondains est bien l'identité, et que celle-ci est relative. De fait, au cours des échanges de propriétés entre Socrate et Platon, les propriétés qui font l'enjeu des transactions deviennent successivement différentes. Ainsi, si l'identité n'est pas absolue mais toujours relative à une catégorie sortale, alors les relations contextuelles d'identité, qui sont médiatisées par comparaison à ces relations, sont elles-mêmes distinctes. Il pourra donc arriver que Socrate-dans-M_1 soit F-identique à Socrate-dans-M_2 et que Socrate-dans-M_2 soit G-identique à Socrate-dans-M_3. Cela n'impliquera pas l'identité entre Socrate-dans-M_1 et Socrate-dans-M_3 puisque la transitivité ne s'applique pas, les relations d'identité utilisées étant différentes. Le relativiste devrait ainsi dire que le paradoxe part d'un monde pour aboutir au même monde de départ. Il y a donc identité numérique entre les deux mondes observés, ils ne sont pas indiscernables mais distincts.

Une autre stratégie déjouant la transitivité consiste à laisser intacte la logique classique de l'identité, mais à bloquer certaines relations d'accessibilité dans la structure des mondes possibles. Hugh Chandler (CHANDLER, 1976) a développé cette stratégie. Afin d'exposer sa propre solution, Chandler décrit une version alternative du paradoxe de Chisholm. Il ne nécessite plus de considérer deux individus du monde actuel, mais un seul, dont on modifie les parties constitutives un peu à la manière du paradoxe du bateau de Thésée. Il rapproche ainsi le paradoxe de Chisholm des paradoxes de constitution matérielle.

Supposons que ce qu'on désigne habituellement par « bicyclette » soit la somme méréologique de trois parties constitutives A, B et C (voir figure 2.3). Admettons en outre le principe de tolérance constitutive suivant : les bicyclettes persistent au changement d'une de leurs parties. Admettons également le principe de restriction constitutive suivant : si l'on change au moins deux des parties d'une bicyclette, l'identité est rompue. Considérons maintenant que dans le monde actuel M_1 une bicyclette est constituée des parties (A_1, B_1, C_1) comme représenté sur la figure 2.3. Puisqu'une partie de la bicyclette aurait pu être échangée, cela implique qu'on a accès, depuis M_1 à un monde possible dans lequel la bicyclette est constituée des parties M_2 (A_2, B_1, C_1). Dès lors, depuis ce monde M_2, cette bicyclette aurait pu elle-même être constituée des parties (A_2, B_2, C_1). Il y a donc un monde M_3 qui représente cette possibilité. Dans ce cas, la transitivité de l'identité nous mène à l'identité des bicyclette entre les mondes M_1 et M_3, ce qui est contradictoire par notre principe de restriction matérielle.



Monde M_1 Monde M_2 Monde M_3

FIGURE 2.3 : Paradoxe de Chandler

Selon Chandler, il y a certains mondes possibles qui ne sont pas accessibles depuis d'autres. Les multiples cheminements qui sont utilisés dans le mécanisme du paradoxe sont peut-être la cause des contradictions qu'il expose. Chandler avance l'idée que les relations d'accessibilité sont dynamiques, qu'elles varient selon le monde possible considéré :

> I want to suggest that what is possible varies from world to world. It should be noticed at the outset that the possibility in question is possibility 'in a broadly logical sense', to use Plantinga's words; not, for example, physical possibility. (CHANDLER, 1976, p. 106)

Ainsi, il se peut que, relativement au monde actuel M_1, le monde M_3 ne soit tout simplement pas accessible, même si, en revanche, on admet que relativement à M_1, le monde M_2 demeure bien accessible.

Comment procéder de manière technique ? Mettons à jour les mécanismes du paradoxe. Généralisons sa forme en considérant qu'un vélo v est constitué, à un moment t, d'un amas matériel p_1. Pour tout entier $i > 1$ on définit de manière récursive p_i comme étant un amas matériel qui est presque, mais pas complètement, coextensif avec p_{i-1}. Finalement, il existe une valeur entière m telle que p_m et p_1 ne recouvrent aucune partie commune. Si on note $\mathrm{M}xy$ pour « x est matériellement constituée par y », on peut alors interpréter de manière formelle le paradoxe de Chisholm de la manière dont le fait Graeme Forbes (FORBES, 1984) :

$$\begin{array}{ll} (\mathrm{PC}_1) & \Diamond \mathrm{M}vb_1 \\ & \Box(\mathrm{M}vp_1 \rightarrow \Diamond \mathrm{M}vp_2) \\ & \ldots \\ & \Box(\mathrm{M}vp_{m-1} \rightarrow \Diamond \mathrm{M}vp_m) \\ \hline & \Diamond \mathrm{M}vp_m \end{array}$$

Quand on examine plus finement ce qui permet de passer d'une étape à l'autre dans l'argumentaire ci-dessus, on constate l'importance de l'inférence modale suivante :

$$\begin{array}{ll} (\mathrm{C}) & \Diamond \mathrm{A} \\ & \Box(\mathrm{A} \rightarrow \Diamond \mathrm{B}) \\ \hline & \Diamond \mathrm{B} \end{array}$$

Cette inférence est valide dans la logique S5, qui admet une relation totale, c'est-à-dire, des mondes uniformément accessibles entre eux. Une première façon de bloquer le paradoxe est donc de refuser cette accessibilité. Il y a bien un monde possible dans lequel la bicyclette a été construite avec les parties $(\mathrm{A}_2, \mathrm{B}_2, \mathrm{C}_1)$, mais en revanche, ce monde n'est pas accessible depuis le monde actuel où elle a été construite avec les parties $(\mathrm{A}_1, \mathrm{B}_1, \mathrm{C}_1)$:

> [...]'it is not possible that this bicycle should have come into existence made up of entirely different parts' cannot be interpreted to mean that there are no worlds of any sort in which the bicycle came into existence made up of entirely different parts. It can only be taken to mean that worlds in which this occurs, if there are any, are not possible relative to the actual world. (CHANDLER, 1976, p. 108)

Dans le paradoxe de Chisholm, nos progressions modales reviennent à contrefaire des contrefactuels, à possibiliser du possible. Autrement dit, nos hypothèses confèrent des possibles au second degré. La stratégie de Chandler consiste à dire que l'on peut faire un pas dans le possible, mais pas deux. La solution préconisée par Chandler a néanmoins un coût important, le rejet de S5. En outre, en dépit de sa similitude avec le paradoxe temporel du bateau de Thésée, il semble que la résolution du paradoxe modal s'écarte d'emblée de son modèle temporel, ce qui peut paraître étrange. En effet, nous pourrions nous attendre à ce qu'une solution au paradoxe modal soit également applicable au paradoxe temporel, par exemple en admettant que les mondes possibles servent à représenter des parties temporelles du monde.

De son côté, Graeme Forbes (FORBES, 1984, p. 171 sq.) fait remarquer que le paradoxe de Chisholm possède une forte coloration sorite. Il peut alors être utile de faire ressortir la caractéristique vague en reformulant le paradoxe. Or, dans S5, la formule $\Box(\mathrm{A} \rightarrow \Diamond \mathrm{B})$ est équivalente à la formule $\Diamond \mathrm{A} \rightarrow \Diamond \mathrm{B}$. Autrement dit, (PC) peut s'interpréter ainsi :

$$\begin{array}{ll} (\mathrm{PC}_2) & \Diamond \mathrm{M}sb_1 \\ & \Diamond \mathrm{M}sb_1 \rightarrow \Diamond \mathrm{M}sb_2 \\ & \ldots \\ & \Diamond \mathrm{M}sb_{m-1} \rightarrow \Diamond \mathrm{M}sb_m \\ & \hline \\ & \Diamond \mathrm{M}sb_m \end{array}$$

Sous cette forme, le paradoxe est exprimé, aux opérateurs modaux près, sous une forme sorite dont la forme générale est la suivante :

$$\begin{array}{ll} (\mathrm{PS}) & \mathrm{P}a_1 \\ & \mathrm{P}a_1 \rightarrow \mathrm{P}a_2 \\ & \ldots \\ & \mathrm{P}a_{m-1} \rightarrow \mathrm{P}a_m \\ & \hline \\ & \mathrm{P}a_m \end{array}$$

Dans sa formulation sorite, le principe de tolérance constitutive est maintenant énoncé de la manière suivante :

$$(\mathrm{TCS}) \quad \forall s \forall b_1 \forall b_2 \, (Sb_1b_2 \rightarrow (\Diamond Msb_1 \rightarrow \Diamond Msb_2))$$

Notons que, contrairement à (TCM), cette nouvelle formulation ne contient pas d'opérateur modal emboîté dans la conséquence du conditionnel. Les deux principes (TCS) et (TCM) demeurent équivalents dans S5, du fait de l'équivalence, dans cette logique, entre $\Box(A \rightarrow \Diamond B)$ et $\Diamond A \rightarrow \Diamond B$. En revanche, si on introduit une relation d'accessibilité non transitive, l'équivalence n'est plus valide.

Forbes fait remarquer que dans sa forme sorite, le paradoxe ne provient cependant pas d'un problème modal. Cela revient à dire qu'il ne faut donc pas trop prématurément rejeter S5 sous le prétexte de résoudre le paradoxe de Chisholm : en effet, la forme sorite peut laisser croire que le cœur du problème réside peut-être ailleurs que dans les complexités modales. Forbes reprend ainsi l'idée, avancée par Kripke, que le paradoxe de Chisholm est un paradoxe du vague. Reste à savoir où doit exactement être placé le vague dans cette affaire. Étant donné qu'il n'y a de vague ni dans les limites d'un amas matériel particulier ni dans la relation de constitution, il doit provenir du concept d'être possiblement identique au vélo de départ. Mais dans un contexte modal standard, ce vague devrait obligatoirement être attribué aux conditions d'identité transmondaine. La solution que privilégie Forbes consiste à laisser intacte la logique de l'identité et à privilégier, pour les individus transmondains, la théorie des contreparties de Lewis.

Forbes (FORBES, 1984) en déduit que c'est le *modus ponens* qui doit être remis en cause dans son usage en logique modale. Dans le cas d'arguments liés au vague, le *modus ponens* ne s'applique plus. Forbes développe une logique qui invalide le principe de bivalence pour les énoncés vagues. L'idée est que les énoncés relatifs au vague peuvent s'évaluer selon des valeurs intermédiaires entre le vrai et le faux, de manière continue sur l'intervalle [0,1]. Dans une telle logique du vague, le degré de vérité d'une conjonction (respectivement d'une disjonction) est égal au degré minimum (respectivement maximum) de vérité des conjoints (respectivement des disjoints). On définit la vérité de l'implication matérielle selon le degré d'écart à la vérité entre l'antécédent et le conséquent. L'implication n'est absolument vraie que si le degré de vérité de l'antécédent est inférieur ou égal au degré de vérité du conséquent. En revanche, si l'antécédent à un degré de vérité supérieur au conséquent, le conditionnel n'est pas absolument vrai, sa vérité est égale au complémentaire par rapport à 1 du degré de différence excédentaire. Plus formellement,

$$\mathrm{Deg}(A \rightarrow B) = \begin{cases} 1 & \text{si Deg(A)} \leqslant \text{Deg(B)} \\ 1 - (\text{Deg(A)-Deg(B)}) & \text{sinon} \end{cases}$$

Avec cette manière de procéder, on voit qu'une inférence n'est absolument vraie que si le degré de vérité de la conclusion d'un argument n'a jamais un degré moins élevé que le plus fort degré de ses prémisses. Or, dans l'acheminent du paradoxe de Chisholm, le degré de vérité de $Pa_{n-1} \rightarrow Pa_n$ est toujours inférieur à 1. En outre, en augmentant le nombre de prémisses de cette forme, le conditionnel voit progressivement sa valeur de vérité approcher de 0. Cela a donc pour effet d'invalider le principe de *modus ponens*, car, comme le remarquent Pascal Engel et Frédéric Nef,

> Quand les prémisses sont plus proches du degré 1, le raisonnement est valide, mais il n'est pas cohérent (*sound*), puisque toutes les prémisses ne sont pas (absolument) vraies. Quand elles approchent de 0, le raisonnement est valide. (ENGEL & NEF, 1988, p. 486)

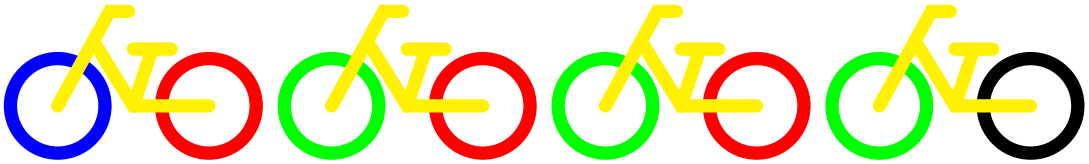
La réévaluation du *modus ponens* est accompagnée chez Forbes d'un rejet de la sémantique

modale classique au profit de la théorie des contreparties de Lewis. En effet, dans la sémantique modale standard, l'identité transmondaine est absolue, il n'y a pas de degré d'identité : pour tout individu a et tout individu b, a est ou n'est pas identique à b. Forbes rejette ce caractère absolu en lui substituant une relation relative par degré de ressemblance. Ce n'est pas l'identité qui est ainsi relativisée par Forbes, c'est l'identité entre mondes qui est perçue comme vague.

Cette stratégie est cependant coûteuse puisqu'elle demande d'abandonner le *modus ponens* en logique modale. En outre, soulignons qu'il n'est pas du tout certain que le paradoxe de Chisholm soit un type de paradoxe sorite. D'autres versions du paradoxe ne font pas intervenir cette composante, autrement dit, évitent l'artifice des petits changements successifs ((MACKIE, 1987), (ADAMS, 2007)). L'argument est le suivant : on peut affirmer que toutes les manières dont Socrate aurait pu être sont aussi des manières dont Platon aurait pu être. Et vice versa. Or l'une des façons dont Socrate aurait pu être est justement la façon dont Platon est actuellement. Et inversement, l'une des façons dont Platon aurait pu être est justement la façon dont Socrate est actuellement. Il y a donc un monde possible dans lequel Socrate joue le rôle de Platon et un monde possible dans lequel Platon joue le rôle de Socrate. Mais il n'y pas de raison pour que ces deux mondes diffèrent par ailleurs. Dès lors, nous obtenons un monde possible dans lequel Socrate joue le rôle de Platon et un monde possible, en tout point identique au précédent mis à part le fait que Platon y joue le rôle de Socrate. Nous retrouvons ainsi des mondes indiscernables quoique numériquement différents sans même être passés par des mondes intermédiaires. La composante sorite du paradoxe semble ainsi être superfétatoire.

Il faut souligner que ce genre de stratégie laisse intactes les situations d'indiscernabilité. C'est du moins ce qui ressort de la formulation que Nathan Salmon (SALMON, 1981, p. 230) donne lui-même au paradoxe de Chisholm, version qualifiée de paradoxe des quatre mondes. Considérons à nouveau que dans le monde actuel M_1 une bicyclette soit constituée des parties $(\mathrm{A}_1, \mathrm{B}_1, \mathrm{C}_1)$. Nous avons l'intuition modale que C_1 aurait pu servir de partie à une autre bicyclette. Autrement dit, dans le langage de la sémantique modale, il y a un monde possible M_4 dans lequel une bicyclette est constituée des parties $(\mathrm{A}_2, \mathrm{B}_2, \mathrm{C}_1)$. Notre principe de restriction assure qu'une telle bicyclette serait numériquement différente de la bicyclette de M_1, puisqu'elle s'en distinguerait par les deux parties A_2 et B_2. Par ailleurs, en appliquant le principe de tolérance constitutive à la bicyclette de M_4, nous pourrions alors affirmer que cette bicyclette pourrait avoir l'une de ses parties échangée. En particulier, il y a un monde possible M_3 dans lequel cette bicyclette est constituée des parties $(\mathrm{A}_2, \mathrm{B}_1, \mathrm{C}_1)$. Et de façon similaire, puisque selon le principe de tolérance constitutive l'une des pièces de la bicyclette de M_1 aurait pu être échangée, il y a donc un monde possible M_2 dans lequel cette bicyclette est constituée des parties $(\mathrm{A}_2, \mathrm{B}_1, \mathrm{C}_1)$.

À quoi aboutissons-nous ? Nous avons deux mondes M_2 et M_3 qui sont qualitativement indiscernables, quoique habités par des individus différents. Car si les bicyclettes de M_2 et M_3 étaient identiques, la transitivité de l'identité impliquerait alors que les bicyclettes contenues dans les mondes M_1 et M_4 le seraient également, ce qui contredirait notre hypothèse de départ. Nous sommes donc en présence de deux mondes indiscernables, mais numériquement différents, ce que les philosophes appellent des mondes *haeccéitiques*.



Monde M_1 Monde M_2 Monde M_3 Monde M_4

FIGURE 2.4 : Paradoxe des quatre mondes

2.2.5 Les haeccéités ou essences individuelles

Comme nous l'avons vu, une conséquence importante du paradoxe de Chisholm réside dans la possibilité d'individus transmondains indiscernables et celle de mondes globalement indiscernables. Or pour certains philosophes, les identités nues sont tout simplement inacceptables. On peut se demander pourquoi ces auteurs sont aussi réfractaires à cette idée. Dans bien des cas, la raison n'est pas indiquée, comme s'il était absolument manifeste que les identités nues constituent d'évidentes et importantes entorses à la raison :

> Is it really the case that Queen Victoria could have had all your actual properties (except for identity with you) while you had all of hers (except for identity with her)? (MACKIE & JAGO, 2017, p.)

Il semble alors qu'il y ait ici chez eux une forme de concession modale au principe métaphysique d'identité des indiscernables, dans l'idée que deux mondes ne pourraient être numériquement distincts en étant qualitativement identiques. Il faut néanmoins distinguer ici deux phénomènes : l'un est épistémique, il concerne la capacité à réellement penser à deux mondes indiscernables. L'autre est métaphysique, il a trait à l'existence de deux mondes indiscernables. Quatre positions peuvent alors être occupées concernant la questions des individus maximaux indiscernables, selon qu'on admet ou non les mondes indiscernables sous un mode métaphysique ou épistémique. Par exemple, Lewis admet bien qu'on puisse concevoir des mondes indiscernables, mais il rejette la possibilité métaphysique de tels mondes.

Jusqu'à présent, nous avons eu surtout eu affaire à trois types d'essentialisme, qu'on peut résumer ainsi :

- l'essentialisme maximal : toutes les propriétés d'un individu lui sont essentielles ;
- l'essentialisme minimal : aucune des propriétés d'un individu ne lui sont essentielles ;
- l'essentialisme sortal : les propriétés essentielles d'un individu relèvent de la sorte d'individu qu'il est ;

Pour échapper au problème des identités nues, une quatrième forme d'essentialisme stipule des essences individuelles. Cela consiste à particulariser les propriétés aux individus.

Selon l'essentialisme de Plantinga, lorsqu'on considère une proposition comme « Emmanuel Macron a été élu président », il faut accorder au moins deux propriétés à Macron, l'une accidentelle, à savoir, « avoir été élu » et l'autre essentielle, « être identique à Macron ». Or d'après la théorie de Lewis, aucune des contreparties ne peut exemplifier cette dernière propriété, autrement dit, aucune des contreparties de Macron ne possède la propriété d'être identique à Macron. Il faudrait par conséquent rejeter la vérité de la proposition « Emmanuel Macron aurait pu être élu président », qui traite d'un monde alternatif où Macron est à nouveau élu. Le partisan des contreparties ne pourrait admettre la vérité littérale de cette proposition. Or cette proposition a été construite de manière arbitraire à partir d'un objet exemplifiant une certaine propriété du monde actuel, en généralisant, on peut constater que la théorie des contreparties est incapable, pour Plantinga, d'accommoder nos intuitions modales à l'égard des particuliers concrets.

La solution de Plantinga au paradoxe de Chisholm consiste à bloquer la conclusion des identités nues, en affirmant que, quelque part dans nos transactions modales, nous avons pris la liberté d'échanger certaines propriétés qui ne sont pas attribuables à un autre individu que celui qui les exemplifie. Parmi les propriétés de Platon, il doit exister une essence individuelle, autrement dit, une ou plusieurs propriétés qui sont à la fois des conditions nécessaires et suffisantes pour être Platon. Cette précaution permet ainsi d'assurer que si Socrate et Platon échangent toutes leurs propriétés accidentelles, il resterait toutefois au moins une propriété essentielle que l'un mais pas l'autre posséderait.

Mais que serait une telle propriété ? Répétons qu'elle devrait appartenir à l'un des deux individus et exclusivement à lui, tout en étant une propriété essentielle. Or il paraît bien difficile de concevoir une propriété non triviale qui remplirait une telle condition. En tout cas, elle ne saurait être une propriété essentielle que Socrate et Platon partagent actuellement, comme être des humains, être issus d'un ovocyte fécondé, etc. La seule possibilité est alors de faire intervenir des propriétés comme « être identique à Socrate », « être le produit du développement de tel ovocyte fécondé par tel spermatozoïde », etc. Cette idée consiste en réalité à admettre la thèse d'essences individuelles, c'est-à-dire des propriétés ou des faisceaux de propriétés dont 1) chaque individu serait l'unique possesseur, et 2) qui assurerait à son possesseur son identité propre. L'essence individuelle P d'un individu a serait ainsi caractérisée par la condition suivante : $\forall x(\mathrm{P}x \leftrightarrow x = a)$

La question qui est soulevée ici est finalement celle des propriétés *pures*. Les essences individuelles, qui sont des propriétés impures, sont généralement appelées des *haeccéités* selon l'appellation de Chisholm (CHISHOLM, 1967) que David Kaplan (KAPLAN, 1979) attribue à Robert Adams (ADAMS, 1979) [38].

Adams (ADAMS, 1979) utilise de son côté le terme de *thisness*, ou propriété d'être tel individu particulier plutôt que tel autre opposé à *suchness*, ou *quiddité*, autrement dit, l'ensemble des propriétés qualitatives d'un individu. L'*haeccéité* d'un individu serait la propriété d'être identique à lui-même, en tant qu'individu particulier.

> A thisness is the property of being identical with a certain particular individual-not the prop-

[38] Adams attribue le terme à Duns Scot en suivant Johannes Kraus (KRAUS, 1927, pp. 93-94). Mais soulignons que l'authenticité scotiste du terme est néanmoins discutée. Quoique l'usage du terme soit popularisée au sein de son école, les occurrences du terme d'*haecceitas* sont plutôt rares dans le corpus de Duns Scot. Pour une introduction historique à la notion d'haeccéité, on pourra consulter l'article (CROSS, 2014).

> erty that we all share, of being identical with some individual or other, but my property of being identical with me, your property of being identical with you, etc. (ADAMS, 1979, p. 6)

La relation P recherchée pour constituer une essence individuelle serait donc « être identique à un particulier a » (P$x := (x = a)$). Il faut dissocier cette propriété de celle partagée par tout ce qui est, d'être identique à soi-même. L'*haeccéité* n'est pas la relation réflexive $x = x$. Ici, on a bien affaire à une propriété individuelle, instanciée exclusivement par l'individu a à l'exception de tous les autres, et qui fournit à celui-ci l'identité avec lui-même. En outre, cette propriété est impure puisqu'elle dépend d'un particulier.

Nous venons de voir que cette tentative d'individuer les particuliers consistait à utiliser des propriétés qui sont des relations à eux-mêmes. Pour admettre des essences individuelles, une autre possibilité consiste à individuer les particuliers par des relations à d'autres particuliers, et en particulier par leur origine. C'est notamment la position que défend Kripke dans la troisième conférence de *Naming And Necessity*, où les attaques à l'égard de la thèse sémantique de Russell prennent une coloration plus métaphysique :

> La reine - cette femme même - aurait-elle pu avoir des parents différents des parents qu'elle a eus en réalité ? Aurait-elle pu être, par exemple, la fille de M. et Mme Truman [39] ? (KRIPKE, 1982, p. 100)

Selon Kripke, nous avons deux principes qui régissent les affectations modales d'un individu transmondain et qui assurent son identité de monde à monde.

1. Nécessité sortale : un individu est nécessairement de la même sorte que celle qu'il instancie actuellement.
2. Nécessité d'origine : un individu a nécessairement la même origine que son origine actuelle.

Soulignons qu'en ce qui concerne les espèces biologiques, le deuxième principe implique le premier. L'origine des êtres vivants ne peut être modifiée en pensée sans rompre la persistance de l'individu entre deux mondes. Socrate n'aurait pas pu avoir d'autres parents.

De fait, nier que l'origine biologique soit une propriété nécessaire pose de manifestes problèmes pour démêler l'écheveau des qualités nécessaires et contingentes. Car si on échange les parents d'un individu entre deux mondes, alors certaines qualités accidentelles dans ce monde peuvent devenir nécessaires dans l'autre. Par exemple, la couleur des yeux peut être accidentelle dans le cas de la présence simultanée de gènes dominants et récessifs dans les gamètes des deux parents, mais elles peuvent s'avérer nécessaires si les deux parents possèdent seulement les allèles dominants. Il devient alors clair que Socrate et Platon ne peuvent pas échanger inpunément toutes leurs propriétés accidentelles *ceteris paribus*.

Kripke fait remarquer que nos raisonnements contrefactuels portent toujours sur des individus déjà pris dans une certaine histoire. C'est à partir de l'instant t de la carrière temporelle d'une personne qu'on peut éventuellement lui attribuer des événements contrefactuels. Si ce moment t peut

[39] Kripke fait référence à la reine Elisabeth d'Angleterre, sœur de la princesse Maragaret. Les Truman avait une fille dénommée elle aussi Margaret. En faisant l'hypothèse d'un changement de parents, Elisabeth deviendrait ainsi la sœur de Margareth Truman et non de la princesse d'Angleterre.

être plus ou moins avancé dans la vie de l'individu, il semble toutefois impossible de le ramener au moment originel de la venue au monde sans répercussion sur son identité. Il y a donc une limite temporelle aux transactions modales à propos d'un même individu. Une phrase comme « Emmanuel Macron aurait pu avoir d'autres parents biologiques » est une illusion modale, elle exprime une proposition qui est fausse. L'identité dépend étroitement des conditions généalogiques. Pour qu'un individu x soit le même qu'un individu y, il est donc tout à fait nécessaire, selon Kripke, que x et y aient la même origine biologique.

Est-ce que cette condition est aussi suffisante ? Selon Graeme Forbes, c'est bien le cas, l'origine est une condition à la fois nécessaire et suffisante d'individuation. Dans *Origin and Identity*, Forbes dénomme « propagule » l'entité générée par un individu biologique (FORBES, 1980). Une entité a est une propagule d'une entité b signifie que a est une des entités qui a grandi ou s'est développée dans b. Le critère d'identité transmondaine constitué par Forbes est le suivant :

- (Individuation par l'origine) Un individu biologique x dans un monde M_1 est identique à un individu y dans un monde M_2 si et seulement si :
  1. x est de la même sorte que y et
  2. Les propagules de x dans M_1 sont identiques aux propagules de y dans M_2.

En outre, la question est finalement de savoir si les conditions d'identité sont réellement valides. L'origine peut-elle véritablement servir de critère à l'individuation ? Constatons d'abord l'asymétrie entre les artefacts et les êtres biologiques. En effet, en faisant appel à la nécessité de l'origine, on trouve peut-être une issue au paradoxe de Chisholm en ce qui concerne les êtres vivants, en revanche, cela ne semble pas fonctionner pour les artefacts. On peut légitimement admettre par exemple que le plastique qui servit à la constitution d'un objet donné aurait pu aussi servir à la confection d'un autre.

Un autre problème auquel se heurte cette conception est celui de la dépendance et de la régression à l'infini. Étant individué par une entité génitrice, l'identité d'un individu dépend d'autre chose que lui. Or cet autre doit, bien sûr, lui-même répondre aux mêmes conditions d'individuation par l'origine, autrement dit, provenir d'une cellule fécondée, qui dépende elle-même de gamètes provenant d'un spermatocyte et d'un ovocyte. Pourtant, il faudrait que les conditions d'identité débouchassent sur quelque entité supportant la chaîne entière, mais l'on voit que la régression semble s'échapper à l'infini.

Il y a encore un autre problème. Certains organismes sont capables de régénérer leur constitution après avoir subi une section. On peut citer des exemples de plantes et d'animaux, qui, dans des conditions favorables à la régénérescence de leurs parties, peuvent produire deux fois le même individu à partir de deux parties amputées. D'autre part, on peut aussi mentionner le cas des jumeaux homozygotes, qui résultent de divisions mitotiques d'une seule et même cellule originelle. Si l'on suppose que le zygote dont est issu Platon aurait pu donner naissance à deux frères jumeaux, le critère d'individuation de Kripke-Forbes implique alors que Platon est identique aux deux jumeaux à la fois.

De telles objections incitent donc à considérer que la condition sur l'origine n'est pas une condition suffisante de l'individuation des particuliers. De toute façon, la phrase « Macron ait pu

avoir d'autres parents » semble bien être compréhensible, et cela, même au sens biologique du terme. La stratégie de Kripke consiste à embrigader les possibilités logiques dans des possibilités métaphysiques, en rejetant comme illusions celles qui ne conviennent pas à l'essentialisme des origines.

L'une des conséquences de l'haeccéitisme concerne les possibilités de métamorphoses. De monde à monde, il devient difficile de contraindre certaines propriétés à demeurer essentielles. Or si toutes les propriétés sont accidentelles, par petites touches successives entre monde, il devient alors possible de donner du crédit à n'importe quelle idée de métamorphose. Le problème n'est pas seulement que Socrate et Platon auraient pu échanger leurs rôles respectifs. Socrate pourrait aussi échanger son rôle avec celui d'un œuf poché. N'importe quoi aurait pu être transformé en n'importe quoi d'autre, ou comme l'écrit Lewis, n'importe qui aurait pu être un œuf poché. Et même s'il l'on admet qu'il n'est pas actuellement possible pour Socrate d'être un œuf poché, Lewis souligne qu'il est néanmoins possiblement possible qu'il le soit, conclusion qu'il reconnaît comme la plus gênante de l'haeccéitisme (LEWIS, 2007, p. 365).

De fait, si l'on admet des haeccéités au niveau des individus, échapper aux possibilités de métamorphoses nécessite aussi d'étendre les haeccéités aux catégories sortales, sans quoi il faut se résigner à admettre qu'elles-mêmes pourraient s'avérer différentes. Sans doute que Macron n'aurait pas pu ne pas être un homme, puisqu'on y réfère par la médiation de cette sorte dans le monde actuel. Mais être un homme aurait toutefois pu relever de choses très différentes. Ainsi, si l'on admet des haeccéités au niveau des individus afin de bloquer la conclusion de permutation qui apparaît dans le paradoxe, il s'agit aussi de les admettre au niveau des catégories sortales. Car ce ne sont pas seulement les individus qui sont susceptibles de modifications de leurs propriétés, mais également les sortes.

Pour Kripke, certaines propriétés demeurent essentielles pour maintenir monde-à-monde l'identité de l'individu considéré. Macron aurait pu perdre les élections, mais en revanche, il n'aurait pas pu être un presse-purée ni un ornithorynque. Macron aurait certes pu être différent, mais il n'aurait pas pu être n'importe quoi. Si l'intuition nous autorise à croire qu'il aurait pu être différent en bien des façons, les marges de ces altérations sont cependant bornées par le maintien de propriétés essentielles, comme par exemple instancier le prédicat « humain ». Dans aucun des mondes possibles où il existe, Macron n'aurait pu être un *Ungezieffer*.

On pourrait néanmoins apprendre que Macron est en réalité un extraterrestre ou un robot très perfectionné comme le sont les répliquants dans le film *Blade Runner* de Ridley Scott. Il n'y aurait aucune impossibilité logique à cela. Mais il faut bien dissocier ici les notions de possibilité et de nécessité épistémiques des notions de possibilité et de nécessité métaphysiques. D'après Kripke, Macron n'aurait pas pu être autre chose qu'un homme *étant donné* qu'il en est un :

> La question de savoir si Nixon aurait pu ne pas être un être humain est un bon exemple de question qui ne relève pas de la théorie de la connaissance. Supposons que Nixon se révèle être un robot. Cela pourrait arriver. Nous pourrions avoir à déterminer si Nixon est un être humain ou un robot. Mais c'est là une question au sujet de notre connaissance. Par contre, la question de savoir si Nixon aurait pu ne pas être un être humain, étant donné qu'il en est un, cette question ne concerne pas notre connaissance, qu'il s'agisse de notre connaissance *a posteriori* ou de notre connaissance *a priori*. Elle concerne la façon dont les choses auraient

pu se passer si elles ne se passaient pas comme elles se passent. (KRIPKE, 1982, p. 35)

Dans *Naming and Necessity*, Kripke utilise trois exemples de contrefaçons portant sur des classes : l'illustration du faux-or, du faux-tigre, de la fausse-eau. Jusqu'à quel point une catégorie sortale peut-elle différer ? Et inversement, que devrions-nous penser d'un métal ayant toutes les propriétés qui ont fixé la référence du terme « or », mais qui ne serait pas de l'or ? Nous aurions une forme de double par imitation. Toutes les propriétés servant à reconnaître l'or seraient représentées. Mais il se pourrait que la structure interne ne soit pas celle de l'or. Dirions-nous que nous avons découvert que ce que nous pensions de la structure de l'or est finalement faux ? Pour Kripke, ce ne serait vraisemblablement pas le cas. Nous dirions plutôt avoir découvert une nouvelle substance ressemblant fortement à de l'or.

Kripke conteste que les catégories sortales relèvent de propriétés contingentes qui seraient référencées par des descriptions. Est-ce que nous apprenons ce qu'est un tigre par une définition ? Est-ce par la description qu'en donnerait un dictionnaire ? Une telle description pourrait par exemple être : « grand félin carnivore quadrupède à rayures, etc. ». Mais nous pourrions rencontrer un tigre à trois pattes. S'agirait-il encore seulement d'un tigre ? Nous devrions bien sûr répondre par l'affirmative. Il s'agirait d'un tigre à trois pattes. Mais comment ne pas y voir une contradiction dans les termes, puisque selon notre définition, être un « tigre » implique d'avoir quatre pattes ? Avec la théorie descriptiviste, sommes-nous encore seulement en mesure de reconnaître un tigre ?

Admettons plus encore que cette rencontre avec un tigre à trois pattes soit l'occasion de nous révéler que nous avions une croyance erronée au sujet du nombre de pattes des tigres, et qu'il faille finalement admettre que ces animaux n'en possèdent bel et bien que trois. Devrions-nous dire alors que ce que nous croyions être des tigres jusque-là n'en sont finalement pas ? Bien plutôt, constate Kripke, nous réviserions notre jugement au sujet des tigres sans contester le moins du monde qu'il s'agisse là de la même espèce animale.

Certains animaux, très ressemblants aux tigres, pourraient s'avérer être en réalité des sortes de reptiles. Devrions-nous en conclure que certains tigres sont des reptiles ? Il semble plus probable d'admettre que de tels animaux, en dépit de leur ressemblance frappante, ne seraient tout simplement pas des tigres.

> Même si nous ne connaissons pas la structure interne des tigres, nous supposons (à juste titre, probablement) que les tigres forment une espèce naturelle. Nous pouvons, dès lors, imaginer une créature qui, tout en ayant toutes les apparences extérieures des tigres, en différerait suffisamment du point de vue interne pour que nous disions que ce n'est pas une chose de la même espèce. Nous pouvons imaginer cela - sans savoir ce qu'est cette structure interne. Nous pouvons dire à l'avance que nous employons le mot « tigre » pour désigner une espèce, et que tout ce qui n'est pas de cette espèce, même si cela ressemble à un tigre, n'est pas un tigre. (KRIPKE, 1982, p. 109)

Il y a donc, selon Kripke, des essences, données par des structures internes qui organisent les êtres vivants. La science consiste à découvrir ces propriétés essentielles qui ne sont pas seulement nominales, mais, pour Kripke, bien réelles. De ce point de vue, les catégories sortales sont des désignateurs rigides non seulement au sens sémantique, mais plus encore, au sens métaphysique.

Nous savons aujourd'hui que l'eau est la molécule H_2O. Mais nous avons d'abord appris à désigner l'eau d'une manière différente. Initialement, on a désigné la substance par les impressions caractéristiques qu'elle procure au toucher, au goût, etc. Que signifierait découvrir une substance en tout point identique avec l'eau, mais dont la structure atomique ne serait pas H_2O ? Dirions-nous que nous avons découvert que l'eau n'est pas nécessairement H_2O ? Serait-ce réellement de l'eau ? Là encore, Kripke soutient que « nous dirions plutôt qu'il y a une fausse eau, comme il y a du faux or ; c'est-à-dire une substance qui, tout en ayant les propriétés par lesquelles nous avons identifié l'eau à l'origine, ne serait pas vraiment de l'eau » (KRIPKE, 1982, p. 117). Pour Kripke, lorsque la science affirme que l'eau, c'est H_2O elle ne procède pas seulement par description, elle détermine plutôt une essence. La science déterminerait ainsi ce qui est nécessaire pour être de l'eau. Mais cette nécessité est donc, pace Kant, établie *a posteriori*. Ce point de vue induit, cette fois paceHume, une séparation entre les modalités épistémiques et aléthiques. Car pour Kripke, croire que l'eau aurait pu ne pas être H_2O n'est rien d'autre qu'une illusion modale : il n'y a pas d'autre possibilité, pour l'eau, que d'être H_2O. Les modalités épistémiques sont donc désolidarisées des notions de nécessité et de contingence. Il y a des choses qui sont concevables, mais qui ne sont pas possibles. L'application de la théorie des désignateurs rigides aux sortes encourage donc l'idée que nos conceptions peuvent évoluer - nous pouvons par exemple découvrir que l'eau est H_2O mais sa signification restera cependant la même. Dans la théorie de Kripke, l'extension d'un terme et son intension sont ainsi fortement désolidarisées l'une de l'autre.

Un essentialisme reposant sur des arguments sémantiques a également été défendu par Hilary Putnam, qui nous invite à considérer l'expérience de pensée suivante. Imaginons un double de la Terre, que nous appellerons Terre-Jumelle. Cette planète serait en tout point ressemblante à la nôtre, excepté que ce qui apparaît comme de l'eau et qui est appelé là-bas « eau » consisterait en fait en une masse constituée de molécules XYZ différentes de notre H_2O terrestre. Selon Putnam, si des Terriens venaient à découvrir la différence interne des molécules d'eau de Terre-jumelle, ils en concluraient que l'eau terrienne et l'eau de Terre-Jumelle sont en fait deux choses différentes, quoiqu'elles porteraient le même nom.

Mais l'expérience de pensée se poursuit ainsi : imaginons maintenant que les Terriens viennent à découvrir l'eau-de-Terre-jumelle à une époque où la structure chimique de la matière n'était pas encore connue, ni sur Terre, ni sur Terre-jumelle. Remarquons d'abord que le changement d'époque n'affecterait en rien les extensions respectives du terme « eau » sur Terre et sur Terre-jumelle. Si l'on admet leur distinction numérique dans le présent, nous devrions alors aussi admettre leur distinction dans le passé. Considérons alors un Terrien de cette époque et son doppelgänger synchronique sur Terre-Jumelle. Admettons en outre que les deux individus entretiennent exactement les mêmes croyances au sujet de ce qu'ils appellent l'un et l'autre « eau ». Leurs intentions respectives à propos de l'eau pourraient s'avérer exactement identiques, quoique les extensions du terme s'avérerait évidemment distinctes. Les doppelgängers ne différeraient pas par des relations intrinsèques [40]. mais par des relations extrinsèques, leurs rapports respectifs à H_2O et à XYZ. L'expérience de pensée de Terre-Jumelle tend ainsi à montrer que les intensions d'un terme peuvent être identiques sans que le soient leurs extensions respectives. Autrement dit, et *pace* Frege, le sens d'un terme ne détermine pas son extension.

Une objection à ce constat consisterait à nier que les extensions ne soient pas affectées par les

[40] A l'exception près que la composition de leurs sangs respectifs.

changements d'époques. Putnam considère alors la référence du terme établie par échantillonnage : « voici un verre contenant de l'eau ». Cet énoncé consiste en une identité ostensive, contenant une condition nécessaire et suffisante pour que la communauté qui use du même terme avec la même référence accepte ou rejette sa véracité. Elle pourrait être par exemple rejetée parce que le verre pourrait contenir en réalité du gin :

> La définition ostensive contient ce que l'on pourrait appeler une condition nécessaire et suffisante dont on peut se « défausser » ; la condition nécessaire et suffisante pour être de l'eau est d'entretenir avec la relation *mêmes*$_L$ avec la substance qui se trouve dans le verre ; mais c'est là une condition nécessaire et suffisante uniquement si la présupposition empirique est vérifiée. Dans le cas contraire intervient une condition appartenant à une série de conditions pour ainsi dire en « réserve ». (PUTNAM, 1985, p. 24)

Mais la *mêmeté*$_L$ entendue ici est une relation théorique dont l'usage peut nécessiter une quantité indéterminée de recherches scientifiques. Que le terrien d'une époque plus avancée refuse d'appeler « eau » la substance XYZ là où le terrien d'une époque plus reculée en aurait admis l'usage, cela n'implique pourtant pas que, dans l'intervalle, la signification du terme ait changé. L'eau du Michigan de 1750 est identique à l'eau du Michigan de 2016. Ce qui change, ce sont les croyances au sujet de la relation de *mêmeté*$_L$ entretenue entre l'eau du Michigan et l'eau de Michigan-de-Terre-jumelle. D'une époque l'autre, cette croyance fausse ne serait plus entretenue.

Putnam prolonge son expérience de pensée ainsi : admettons que seul un expert puisse différencier les marmites de molybdène des marmites d'aluminium. Admettons en outre que leur rareté soit inverses sur Terre et Terre-jumelle, autrement dit que, sur Terre-Jumelle l'aluminium soit rare alors que le molybdène, répandu. Les marmites de Terre-jumelle seraient donc faites en molybdène. Enfin, admettons que les termes soient eux-mêmes inversés d'une planète à l'autre. L'aluminium terrestre serait appelé molybdène sur Terre-jumelle, et le molybdène terrestre serait nommé aluminium sur Terre-jumelle. Leurs rôles respectifs seraient ainsi échangés d'une planète à l'autre. Des visiteurs non experts d'une planète en provenance de l'autre planète ne différencieraient pas ce qu'ils nomment aluminium de ce que leurs hôtes nomment molybdène. Mais une expertise du métal pourrait toutefois révéler leur méprise.

Remarquons que l'erreur d'identification ne concernerait pas toute la communauté linguistique, mais seulement une partie. Là encore, des individus non-experts de planètes différentes auraient le même genre d'état psychologique au sujet du constituant des marmites, alors que le terme utilisé pour qualifier leur matériau n'aurait pas les mêmes extensions. Putnam conclut par le slogan que « les significations ne sont pas dans la tête ». Deux individus pourraient avoir les mêmes intensions sans qu'il y ait identité entre les extensions. On pourrait donc avoir affaire à une indiscernabilité sémantique sans avoir affaire à une indiscernabilité ontologique, ce que Putnam désigne sous la notion de *stéréotype*, « A standardized description of features of the kind that are typical, or 'normal', or at any rate, stereotypical » (PUTNAM, 1975, p. 230). Chez Putnam, le stéréotype tend à désolidariser le contenu épistémique de la signification de son contenu référentiel métaphysique. L'« eau » et « l'eau-de-terre-jumelle » ne sont rien d'autre que des descriptions standardisées, les descriptions n'intervenant finalement que comme des composantes secondaires dans la théorie cognitive de la référence de Putnam. Le stéréotype joue ainsi une sorte de rôle sémantique équivalent à celui du mode de présentation frégéen, sinon que la dénotation n'est plus

déterminée par le sens.

2.2.6 Individuation par stipulation et individuation par accointance

Le mérite de la conception de Putnam est de montrer le rôle que joue l'environnement dans la signification d'un terme d'espèce naturelle en dissociant le rôle épistémique joué par le stéréotype du rôle référentiel. Mais les théories sémantiques de Kripke et de Putnam sont fortement influencées par une propension métaphysique envers l'essentialisme. La description n'y joue jamais qu'un rôle d'auxiliaire dans la dimension cognitive de la signification.

Les explications de Kripke et Putnam, qui reposent sur l'hypothèse d'entités indiscriminables et numériquement différentes, sont, pour cette raison, exposées à des objections. En concevant des phénomènes indiscriminables mais numériquement distincts, Kripke et Putnam bâtissent la différence numérique sur des haeccéités qui présupposent d'emblée ce que leurs théories tentent de démontrer. En ce sens, l'argument paraît circulaire. Si la science est sensée déterminer des essences, elle ne peut le faire que par observation et expérimentation. Si les entités XYZ et H_2O étaient, dans la représentation, parfaitement indiscriminables, on voit mal alors en quoi elles pourraient posséder des propriétés « scientifiques » objectivement différentes. Il semble même assez raisonnable de prétendre que les expérimentations faites par les scientifiques d'une planète avec un échantillon provenant d'une autre planète mèneraient sans doute strictement aux mêmes résultats que celles faites avec les échantillons d'eau de leur propre planète. De fait, on peut avoir des doutes légitimes sur le fait qu'un scientifique de Terre à qui l'on soumettrait l'examen d'un échantillon d'eau de Terre-jumelle puisse aboutir à une autre conclusion qu'une structure similaire à H_2O. Et indiquer un désaccord qui ne soit pas seulement linguistique au sujet de la nature des atomes composant l'eau de Terre et celle de Terre-jumelle impliquerait de toute façon une régression à l'infini, parce qu'il faudrait encore déterminer les essences de X, Y et Z et montrer en quoi elles différeraient effectivement de celles de H et O.

Dans d'autres mondes possibles, l'eau pourrait avoir des propriétés phénoménales différentes, une appellation différente, voire, ne pas exister. Mais ce que Kripke et Putnam refusent, c'est qu'un échantillon d'eau d'un monde possible ramené dans le monde actuel soit différent de ce qu'il est actuellement du point de vue scientifique. L'eau est nécessairement H_2O. Le stéréotype n'intervient donc pas dans les propriétés nécessaires de l'eau, et, comme l'écrit Manuel Rebuschi,

> Modalités épistémiques et modalités métaphysiques sont donc [...] clairement dissociées. Les propriétés essentielles sont véritablement possédées par les choses, et en dehors de la signification : on ne peut pas être plus éloigné des essences nominales lockéennes. Quel est l'impact d'une découverte comme celle de la structure chimique de l'eau ? Il s'agit d'une découverte empirique, donc *a posteriori*, d'une propriété essentielle de l'eau, c'est-à-dire d'une vérité nécessaire. Ce que Putnam défend, c'est qu'une telle découverte modifie notre connaissance du monde, mais ne modifie pas en premier lieu la signification linguistique du terme « eau », même si par la suite, la composition chimique de l'eau est peut-être venue renforcer le stéréotype. [...] Putnam conteste vivement les formes de conventionnalisme ou de relativisme selon lesquelles des modifications de nos croyances entraînent des changements de signification et d'extension de nos termes. A l'encontre d'un Thomas Kuhn prônant l'incommensurabilité de théories rivales car parlant de sujets distincts, Putnam objecte que nos si conceptions passent,

l'extension demeure. (REBUSCHI, 2008, p. 77)

C'est la présupposition métaphysique des essences qui guide la conception de doubles qualitatifs parfaits. Si les propriétés sont véritablement possédées par les choses, elles partagent donc entre elles une ressemblance qui est indépendante de nos manières de les représenter. La représentation est ici minée dès le départ par une sorte d'imitation qui se manifeste objectivement entre des choses, quoique cette ressemblance soit indépendantes de nos manières de les concevoir. Cela conduit Kripke et à stipuler l'hypothèse de choses numériquement différentes, quoique indiscriminables. Concevoir de telles entités n'a rien de bien fâcheux, mais les admettre au sein de l'ontologie, c'est une toute autre histoire.

Les théories de l'identité transmondaine (*cross-world identities*) peuvent être alors globalement partagées en deux courants majeurs. Les partisans du premier courant, essentiellement représentés par la sémantique des mondes possibles défendue par Kripke, souscrivent à la thèse de la référence rigide et stipulent qu'il n'y a pas de problème d'identification transmondaine pour la raison que les énoncés modaux sont toujours focalisés sur un objet particulier qui est un objet *donné*. Il n'est pas possible de perdre de vue cet objet, parce que, dans l'esprit, nous passons du monde actuel à un monde possible en reconstruisant ce dernier autour de l'objet actuellement visé.

La seconde théorie sémantique des mondes possibles, inaugurée par Jaakko Hintikka, rejette cette conception en proposant de considérer les individus de manière transmondaine comme des fonctions reliant des particuliers résidant dans différents mondes possibles. Selon Hintikka, l'erreur fondamentale de Kripke, c'est de croire que nous pouvons donner un nom à un individu dans ce monde, et de le promener dans d'autres mondes en employant ce nom comme une sorte de panneau indicateur. Parallèlement à la conception modale essentialiste de Kripke, Jaakko Hintikka a proposé une version épistémique qui, elle, ne rompt pas le lien entre les modalités épistémiques et les modalités métaphysiques. Hintikka reconnaît un héritage kantien, soulignant que ce que nous appelons réalité ne peut être complètement désintriqué de nos concepts :

> My reasoning ends on a distinctly Kantian note. Whatever we say of the world is permeated throughout with concepts of our own making. Even such prima facie transparently simple notions, as that of an individual, turn out to depend on conceptual assumptions dealing with different possible states of affairs... reality cannot be, in principle, wholly disentangled from our concepts. A Ding an sich, which could be described or even as much as individuated without relying on some particular conceptual framework, is bound to remain an illusion. (HINTIKKA, 1969a, p. 109)

Ainsi, la notion d'individu *simpliciter* n'a pas de signification, elle est toujours relative à un aspect, à une manière particulière de le concevoir. La réalité peut être conceptualisée de bien des manières dont les unes ne sont pas réductibles aux autres. La référence à un individu ou à une espèce naturelle n'est jamais rigide, mais toujours relative à une modalité particulière.

Hintikka utilise les mondes possibles pour interpréter des assertions épistémiques comme « Jean sait que Shakespeare a écrit *Le Roi Lear* » : dans tous les mondes possibles compatibles avec l'ensemble complet des croyances de Jean, il est vrai que Shakespeare a écrit *Le Roi Lear*. Ainsi, l'analyse des connaissances de Jean consiste ainsi à considérer tout un ensemble de mondes possibles qui lui sont indiscriminables du monde réel. Hintikka parle de multiplicité référentielle

pour qualifier le comportement des termes singuliers dans le champ des opérateurs intensionnels. De fait, la référence du nom Shakespeare dans le monde actuel n'est donc pas la seule chose qui compte puisque les références de ce nom dans les mondes possibles compatibles avec les croyances de Jean sont également importantes. Jean pourrait par exemple croire que Marlowe et Shakespeare ne sont qu'une seule personne. En particulier, Hintikka reformule ainsi le principe d'indiscernabilité des identiques :

PIdI* $a = b \rightarrow \forall \mathrm{F}(\mathrm{F}a \leftrightarrow \mathrm{F}b)$, à condition que les termes a et b réfèrent au même individu dans tous les possibles que F nous encourage à considérer.

La sémantique d'Hintikka s'oppose ainsi à celle de Kripke en ce qu'elle maintient un mode de présentation à travers la référence : un individu est compris comme une fonction qui prélève un objet du domaine d'un monde possible donné, objet appelé « manifestation » de cet individu dans le monde en question, et un autre objet du domaine d'un autre monde. Une ligne de monde relie ces deux objets issus de scénarios différents de telle manière que deux individus différents puissent partager les mêmes propriétés dans un scénario donné, sans néanmoins les partager toutes entre les autres scénarios.

> What becomes of an individual is now an atemporal and non-spatial 'world line' through several possible worlds. In equivalent terms each individual in the full sense of the word is now essentially a function which picks out from several possible worlds a member of their domains as the 'embodiment' of that individual in this possible world or perhaps rather as the role which that individual plays under a given course of events. These functions I have called individuating functions. They may of course be only partial ones: a well-defined individual existing in one world may fail to exist in another. (HINTIKKA, 1970, p. 412)

Un individu n'est donc pas la résultante d'une individuation métaphysique mais d'une identification entre des scénarios différents représentés comme des mondes possibles. Dans la sémantique d'Hintikka, les domaines sont relativisés aux mondes possibles.

La sémantique d'Hintikka a donc pour vocation de maintenir aussi fermement que possible le rôle cognitif que joue le mode de présentation. En ce sens, la ressemblance ne peut donc être désolidarisée, comme le font Kripke et Putnam, de la représentation. L'individuation relève chez Hintikka de deux types de processus, d'une part une identification par stipulation, publique basée sur les informations qui font consensus, et d'autre part une identification perspectiviste, basée sur une relation d'accointance d'un agent à un objet [41]. Ces deux modes peuvent intervenir de manière concomitante à l'intérieur d'un même processus d'identification.

L'avantage de cette théorie vient du fait que dédoubler les modes de présentation d'un objet ne pèse pas bien lourd dans l'ontologie.

2.3 La question des indiscernables

Dans le monde, de nombreuses choses se ressemblent, comme par exemple deux feuilles d'une glycine ou deux galets d'un rivage. Certaines se ressemblent même très fortement, comme deux

[41] A ce sujet, voir aussi (LEWIS, 1983b).

stylos élaborés sur une même chaîne de production industrielle ou comme deux clones. Mais que serait une ressemblance qui n'exemplifierait pas la moindre différence entre les occurrences comparées ? Que seraient des choses *parfaitement indiscernables* [42] ? Il peut être tentant d'affirmer qu'en fait de ressemblance exacte, nous aurions plus vraisemblablement affaire à l'identité. Partant, nous n'aurions pas *deux choses répétées*, mais une seule, *ré-identifiée*.

La possibilité d'indiscernables, autrement dit, la possibilité de choses exactement ressemblantes mais numériquement différentes, est contestée par le principe dit « d'identité des indiscernables ». Quoique ce principe soit historiquement attaché à la figure de Leibniz [43], il fut défendu par d'autres philosophes avant lui et occupe une place importante dans les débats contemporains de métaphysique. Pour en donner un premier aperçu, disons que ce principe stipule que si x et y partagent exactement les mêmes propriétés, alors x est identique à y. Exprimé de manière contraposée, le PIdI revient à dire que deux particuliers différents se différencient toujours par quelque propriété. Le principe rabat ainsi la ressemblance exacte sur l'identité.

2.3.1 Pourquoi devrions-nous rejeter la possibilité d'indiscernables ?

La thèse de discernabilité nécessaire chez Platon

La question de la ressemblance parfaite est vue, chez Platon, comme une question d'imitation, c'est-à-dire en relation avec la notion d'originalité et d'authenticité. C'est ainsi qu'elle apparaît par exemple dans le *Cratyle* (PLATON, 1959) où Platon analyse cette notion à partir de celle de rectitude, relation entre les choses et leur dénomination. La problématique centrale du *Cratyle* est en effet de déterminer si la relation de dénomination est naturelle, comme le prétend Cratyle ou seulement conventionnelle comme le considère Hermogène. La rectitude entre un mot et la chose qu'il désigne procède-t-elle de nécessités naturelles, ou relève-t-elle au contraire de conventions linguistiques et sociales ? Platon considère que la dénotation est une forme d'imitation, que le nom imite la chose comme le portrait imite ce qu'il dépeint. C'est donc une théorie picturale du langage qui est présentée ici. Le nom est une imitation d'une chose. Ce qui nomme témoigne de ce qui est.

Platon relève deux types de rectitude, l'exactitude et la fidélité. L'exactitude se rapporte à l'ordre et au dénombrement. Il s'agit d'une relation quantitative. Platon donne en exemple le cas de la rectitude orthographique : bien des erreurs sont possibles dans l'écriture d'un mot. Un mot n'a en revanche qu'une seule graphie correcte. Les raisons pour lesquelles un mot est mal écrit sont soit qu'il ne contient pas le bon nombre de lettres, soit que l'enchaînement de celles-ci n'est pas adéquat. Mais prenons garde que, lorsque nous disons improprement que « le mot est mal écrit », nous ne parlions pas en réalité du mot lui-même, qui serait mal orthographié. Car mal écrire un mot, ce n'est justement pas écrire le mot en question. Ce n'est pas le *même* mot, mal écrit. C'est plutôt un *autre* mot, sinon une vaine suite de lettres.

En quelque sorte, Platon applique ici la contraposée du principe d'indiscernabilité des identiques : si le mot m_1 est dissemblable du mot m_2, alors m_1 est numériquement différent de m_2.

[42] Dans tout ce qui suit, j'emploie de manière synonyme les adverbes « parfaitement » et « exactement » de même que les adjectifs « parfait » et « exact ».

[43] Pour une étude détaillée de la place de l'identité des indiscernables dans la philosophie de Leibniz, voir (RODRIGUEZ-PEREYRA, 2014).

Ainsi, s'il y a bien des façons de mal écrire un mot, toutes ces manières consistent à écrire, en fait, d'autres mots, et de là, dit Platon, de signifier d'autres choses. Platon dit que dans ce genre de situation, la rectitude est mise en défaut parce que quelque chose est retranché ou ajouté, que le mot est pris en défaut ou en excès, ou parce qu'il subit un agencement inadéquat. Cette entorse faite à la relation entre ce qui est dit et ce qui est, Platon la désigne comme « défaut d'exactitude ». En revanche, il y a exactitude lorsque la relation est adéquate. De fait, on pourrait cependant objecter à Platon l'existence de synonymes.

L'exactitude pose aussi, et en retour, la question de l'identité des indiscernables : si un mot m_1 est indiscernable d'un mot m_2, a t-on nécessairement affaire à un seul mot ? Pourrait-il y avoir au contraire des mots indiscernables quoique n'ayant pas la même relation aux choses, autrement dit, pourrait-il y avoir des mots homonymes ? Platon transpose alors la question du mot à celle de l'image. Si deux images peuvent être indiscernables, alors deux mots peuvent sans doute l'être aussi, ce qui indiquerait certainement un indice en faveur de l'hypothèse conventionnaliste du langage. Car il semble que dans la théorie des idées de Platon, seule une relation univoque entre mots et choses permettrait de trancher en faveur de la conception naturelle du langage.

Nous pouvons copier plusieurs fois le même mot, ce qui induit que l'exactitude est une relation de reproduction à l'identique. Le mot n'imite pas la chose mais la désigne. Mais lorsqu'il s'agit d'image, la notion de rectitude ne peut plus coïncider sur celle d'exactitude, elle ne peut plus être une relation quantitative. Aussi, cette nouvelle instance de rectitude, que Platon appelle *fidélité*, concerne les situations où la relation ne se rapporte ni à l'ordre, ni au dénombrement. La fidélité est relative à la relation entre des images et les choses qu'elles reproduisent. La fidélité d'une image caractérise le fait qu'elle peut s'avérer suffisante à rendre reconnaissable la chose qu'elle représente, sans en épuiser la richesse qualitative. Mais jusqu'où peut aller la fidélité ? Y a-t-il une limite dans l'ordre du représentable ? La défaillance de l'image par rapport à son modèle est-elle seulement contingente, ou est-elle nécessaire ?

Pour répondre à cette question, Platon décrit l'expérience de pensée suivante : admettons qu'un dieu fasse une « reproduction » de Cratyle, un peu comme un peintre ferait son portrait, à la différence près que le dieu ne limiterait pas sa reproduction aux seules formes et couleurs, mais l'étendrait aux organes, aux pensées, aux humeurs, aux sensations, bref, à tout ce qui peut véritablement se dire de Cratyle. L'intervention d'une entité divine joue bien sûr un rôle important dans l'hypothèse, à savoir, celui de ne pas imputer l'inexactitude de la représentation à la défaillance de l'artisan. Il s'agit d'admettre ici que l'auteur de la copie soit entièrement apte à produire les gestes menant à une copie parfaite. La question est donc essentiellement centrée sur la possibilité métaphysique d'une telle reproduction, et non sur la capacité qu'aurait l'artisan de reproduire son modèle.

Selon Platon, il n'y aurait pas une image parfaitement fidèle de Cratyle en plus de son modèle orignal. La reproduction exacte d'une chose ne pourrait pas être sa représentation comme le serait une statue. Il s'agirait pas d'une duplique, d'une copie ou d'un double. Il y aurait en fait là « deux fois Cratyle », c'est-à-dire deux fois une seule et même chose, et non « deux Cratyles », autrement dit il n'y aurait pas deux choses indiscernables, mais une seule, deux fois présente. Ce qui semble parfaitement absurde. Pour Platon, l'impossibilité que l'image reproduise complètement la chose n'est donc pas une incapacité technique, mais une *impossibilité logique* qui condamne toute production à ne jamais pouvoir doubler exactement la chose qu'elle représente. Deux choses

peuvent certes être ressemblantes, mais *deux fois la même chose* est une absurdité. On voit que Platon donne bon droit au principe d'identité des indiscernables, qu'il utilise pour conclure : si a est indiscernable de b, alors a est b. Comme l'écrit Vincent Descombes, « le dieu lui-même ne peut produire un second Cratyle, parce que le double est une entité impossible [...] Il y a ici un théorème présupposé » (DESCOMBES, 1971, p. 46). Ce théorème, c'est bien sûr le principe d'identité des indiscernables.

En conséquence, la reproduction d'une chose par les moyens de la représentation est toujours, par nature, limitée. L'image peut se satisfaire, pour parvenir à représenter, d'une ressemblance imparfaite. Entre l'image et la chose, nombre de détails diffèrent sans que cela impose que la première cesse de représenter la deuxième. Plus encore, l'essence de l'image est de ne pas être exactement la chose dont elle est l'image, sans quoi elle ne serait plus une image, mais à la fois la même chose et à la fois autre chose, ce que rejette Platon. L'image est donc toujours en défaut par rapport à ce dont elle est image. Elle est toujours en relation d'approximation, et jamais en relation d'exactitude. Car si la rectitude exigeait la ressemblance parfaite, le monde serait alors rempli de doubles :

> Il y aurait à coup sûr de quoi rire dans la situation où les noms mettraient les choses dont ces noms sont les noms s'ils devaient en tout et de tout point leur être semblables : elles deviendraient toutes en effet, je pense, des doubles, et d'aucune d'elles il n'y aurait moyen de dire lequel de ces doubles est la chose elle-même et lequel est son nom. (PLATON, 1955, p. 678-679)

Ce que Platon reproche à l'imitation c'est de relever de l'ordre du simulacre, autrement dit, d'être une forme d'usurpation, une manière de se tenir à la place d'autre chose sans parvenir néanmoins à épuiser complètement la richesse constitutive de la chose remplacée. L'hypothèse de non réplication est, selon Descombes, une règle fondamentale du platonisme :

> Tout ce qui peut être institué en modèle pour une imitation est doué d'unicité. (DESCOMBES, 1971, p. 46)

L'indiscernabilité implique chez Platon l'identité, autrement dit, il est nécessaire que les choses diffèrent par quelque qualité afin de différer numériquement.

> De même que [...] les dissemblances se trouvent être semblables par leur dissemblance même, en ce qu'elle leur est commune, de même [...] les semblables se révèlent dissemblables par leurs ressemblances. (DESCOMBES, 1971, p. 48)

Selon Vincent Descombes, le refus d'admettre la possibilité de duplication de Cratyle indiquerait que le platonisme n'est pas une philosophie du double, mais au contraire, une philosophie de la singularité. Or la singularité implique une hiérarchie, et de fait, le platonisme est une philosophie axiologiquement hiérarchisée selon l'axe ontologique allant du modèle à son image et de celle-ci à son image, voire aux images fausses, ou simulacres, qui prétendent se substituer aux choses elles-mêmes.

Contrairement à Descombes, Clément Rosset insiste au contraire sur l'idée que la singularité intrinsèque des choses réelles chez Platon n'implique pas nécessairement une philosophie de

l'unicité. Car il faut distinguer entre deux niveaux de duplication, le niveau sensible et le niveau métaphysique. Alors que la thèse que les choses sensibles ne peuvent être dupliquées est exposée dans le *Cratyle*, cet argument doit être complété par une deuxième thèse, présentée dans le *Parménide* et qui affirme que les objets sensibles ne peuvent apparaître comme les dupliques de leurs modèles réels, suprasensibles. Ainsi, nous avons deux thèses. La première, positive et exposée dans le *Cratyle*, affirme le caractère foncièrement non-répétable des choses du monde. Mais ces choses ont, selon le jugement de Platon, un poids ontologique faible. La deuxième, négative et exposée dans le *Parménide*, nie la reproductibilité des choses que Platon conçoit comme véritablement réelles, à savoir, les idées. Le monde sensible est fustigé non parce qu'il est une duplique, mais justement parce qu'il est incapable d'être une duplique parfaite du monde réel, autrement dit, du monde intelligible. Au demeurant, il n'en est pas même une image. Il s'agit au mieux d'un simulacre, de jeux d'ombres projetées ou de peintures rupestres délinéant péniblement les contours d'objets aux dimensions véritablement plus complexes. D'où la condamnation, paradoxale, que Platon lance au monde sensible et que résume ainsi Clément Rosset :

> Le caractère non duplicable de la réalité aboutit à une dépréciation de l'objet sensible, auquel il est précisément reproché de ne pouvoir être le double, ni de lui-même en tant que sensible, ni de de l'autre en tant que réalité primordiale. Ce qui signifie que l'impossibilité du double vient paradoxalement démontrer que ce monde-ci n'est justement qu'un double, ou plus précisément un mauvais double, une duplication falsifiée, incapable de redonner ni l'autre, ni elle-même, bref, une réalité apparente, entièrement tissée dans l'étoffe d'un moindre « être » qui est à l'être ce que le succédané est au produit véritable. (ROSSET, 1979, p. 60)

Arguments pour le PIdI

Nous pouvons mentionner quelques arguments qui plaident en faveur du PIdI.

Premièrement, à l'échelle des individus macroscopiques et microscopiques, même pour les plus similaires d'entre eux, nous ne cessons, à l'observation patiente, de constater quelque dissemblance liée à leur structure complexe. L'examen attentif de deux stylos du même modèle révélera inexorablement certaines différences qualitatives. Ce constat pourrait nous pousser à induire que la complexité des entités matérielles leur sert de rempart contre la possibilité d'être des copies parfaites et à convenir, avec Leibniz, qu'« il n'y a jamais dans la nature, deux êtres qui soient parfaitement l'un comme l'autre » (Gottfried Wilhelm LEIBNIZ, 1995, p. 96).

Dans une lettre à Samuel Clarke, Leibniz écrit que :

> There is no such thing as two individuals indiscernible from each other. An ingenious gentleman of my acquaintance, discoursing with me in the presence of His Electoral Highness, The Princess Sophia, in the garden of Herrenhausen, thought he could find two leaves perfectly alike. The Princess defied him to do it, and he ran all over the garden a long time to look for some; but it was to no purpose. (G. W. v. LEIBNIZ & CLARKE, 2000, p. 22)

On notera que Leibniz n'hésite pas à appliquer le principe d'identité des indiscernables à des objets de différentes tailles, allant des objets les plus communs en passant par ceux qui sont difficilement observables à l'œil nu, jusqu'aux plus volumineux, voire à la totalité de l'univers.

Le PIdI est vrai pour les objets les plus petits, comme les gouttes de lait, qui, en dépit d'une forte similitude entre elles, révéleront toutefois, au microscope, de manifestes différences. Ce qui nous fait croire en leur parfaite similarité n'a donc pour raison qu'une observation trop hâtive ou imparfaite. En remédiant à la déficience de nos sens par une attention prolongée ou par l'usage d'appareils idoines, la dissimilitude se manifestera à notre observation.

> To suppose two things indiscernible is to suppose the same thing under two names. And therefore, to suppose that the universe could have had at first another position of time and place than that which it actually had, and yet that all the parts of the universe should have had the same situation among themselves as that with they actually had, such a supposition, I say, is an impossible fiction. (G. W. v. LEIBNIZ & CLARKE, 2000, pp. 22-23)

Deuxièmement, le PIdI permet de fournir une explication simple à l'individuation des particuliers, c'est-à-dire à la question de déterminer en quoi une chose donnée est différente de toute autre. Rappelons qu'en métaphysique, il existe principalement deux théories concurrentes de l'individuation. La première affirme que les particuliers sont des porteurs de propriétés, mais qu'ils ne s'y réduisent pas (je laisse pour l'instant de côté ce que l'on doit précisément entendre par « propriété », j'y reviendrai plus tard). Un *substratum* ou « particulier nu » est donc requis pour supporter les propriétés. Il s'agit d'une sorte de résidu non qualitatif qui subsisterait si l'on pouvait retirer à l'individu toutes ses propriétés. Cette première théorie est indépendante du PIdI.

La seconde théorie de l'individuation soutient que les particuliers ne sont rien d'autre que des faisceaux de propriétés. Par comparaison avec la théorie précédente, cette conception a pour mérite d'éviter l'hypothèse des particuliers nus, et donc de fonder l'individualité exclusivement sur des faits qualitatifs, autrement dit, accessibles de manière empirique. Néanmoins, cette théorie est tributaire de l'acceptation du PIdI. En effet, le fait d'instancier exactement les mêmes propriétés implique l'identité numérique des faisceaux, autrement dit, des individus : un individu a est identique à un individu b si et seulement si a instancie exactement le même faisceau de propriétés que b, ce qui n'est rien d'autre qu'une reformulation du PIdI et de sa réciproque. En conséquence, le rejet du PIdI implique inévitablement le rejet de la théorie du faisceau [44], autrement dit, l'abandon d'une théorie ontologiquement moins coûteuse que celle du *substratum* [45].

Troisièmement, la ressemblance exacte semble satisfaire deux caractéristiques formelles importantes de l'identité qui ne sont pas remplies par la ressemblance partielle. Cela pourrait inciter à identifier la ressemblance exacte avec l'identité en donnant ainsi raison au PIdI. La première caractéristique formelle est la transitivité. En effet, si a ressemble exactement à b et si b ressemble exactement à c, alors évidemment il est aussi le cas que a ressemble exactement à c. En revanche,

[44] Pour un point de vue divergent sur la question, voir (RODRIGUEZ-PEREYRA, 2004).

[45] Pour Frédéric Nef, ces deux solutions semblent aussi désespérées l'une que l'autre : « Individuer les particuliers par des substrats particuliers dénués de propriétés et donc de structure est une tentative désespérée et vouée à l'échec. Tout ce que nous savons de l'individuation des contenus dans la science, la perception ou le langage va contre l'idée d'une individuation par un substrat nu. Cependant si nous admettons que les objets sont des réunions éphémères, contingentes et irrégulières de propriétés particulières existant indépendamment, nous nous exposons à une difficulté aussi insurmontable que la précédente. Le prix d'une telle doctrine est extraordinairement élevé. Comment distinguer les réunions consistantes des inconsistantes ? Les complètement impossibles des relativement impossibles ? » (NEF, 2006). Nef propose une conception intermédiaire visant à considérer les individus comme des faisceaux de propriétés particularisées, ou tropes.

ce n'est généralement pas vrai pour la ressemblance partielle.

Trois régimes de doubles

En dépit des arguments précédents, les discussions menées jusqu'à présent autour du PIdI s'avèrent pourtant non concluantes. Le PIdI semble doté d'une certaine plasticité qui lui permet, certes, de toujours échapper aux attaques, mais qui, en même temps, ne l'en dégage jamais complètement. Pour mettre à jour les articulations qui justifient une telle souplesse dialectique, il vaut la peine de passer par une interprétation plus formelle du PIdI. En logique du second ordre, autrement dit, en admettant la quantification sur les propriétés, le principe peut être formalisé de la manière suivante :

$$\textbf{[PIdI]}: \ \forall x \, \forall y \, [\forall \mathrm{F}(\mathrm{F}x \leftrightarrow \mathrm{F}y) \rightarrow (x = y)]$$

Le principe est également utilisé sous sa forme contraposée, baptisée « principe de dissemblance du divers ». Sous cette seconde formulation, il affirme que des objets numériquement distincts diffèrent par quelque propriété. Plus formellement :

$$\textbf{[PDisD]}: \ \forall x \, \forall y \, [(x \neq y) \rightarrow \exists \mathrm{F} \, (\mathrm{F}x \wedge \neg \mathrm{F}y) \vee (\neg \mathrm{F}x \wedge \mathrm{F}y)]$$

Notons tout de suite deux caractéristiques qui ressortent de cette formulation. Premièrement, on peut se poser la question du domaine de quantification des variables x et y. Si nous demandons s'il peut exister des indiscernables, un premier embarras vient de ce que nous entendons au juste par « exister ». Si l'existence d'entités comme les cailloux, les rhododendrons, les charançons et les tasses à café ne semble pas trop problématique, en revanche, celle des cercles, des nombres, des personnages de fiction, des ombres et des trous soulève sans doute plus de questions. Je me cantonnerai toutefois à une conception de l'existence telle que l'entend Quine quand il écrit qu' « une théorie est engagée pour les entités, et celles-là seules, que ses variables liées doivent avoir comme références possibles pour que les affirmations faites dans la théorie soient vraies » (W. V. O. QUINE, 2003b, p. 41). Précisément, dans ce cadre, les théories en question seront celles des sciences de la nature.

La deuxième caractéristique qui ressort de cette formalisation du PIdI est le domaine de quantification des prédicats. Jusqu'à présent, nous avons fait comme si la nature des propriétés allait de soit, mais c'est loin d'être le cas. De quelles sortes de propriétés est tributaire le PIdI ? Doit-on restreindre la quantification à certains prédicats ? Le cas échéant, lesquels ? Selon une conception que David Lewis appelle « abondante » des propriétés, n'importe quel prédicat représente une propriété (LEWIS, 1983c), comme « être à moins d'un kilomètre d'un marchand de glace », ou comme « être un nombre premier ou ne pas être la Tour Eiffel ». La profusion de telles propriétés engendre une conséquence importante pour l'évaluation des ressemblances : pour deux choses données, il y a une infinité de propriétés qu'elles ne partagent pas. Dès lors, des entités numériquement distinctes sont toujours discernables par une infinité de propriétés et le PIdI se trouve ainsi trivialisé. C'est ce qui fait dire à Lewis que « [...] les propriétés ne font rien pour capturer des faits de ressemblance » (LEWIS, 1983c, p. 346). En revanche, selon la conception dite « rare » des propriétés que défend David Armstrong, tout prédicat ne représente pas nécessairement une propriété. Un prédicat exprime une propriété seulement s'il dénote un universel, autrement dit, une entité qui

entre dans la composition d'un particulier et qui est susceptible d'être simultanément instanciée à divers emplacements [46]. Ce qui doit ou ne doit pas être admis comme un universel est arbitré principalement par les théories scientifiques.

La validité PIdI doit donc être discutée selon le domaine de quantification des propriétés servant à la comparaison des particuliers. En principe, les auteurs distinguent trois versions du principe : triviale, modérée et faible :

- **[PIdI$_{\text{triviale}}$]** :
  $\forall x\ \forall y\ \Big(\forall$F où F est une propriété quelconque $(\mathrm{F}x \leftrightarrow \mathrm{F}y) \rightarrow (x = y)\Big)$
- **[PIdI$_{\text{modéré}}$]** :
  $\forall x\ \forall y\ \Big(\forall$F où F est une propriété pure $(\mathrm{F}x \leftrightarrow \mathrm{F}y) \rightarrow (x = y)\Big)$
- **[PIdI$_{\text{faible}}$]** :
  $\forall x\ \forall y\ \Big(\forall$F où F est une propriété pure intrinsèque $(\mathrm{F}x \leftrightarrow \mathrm{F}y) \rightarrow (x = y)\Big)$

On peut remarquer les restrictions croissantes imposées aux propriétés : admettre le principe faible implique d'admettre le principe modéré et le principe trivial sans que la réciproque soit vraie. Et de manière contraposée, infirmer le principe modéré implique d'infirmer le principe faible.

Passons en revue ces trois versions. La première version est triviale puisqu'il s'agit d'un théorème de logique de second ordre. En effet, on peut toujours faire appel à un prédicat du type « être identique à a » pour distinguer a de n'importe quelle autre chose. Mais de tels prédicats peuvent-ils être envisagés comme d'authentiques propriétés ? Cela consisterait à stipuler une identité primitive, ou une essence individuelle. Comme nous l'avons vu, l'existence de telles propriétés individuantes est cependant contestable, et c'est pourquoi la version modérée du principe exclut les propriétés qui ne sont pas *purement qualitatives*.

La version faible du principe ne restreint pas le domaine de quantification aux seules propriétés pures, mais plus particulièrement aux propriétés dites *intrinsèques*. Pour en saisir l'intérêt, rappelons une expérience de pensée inventée par David Lewis (LEWIS, 1983c, p. 355). Imaginons une photocopieuse parfaite, capable de reproduire un imprimé à la particule quantique près. La duplique d'un document serait, du point de vue de ses qualités, absolument indiscernable de l'original. Mais nous pourrions encore distinguer les deux documents par d'autres propriétés, dites « extrinsèques », comme le fait d'occuper des emplacements différents. Les propriétés intrinsèques se distinguent donc des propriétés extrinsèques en ce que les premières, contrairement aux secondes, ne dépendent d'aucune autre chose. Par exemple, « être courbé » exprime une propriété intrinsèque, contrairement à « être à gauche d'une fenêtre » qui exprime une propriété extrinsèque.

Ces trois versions du principe d'identité des indiscernables permettent maintenant d'affiner notre terminologie en distinguant plusieurs sortes d'indiscernables. Aussi nous dirons dorénavant que :

[46] Dans la littérature philosophique, d'autres définitions des propriétés existent : elles sont parfois conçues à l'aide de classes ou à partir d'une relation primitive de ressemblance, voire identifiées à particuliers abstraits (tropes) (David M. ARMSTRONG, 2010, p. 77 sqq).

- un objet x est une *réplique pure* d'un objet y si et seulement si x et y contredisent le principe faible d'identité des indiscernables, mais pas le principe modéré, à savoir, si et seulement si x et y partagent toutes leurs qualités intrinsèques pures mais pas toutes leurs qualités pures extrinsèques, sans être identiques.

- un objet x est un *double pur* d'un objet y si et seulement si x et y contredisent le principe modéré d'identité des indiscernables - autrement dit si x et y partagent à la fois toutes leurs qualités pures intrinsèques et extrinsèques sans être identiques - et ne sont pas co-localisés.

- un objet x est un *coïncident pur* d'un objet y si et seulement si x et y contredisent le principe modéré d'identité des indiscernables et sont co-localisés ;

- on conservera au terme d'*indiscernables* un caractère neutre par rapport à la version du principe d'identité des indiscernables qui est contredite.

Dans les sections qui suivent, nous allons examiner, tour à tour, la possibilité des répliques pures, des doubles purs et des coïncidents purs.

2.3.2 Le cas des répliques pures

Existe-t-il des choses qui soient des répliques pures ? Et si tel n'est pas actuellement le cas, pourrait-il ou aurait-il pu en exister ?

Les particuliers nus sont-ils des répliques pures ?

La théorie du faisceau exclut par hypothèse le cas de répliques pures. Au contraire, celle du *substratum* ne les écarte pas d'emblée, et parfois même, cette théorie est choisie justement parce qu'elle les admet. Les particuliers nus servant à justifier la différence numérique entre des objets parfaitement similaires, les ensembles de propriétés qu'ils instancient peuvent en effet s'avérer strictement identiques sans que se pose un problème d'individuation.

Nous pourrions alors nous demander si, dans la perspective de la théorie du *substratum*, les particuliers nus ne relèveraient pas du cas des répliques pures. La réponse, négative, pourrait paraître triviale. Les particulier nus sont justement trop nus pour être des répliques. On admet généralement en effet qu'un particulier nu est, par définition, sans propriété, qu'il est une sorte de résidu obtenu après avoir retiré toutes les propriétés intrinsèques et extrinsèques au particulier « épais », c'est-à-dire à l'objet, qui les instancie. Comme l'écrit Gustav Bergmann, qui a relancé les discussions autour de telles entités dans les écrits de métaphysiques après leur exclusion par Bertrand Russell, « Bare particulars neither are or nor have natures. Any two of them are not intinsically but only numerically different. That is their bareness. » (BERGMANN, 1967, p. 24). Du point de vue de Bergmann, les particuliers nus sont donc entièrement déterminés par les deux caractéristiques suivantes : (1) ils ne diffèrent pas, entre eux, de manière intrinsèque, et (2) ils ne peuvent pas être présents dans plus d'une substance. En ce sens, puisque les particuliers nus n'instancient pas de propriétés intrinsèques, ils ne relèvent donc pas du cas des répliques pures.

Certains auteurs, dont Sellars, ont objecté qu'il n'est pas possible qu'un particulier nu n'instancie aucune propriété, parce que le fait de n'instancier aucune propriété constitue déjà une propriété. La réponse de savoir si des particuliers sont ou non des répliques pures ne serait donc pas si triviale que cela. Mais de toute façon, souligne Sellars, il serait absurde d'envisager l'existence de telles entités, car celles-ci seraient, de manière contradictoire, à la fois munies et démunies de propriétés :

> Perhaps the neatest way in which to expose the absurdity of the notion of bare particulars is to show that the sentence 'Universals are exemplified by bare particulars' is a self- contradiction. As a matter of fact, the self-contradictory character of this sentence becomes self-evident [...] 'If a particular exemplifies a universal, then there is no universal it exemplifies'. (SELLARS, 1952, p. 184)

Comme nous l'avons vu, les partisans de la théorie du substrat répondent habituellement que les prédicats ne réfèrent pas tous à d'authentiques propriétés, que seuls certains, dits « naturels » ou qualitativement purs renvoient à de vénérables universaux. De ce point de vue, les particuliers nus ne sont donc non pas sans propriété *simpliciter*, ils sont sans propriété *naturelle*. Or, les propriétés qui comptent pour parler de répliques pures sont les propriétés naturelles, qui peuvent être partagées, contrairement aux propriétés haeccéitiques comme « être identique à Socrate ». En conséquence, nous devrions bien répondre que les particuliers nus ne sont pas des répliques pures.

Théodore Sider soutient cependant que le fait de ne pas instancier une propriété naturelle constitue bien une authentique propriété naturelle. Car, argumente Sider, avoir une nature n'exige en rien d'instancier au moins une propriété intrinsèque : « a thing could have a nature simply by failing to instantiate monadic universals. » (SIDER, 2006, p. 8). Si l'on accepte ceci, il faut concéder alors que les particuliers nus instancient entre eux exactement une seule propriété naturelle, et chacun la même, à savoir celle de ne pas instancier de propriété naturelle. Nous aurions ainsi affaire à d'authentiques répliques pures :

> On an intuitive level: to have a nature is to "be a certain way". There must be answers to questions like "what is the thing like?", and "to what is the thing similar, and to what is it dissimilar?" Truly bare particulars do have natures in this intuitive sense. Indeed, they all have the same nature, and that nature is exhausted by the fact that they instantiate no monadic universals. (SIDER, 2006, p. 9)

Selon Sider, nous devrions même en déduire l'actualité des particuliers nus, et pas seulement leur pure possibilité. Les particuliers nus étant pensés comme des entités fondamentales, ils sont indépendants des universaux qu'ils instancient, et de ce fait, ils existent sans eux. Mais, s'ils sont actuels, où les trouve-t-on dans le monde ? Sider soutient que la science laisse à penser qu'il existe de telles entités, à savoir, les points spatio-temporels :

> I have in mind points of spacetime... What are the distinctive intrinsic features of points of spacetime? If we look at science for guidance, we find that physical theories require almost nothing of the points intrinsically. I suggest, then, that a natural and economic theory of points of spacetime is that each one is a partless, truly bare particular that stands in a network of spatiotemporal relations. (SIDER, 2006, p. 10)

De fait, dans sa discussion autour des conceptions absolutiste et relativiste de l'espace, John Earman souligne que parmi les caractéristiques que Newton confère à la notion d'un espace absolu figure celle de substance, en ce que l'espace, « forms a substratum that underlies physical events and processes, and spatiotemporal relations among such events and processes are parasitic on the spatiotemporal relations inherent in the substratum of space-time points and regions » (EARMAN, 1989, p. 11). Les particuliers surviennent donc sur les régions de l'espace qui constituent la substance fondamentale du monde. Mais cela n'implique pas que les points spatio-temporel soient des répliques pures. Pour cela, l'espace doit être composé d'une multiplicité de points spatio-temporels qui possèdent chacun la propriété d'être, de manière absolue, différent de tous les autres, autrement dit, une haeccéité. Une telle qualité étant très suspecte, une autre façon d'envisager cette multiplicité serait d'admettre que les points spatio-temporels sont des particuliers nus capables d'instancier des propriétés, mais on obtient alors une conception de l'espace peu intuitive [47] dans laquelle les particuliers sont vus comme des objets hybrides mêlant régions physiques et propriétés qualitatives plutôt que comme des objets localisés dans l'espace par des relations géométriques mutuelles.

Quoi qu'il en soit, nous avons vu que les arguments qu'utilise Sider pour montrer que les particuliers nus sont des points spatio-temporels assimilés à des répliques pures reposent sur l'hypothèse que le prédicat de ne pas instancier de propriété naturelle exprime bien une authentique propriété naturelle. Cette hypothèse semble toutefois peu convaincante, elle est, en tout cas, loin de faire l'unanimité, même parmi les partisans de la théorie du *substratum*. De manière plus générale, le fait qu'un partisan de la théorie du *substratum* débouche sur l'idée de particuliers nus vus comme des répliques pures indique peut-être que de telles entités sont, avec les haeccéités et les tropes, plus des « fictions métaphysiques extravagantes » (J. LOWE, 2003, p. 88) destinées à résoudre des problèmes d'individuation ardus que des entités réelles. Et de telles inventions posent sans doute, au final, plus de problèmes qu'elles n'en résolvent : « We do metaphysics and ourselves no service by inventing mysterious entities to solve deep ontological problems, such as the problems of individuation" (J. LOWE, 2003, p. 83). Le but de Sider est de contrer cette accusation en rapprochant les particuliers nus des points spatio-temporel pour légitimer l'hypothèse de leur existence, ou du moins, pour en rendre l'idée moins fantasque. Mais assimiler ainsi les particuliers nus à des répliques pures semble mener, du point de vue d'un partisan de la théorie du *substratum*, à une stratégie contre-productive. En effet, dans les discussions des réalistes autour des critères d'individuation, l'hypothèse d'un particulier nu sert habituellement de rempart pour contrebalancer le rôle uniformisant des universaux. Il est donc assez étrange, pour un partisan de cette doctrine, de considérer que les particuliers nus sont des répliques pures les uns des autres. Car entendus comme des répliques, comment pourraient-ils encore avoir un effet particularisant ?

Si le réaliste souhaite donner crédit à l'idée de particuliers nus, il doit donc admettre que de telles entités ne pourraient pas, par définition, être des répliques pures les unes des autres, sans quoi elles perdraient inévitablement leur effet individuant.

[47] Peu intuitive, mais sans doute proche de la conception de la théorie de la relativité qui associe matière et espace.

Les emplacements spatiaux sont-ils des répliques pures ?

Indépendamment des l'existence des particuliers nus, la discussion autour de l'argument de Sider a pour mérite de mettre en avant la question de savoir si les points spatio-temporels constituent ou non des cas de répliques pures. Car des points pourraient ne pas être des particuliers nus, même si, comme le reconnaît Michael Loux, l'hypothèse d'un espace absolu composé d'une multiplicité de points spatio-temporels est de toute manière très proche de la question de particuliers nus : « The points and moments of the absolute theory are, however, ontologically on a par with bare particulars » (LOUX, 1978, p. 138).

De fait, un principe fondamental gouverne nos pratiques de dénombrements : quand un emplacement est occupé par une chose, nous savons alors qu'il ne faut la compter qu'une seule fois. L'enfant qui compte ses billes désigne du doigt les billes en indiquant leur emplacement. Un emplacement de bille compte pour une bille et une seule. Il n'y a pas de billes superposées. Selon Locke, ce principe d'impénétrabilité peut être énoncé ainsi :

(Principe d'impénétrabilité) Deux choses de la même sorte ne peuvent occuper synchroniquement le même emplacement.

Cette affirmation pourrait laisser supposer que les entités physiques ne sont pas seulement identifiées par leur emplacement spatio-temporel, mais, au sens métaphysique du terme, qu'elles sont même individuées par leurs emplacements, ce qui engage alors l'idée que les particuliers surviennent sur les emplacements. Admettons par exemple l'hypothèse d'un espace absolu qui laisse la possibilité d'univers parfaitement ressemblants par les propriétés intrinsèques, mais distincts par les emplacements. Cette idée est, rappelons-le, absolument invraisemblable aux yeux de Leibniz. Comment pourrait-on admettre qu'un univers soit numériquement différent d'un autre dans lequel tous les individus auraient été translatés de 1 cm selon une direction donnée ? Selon Leibniz, il ne s'agirait pas d'une différence réelle, mais seulement d'une différence nominale : nous parlerions de façons différentes d'un seul et même univers :

> Supposer deux choses indiscernables, c'est supposer la même chose sous deux noms. Ainsi, supposer que l'univers aurait pu avoir, au commencement, une position temporelle et spatiale différente de celle qu'il a actuellement, et ainsi que toutes les parties de l'univers auraient pu avoir, entre elles, une situation identique à celle qu'elles ont actuellement, une telle supposition, dis-je, est une fiction impossible. (G. W. v. LEIBNIZ & CLARKE, 2000, pp. 22-23)

Dans la *Critique de la Raison Pure*, Kant, qui opte pour le caractère absolu de l'espace, reprend à son compte le principe lockéen d'impénétrabilité. Kant souligne que les emplacements permettent toujours de différencier numériquement des individus qui seraient par ailleurs qualitativement indiscernables. Deux gouttes d'eau, même indiscernables, sont toujours diversement localisées dans l'espace : « La diversité des lieux, sans autre condition, rend la multiplicité et la distinction des objets comme phénomènes non seulement possibles par elles-mêmes, mais aussi nécessaires » (KANT, 2012, p. 238-239). Puisque la diversité des lieux est une condition à la fois suffisante et nécessaire à la diversité numérique, il n'y a pas de contradiction à ce que deux choses soient à la fois indiscernables du point de vue de leurs propriétés intrinsèques et à la fois numériquement différentes en raison de leur dispersion spatiale (à condition d'admettre qu'occuper

un lieu donné n'est pas une propriété pure intrinsèque). Cette contestation du PIdI est soutenue en présupposant qu'« une partie de l'espace, bien que tout à fait semblable et égale à une autre, est cependant en dehors d'elle, et elle est précisément par là une partie distincte de la première » (KANT, 2012, p. 234). De ce point de vue, puisqu'elles sont d'emblée individuées, deux parties non superposées d'espace constituent déjà des cas de répliques pures.

Dans les *Prolégomènes à toute métaphysique future*, Kant avance un deuxième argument en faveur des répliques pures [48], à savoir, l'existence de couples parfaitement énantiomorphes d'objets [49].

> Que peut-il y avoir de plus semblable et de plus égal en tous points à ma main ou à mon oreille que leur image dans le miroir ? Et pourtant je ne puis substituer une main vue dans le miroir à son modèle ; car si c'est une main droite, dans le miroir c'est une main gauche. (KANT, 1986, p. 49)

L'indiscernabilité parfaite entre, disons, une main droite et une main gauche relève seulement de leurs propriétés intrinsèques. Elle ne couvre donc pas leur caractère énantiomorphe. Car entre deux mains chirales « il n'y a pas de différence pure qu'un entendement pourrait à lui seul, penser » (KANT, 1986, p. 49-50 §13). Autrement dit, la diversité numérique des deux mains ne résulte pas d'une différence conceptuelle, mais relationnelle. Ainsi, les objets parfaitement énantiomorphes fournissent un type particulièrement intéressant de répliques pures. De tels objets sont non seulement individués par leur dispersion spatiale (par translation et par rotation), mais également par leur orientation dans l'espace (par réflexion).

Mais l'hypothèse d'un espace absolu et orienté est sans doute contestable, et de toute façon, le contexte historique duquel provient le PIdI est tributaire de la conception leibnizienne d'un espace relatif [50]. Or, dans le cas de la conception relative de l'espace, on voit mal quelle propriété intrinsèque pourraient posséder les points spatio-temporels, puisqu'ils sont définis exclusivement de manière extrinsèque les uns par rapport aux autres. Dans ce cas, la question des répliques pures semble leur échapper complètement.

Y a-t-il des entités concrètes qui sont des répliques pures ?

Existe-t-il des entités concrètes qui soient de bons candidats au rôle de répliques pures ? Que penser par exemple des électrons, s'il existe de telles choses ? Du point de vue de la mécanique

[48] Les arguments usant d'énantiomorphes apparaissent en fait pour la première fois chez Kant dans un court texte de 1768 intitulé *Du premier fondement de la différence des régions dans l'espace* dont l'objet, écrit Kant, est de concevoir que « l'espace absolu a une réalité propre, indépendamment de l'existence de toute matière et même comme premier fondement de la possibilité de sa composition. » (KANT, 1970). Kant demande d'imaginer que la toute première création soit une main humaine. Un telle main serait, dit Kant, ou droite, ou gauche, ce qui récuse selon lui le caractère relatif de l'espace. L'espace est donc, selon Kant, orienté de manière intrinsèque. Remarquons au passage que le contre-exemple qui sert à la démonstration de Kant utilise une réflexion, autrement dit un retournement dans l'espace là où l'argument de Leibniz, considérant qu'un monde dont on aurait bougé tous les objets de 1 cm selon une même direction serait strictement le même monde, recourt quant à lui à une translation.

[49] Le terme se dit de chacun des éléments d'un couple de solides qui sont non superposables et dont l'un est l'image de l'autre dans un miroir. On dit aussi que de tels objets possèdent la propriété de *chiralité*.

[50] A ce sujet, voir la correspondance entre Leibniz et Clarke (G. W. v. LEIBNIZ & CLARKE, 2000).

classique les particules ne sont pas inter-pénétrables. Aussi, si une particule est localisée sur une partie d'espace, toute autre particule est située en dehors de cette partie, ce qui fournit un critère extrinsèque de distinction numérique entre particules par ailleurs intrinsèquement indiscernables. Nous serions donc en présence de répliques pures. Et il pourrait en être de même pour deux atomes d'hydrogène, voire pour deux molécules. Dans ce cas, à quel stade de complexité devrions-nous renoncer aux indiscernables ? Nous avons soupçonné les objets microscopiques d'être suffisamment complexes pour ne pas pouvoir produire des répliques pures. Du point de vue de la physique classique, cela pourrait néanmoins s'avérer faux. Il pourrait exister des répliques pures tant au niveau des entités les plus simples, qu'à l'échelle des ensembles les plus complexes comme les mondes, comme le font remarquer les partisans de mondes dupliqués. Nous avons vu dans la première section que l'atomisme épicurien ou la physique classique pouvait s'avérer favorable à l'idée de répétitions à l'identique d'entités simples et d'entités composées. Nous avons vu également que la théorie d'un univers éternel alternant une succession de Big Bang et de Big Crunch est compatible avec l'idée de répétitions temporelles d'un monde.

Dans *An Inquiry about Meaning and Truth*, Bertrand Russell insiste sur l'importance de la question des répliques pures pour sa propre théorie descriptiviste :

> This leads ultimately to such questions as: if there were in New York an Eiffel Tower exactly like the one in Paris, would there be two Eiffel Towers, or one Eiffel Tower in two places? If history repeated itself, would the world be in two exactly similar states on two different occasions, or would one and the same state occur twice, i.e., precede itself? The answers to such questions are only partly arbitrary; in any case, they are indispensable for the theory of names. (RUSSELL, 1940, pp. 96-96)

Russell considère que des cycles temporels indiscernables ne sont pas des mondes numériquement différents, mais des répétitions temporelles d'un seul et même monde :

> Many writers have imagined that history is cyclic, that the present state of the world, exactly as it is now, will sooner or later recur. [...] We shall have to say that the later state is numerically identical with the earlier state; and we cannot say that this state occurs twice, since that would imply a system of dating which the hypothesis makes impossible. (RUSSELL, 1940, p. 102)

Pour Leibniz, le PIdI est un principe métaphysique. Il y a toujours moyen de discerner deux choses depuis un point de vue. La stratégie leibnizienne de défense du PIdI s'appuie sur la discernabilité des particuliers. Chez Russell, l'hypothèse d'une répétition temporelle ou spatiale des individus relève d'une impossibilité logique. En effet, dans les *Principia Mathematica*, Russell définit l'identité logique directement par le PIdI :

> The prepositional function "x is identical with y" will be written "$x = y$" We shall find that this use of the sign of equality covers all the common uses of equality that occur in mathematics. The definition is as follows: $x = y := \forall\sigma, \sigma(x) \rightarrow \sigma(y)$. This definition states that x and y are to be called identical when every predicative function satisfied by x is also satisfied by y. We cannot state that every function satisfied by x is to be satisfied by y, because x satisfies functions of various orders, and these cannot all be covered by one apparent variable. But in

> virtue of the axiom of reducibility it follows that, if $x = y$ and x satisfies φx, where φ is any function, predicative or non-predicative, then y also satisfies φy. Hence in effect the definition is as powerful as it would be if it could be extended to cover all functions of x. (RUSSELL & WHITEHEAD, 1997, p. 13.01)

Commentant les théories de l'éternel retour, Russell soutient ainsi que, dans le cas de mondes aux répétitions cycliques, tout se passe comme lorsque, partant d'un point d'une trajectoire circulaire, on revient en bout de course au même endroit :

> The situation would be analogous to that of a man who travels round the world : he does not say that his starting-point and his point of arrival are two different but precisely similar places, he says they are the same place. (RUSSELL, 1940, p. 102)

On pourrait toutefois objecter que si, dans la comparaison de Russell, l'emplacement reste numériquement identique, l'instant temporel, lui, est bien différent.

Russell considère qu'un objet n'est qu'une agglomération de qualités, un « faisceau » de propriétés contemporaines, et il a donc tout intérêt à préserver le PIdI, puisque le rejet de ce principe implique aussi le rejet de la théorie des faisceaux. Ainsi, il écrit que :

> The hypothesis that history is cyclic can be expressed as follows : form the group of all qualities contemporaneous with a given quality ; in certain cases the whole of this group precedes itself. Or : in these cases, every group of simultaneous qualities, however large, precedes itself. Such an hypothesis cannot be regarded as logically impossible so long as we say that only qualities occur. To make it impossible, we should have to suppose a momentary subject of qualities, and to hold that this subject owes its identity, not to its character, but to its space-time position. (RUSSELL, 1940, p. 102)

Comme on le voit, la stratégie de Russell consiste à affirmer qu'un individu n'étant qu'un faisceau d'universaux ayant la capacité d'être localisés de manière synchronique ou diachronique à différents emplacements, un même faisceau peut se reproduire à une certaine distance de lui-même et à un moment ultérieur de lui-même. C'est, écrit Russell, l'idée qu'il y a une substance sous les qualités qui implique la contradiction logique, mais dans le cas de la théorie des faisceaux d'universaux, il n'y en aurait pas. De ce point de vue, c'est bien la *même* chose (au sens de l'identité numérique) qui revient entre deux cycles, et non des individus indiscernables et numériquement distincts. Russell considère donc que l'éternel retour ne peut avoir de sens que dans l'hypothèse de qualités regroupées. La totalité de cet assemblage de qualités constitue l'individu lui-même. Le retour de ces qualités impliquerait donc le retour du même individu. Ce qui reviendrait ne serait pas un ressemblant, mais, à strictement parler, le même individu. En revanche, admettre qu'un particulier individué par des conditions temporelles puisse se reproduire serait une absurdité.

Pour Russell, les universaux ne sont pas des répliques pures. Ils sont dotés de l'extraordinaire propriété de pouvoir être à une certaine distance d'eux-mêmes, avec d'autres propriétés extrinsèques aussi étranges qui sont pourtant habituellement réputées être non-réflexives :

> When I say that redness can be in two places at once, I mean that redness can have to itself one or more of those spatial relations which, according to common sense, no "thing" can have

> to itself. Redness may be to the right of redness, or above redness, in the immediate visual field ; redness may be in America and in Europe, in physical space. (RUSSELL, 1940, p. 98)

Par opposition aux universaux, les propriétés spatiales, que Russell considère comme des qualités, ne peuvent pas être à une certaine distance d'elles-mêmes :

> If θ, Φ, h are a latitude, a longitude, and an altitude, we shall find that the bundle (θ, Φ, h) cannot be north or south or east or west or above or below itself, as redness can. (RUSSELL, 1940)

Comme nous l'avons vu plus haut, en ouvrant le domaine de quantification aux entités stipulées par la physique classique, nous devrions sans doute conclure qu'il existe des répliques pures à toutes les échelles, des particules élémentaires aux galaxies en passant par les individus macroscopiques. Le jardin de Herrenhausen dont parle Leibniz est sans doute trop petit pour abriter deux feuilles exactement ressemblantes, mais dans un contexte laplacien et à l'échelle de l'univers, le réaliste des propriétés devrait au contraire être assuré de l'existence d'une infinité de feuilles intrinsèquement indiscernables.

2.3.3 Le cas des doubles purs

Passons maintenant à la question de l'existence de doubles purs. De telles choses pourraient-elles exister ?

On peut contester la validité de la physique classique et considérer alors que le PIdI est vrai de manière actuelle. Or pour accéder au statut d'authentique principe métaphysique, le PIdI ne devrait pas être vrai seulement de manière contingente, mais de manière nécessaire, c'est-à-dire qu'il devrait être vrai dans *tous* les mondes possibles.

La portée modale du PIdI

Dans un article intitulé « L'identité des indiscernables », Max Black prétend avoir décrit un monde possible contenant des doubles purs, c'est-à-dire un monde qui contredit la version modérée du PIdI. Mais depuis sa publication en 1952, l'article a fait couler beaucoup d'encre autour de la question très controversée de savoir ce qu'il prouve au juste. L'auteur y décrit un dialogue entre deux personnages où l'un défend le PIdI et l'autre le rejette. À un moment de l'échange, l'opposant au PIdI propose de considérer un univers exclusivement constitué de deux sphères exactement similaires de centres séparés d'une distance double de leur diamètre, comme illustré sur la figure 2.5. Black conclut que de telles sphères contrediraient la version modérée du PIdI. Si l'on en convient, cela implique que le PIdI est au mieux un principe contingent, mais pas un principe nécessaire. En effet, sa portée modale serait affectée par un monde dans lequel il serait invalidé.

Black estime qu'aucun argument ne permet d'échapper au constat de deux sphères indiscernables. On a trois voies pour rejeter l'indiscernabilité des sphères :

- admettre la distinction numérique des sphères en récusant leur indiscernabilité ;

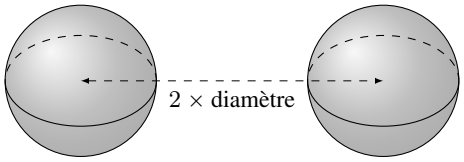


FIGURE 2.5 : Expérience de pensée des deux sphères indiscernables

- refuser la distinction numérique des sphères ;
- refuser d'individuer les sphères.

Examinons, les unes après les autres, ces trois alternatives.

Voie 1 : défense du PIdI par la discernabilité

Commençons par évaluer la résistance du scénario de Black aux objections de discernabilité. Parvenir à discerner les sphères, c'est déterminer au moins une propriété que l'une possède mais pas l'autre. Dans le dialogue, après quelques tentatives non concluantes visant à les discerner par divers moyens sémantiques dont nous ne rappellerons pas le détail, le défenseur du PIdI tente d'utiliser à son profit l'hypothèse de dispersion spatiale. Si les sphères sont séparées (nous dirons dorénavant « non-connexes »), dit-il en substance, c'est qu'elles ne sont pas indiscernables. Le raisonnement procède ainsi :

(1)	x et y sont non-connexes	hypothèse de non-connexité
(2)	si x et y sont non-connexes, alors x et y sont discernables	principe (?) de discernabilité des non-connexes
	———————————————	
(3)	x et y sont discernables	*Modus ponens* appliqué à (1) et (2)

Mais le putatif principe de discernabilité des non-connexes mobilisé ici ne semble pas valide. Car comme l'écrit Black, « dire que les sphères occupent différents emplacements, c'est ne rien dire de plus qu'elles sont deux sphères » (BLACK, 2012, p. 126). Précisément, la non-connexité ne permet pas de déterminer une propriété qui serait instanciée par l'une mais pas l'autre des deux sphères. Le défenseur du PIdI croit discerner les sphères en raison de leur dispersion spatiale, mais sa conclusion devrait vraisemblablement être re-dirigée vers la diversité numérique des sphères.

Une manière d'envisager la discernabilité par la position spatiale nécessite de souscrire à l'idée d'un espace absolu dans lequel les emplacements sont eux-mêmes individués. Stalnaker écrit par exemple qu'« être localisé à un emplacement particulier serait une propriété qualitative (intrinsèque) permettant de distinguer les deux sphères » (STALNAKER, 2012, p. 74). Il me semble au

contaire qu'une telle propriété serait non seulement extrinsèque, mais aussi qualitativement impure puisqu'elle consisterait à faire référence à un autre particulier, en l'occurrence, la position spatiale occupée par la sphère [51]. Ainsi, admettre l'existence d'un espace absolu impliquerait de considérer les sphères, certes non comme des doubles purs, mais au moins comme des répliques pures. Néanmoins, nous poursuivrons notre discussion sans céder à l'idée d'un espace absolu.

Une seconde tentative pour défendre la discernabilité des sphères consisterait à utiliser le critère de discriminabilité faible de Quine (W. v. O. QUINE, 1976a), qui stipule que deux choses sont faiblement discriminables si et seulement si il existe une phrase ouverte à deux variables vérifiée par ces deux choses mais pas par une seule d'entre elles avec elle-même. Plus formellement, la version faible du PIdI peut s'écrire :

- **[PIdI$_{\text{faible}}$]** : $\forall x\ \forall y \left(x \neq y \rightarrow \exists \mathrm{R}\ \mathrm{R}xy \wedge \neg \mathrm{R}xx]\right)$

La propriété « d'être à la distance $2 \times$ diamètre de [...] » semble pouvoir faire l'affaire dans le cas des deux sphères. Mais cela n'est vrai que si l'on accepte qu'aucune sphère ne peut être à une certaine distance d'elle-même, ce qui, on le verra plus loin, est loin d'être évident.

Il se peut donc que la dispersion spatiale ne suffise pas pour discerner les sphères. Or il est probable que nous n'ayons aucune autre manière d'y parvenir. Il nous faut donc envisager une autre voie.

Voie 2 : défense du PIdI par l'identité

Après avoir examiné les défenses du PIdI par la discernabilité des sphères, passons maintenant aux défenses par l'identité. Comment invalider la possibilité qu'il y ait deux sphères plutôt qu'une seule ? Si la diversité numérique est déduite de la dispersion spatiale, cela présume que l'argumentation précédente soit corrigée de la manière suivante :

(1)	x et y sont non-connexes	hypothèse de non-connexité
(2)	si x et y sont non-connexes, alors $x \neq y$	principe de distinction des non-connexes
	———	
(3)	$x \neq y$	*Modus ponens* appliqué à (1) et (2)

Il y a donc deux stratégies pour invalider la conclusion de diversité numérique des sphères :

1) rejeter le principe de distinction des non-connexes ;

[51] La conception de Stalnaker semble être qu'une position spatiale ne serait pas un particulier, mais il ne dit pas pourquoi. Il me semble cependant difficile de concéder cet argument. Étant individuée, une position spatiale ferait bien office de particulier. En outre, j'ai déjà dit plus haut que de tels emplacements seraient eux-mêmes des exemples de répliques pures.

2) rejeter l'hypothèse de non-connexité des sphères.

Considérons d'abord la première alternative. Notons que rejeter le principe de distinction des non-connexes, c'est accepter qu'une même chose puisse occuper simultanément différentes positions spatiales.

John O'Leary-Hawthorne (O'LEARY HAWTHORNE, 1995) propose une stratégie de défense du PIdI qui suit cette voie. Elle est appuyée sur une théorie réaliste des faisceaux d'universaux : un objet est entièrement individué par la collection totale des propriétés qui le caractérisent. Ces propriétés étant considérées comme des universaux, elles sont susceptibles d'être instanciées simultanément à différents emplacements. Selon O'Leary-Hawthorne, cela devrait nous inciter à penser qu'il puisse en être de même pour la totalité du faisceau. En conséquence, le particulier étant complètement individué par le faisceau, il devrait lui-même être capable d'être présent simultanément à différents emplacements. Ainsi, les prétendues « deux » sphères ne s'avéreraient finalement n'être qu'un seul et même objet simultanément présent à deux emplacements différents.

Mais il s'agit sans doute d'une manière bien étrange de concevoir les particuliers. Par opposition à eux, les universaux sont habituellement définis par leur capacité à être instanciés simultanément à différents emplacements. Si un même particulier est diversement localisable de manière synchronique, y a-t-il encore un sens à le distinguer d'un universel ? D'autres problèmes surviennent encore. Comme le souligne Katherine Hawley (HAWLEY, 2009), cette conception implique de notables bizarreries méréologiques. Si l'on suppose que le volume de la sphère est de 1 m^3, qu'en serait-il de celle de deux occurrences simultanées ? D'un coté, s'il n'y a qu'une seule sphère, le volume total d'occupation est de 1 m^3. Mais d'un autre coté, nous avons deux apparitions qui instancient pareillement la propriété d'avoir un volume égal à 1 m^3, ce qui fait un compte de 2 m^3 pour l'ensemble. Un tel objet aurait donc, à l'intérieur de son faisceau, la propriété d'avoir un volume sphérique de 1 m^3 et celle d'avoir un volume non sphérique de 2 m^3, ce qui est contradictoire.

Portons maintenant l'objection sur l'hypothèse de non-connexité des sphères. L'hypothèse repose sur le fait que chaque sphère est située à une certaine distance de l'autre. Il pourrait donc être avancé qu'en raison de cette distance, elles sont inévitablement dispersées. Si l'on veut éviter de recourir à une individuation brute des emplacements, ces derniers doivent dépendre des relations que les sphères entretiennent entre elles, et il semble que la déduction qui soutiendrait la diversité numérique des emplacements aurait inévitablement la forme suivante :

(1)	l'emplacement a de la sphère S_1 et l'emplacement b de la sphère S_2 sont tels que $d(a, b) > 0$ où d est une distance	hypothèse métrique
(2)	Si $d(a, b) > 0$ alors $a \neq b$	principe de diversité des distants
(3)	$a \neq b$.	*Modus ponens* appliqué à (1) et (2)

Comme l'a souligné Ian Hacking (HACKING, 1975), le problème est que le putatif principe

de diversité des distants est tout simplement faux : le lecteur de cet article se trouve à quelques 40 000 km de lui-même sur une géodésique. La diversité numérique des sphères ne peut être déduite de la relation métrique qu'elles entretiennent. Tirer la diversité numérique des objets de leur distance réciproque n'est au fond qu'une pétition de principe. Or Hacking estime que la distinction numérique des sphères ne devrait pas être stipulée, mais *prouvée*. Cela ne peut l'être qu'indépendamment de considérations spatiales. En effet, déterminer des relations spatiales entre des objets demande préalablement d'avoir déterminé les conditions métriques ou topologiques du monde dans lequel vivent ces objets. Il est toujours possible de décrire l'expérience de pensée de Black comme celle d'un monde contenant une unique sphère. En effet, on peut envisager une description dans laquelle l'espace suivrait une géométrie courbe où il n'y aurait pas deux sphères séparées d'une distance d, mais bien une seule sphère, servant de point de départ et d'arrivée à un parcourt de longueur d comme illustré sur la figure 4.5.

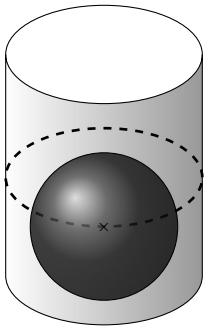


FIGURE 2.6 : Courbure de l'espace et identité

Ian Hacking souligne que le PIdI ne peut être ni infirmé ni confirmé par des arguments basés sur les relations spatio-temporelles. Ni le cas de gouttes indiscernables, ni l'expérience de pensée de Max Black n'infirment le PIdI, parce que dans ces deux exemples, la déduction qu'il y a deux gouttes ou deux sphères provient de la seule raison que les sphères n'occupent pas, synchroniquement, la même position. En dehors de cette différence spatiale, toutes les autres propriétés relationnelles sont identiques. Selon Hacking, il n'est pas suffisant d'attribuer une différence numérique sur le seul fait que les places occupées semblent différentes. Il faut prouver la diversité numérique sans l'appuyer sur une putative diversité spatiale :

> There can be no determination of spatial relations without a study of the laws of nature attributed to objects in space. When such laws are prescribed, most of space falls into place. (...) In arguing that in a certain possible world exist two distinct but indiscernible objects, bland assertion that there are two such objects is not enough. There must be argument. (HACKING, 1975, pp. 250-251)

En outre, Hacking souligne que son objection demeure valide même dans des cas plus sophistiqués où l'on imagine que les sphères se refléteraient les unes les autres. Car à nouveau, le reflet d'une sphère n'impliquerait pas la preuve de l'existence de deux sphères plutôt qu'une. En effet, le reflet pourrait très bien être la projection, sur la sphère, de sa propre image. Hacking conteste aussi l'argument de Strawson dans lequel un observateur placé sur des cases différentes d'un échiquier pourrait avoir deux points de vue absolument indiscernables :

> Think of a chess board. The universe we are to consider is bounded by its edges. The universe consists therefore of a limited arrangement of black and white squares... Even if the view from each square is allowed to comprehend the whole board, [it] is still impossible to differentiate square $(c3)$ from square $(f6)$. The view from each over the whole board is the same: each has two white squares going away from it in one direction and five in the other... (STRAWSON, 1959a, p.)

Selon Hacking, on pourrait tout aussi bien voir les choses ainsi : il pourrait bien n'y avoir en fait qu'une seule case reflétée par un miroir traversant la diagonale du plateau.

Remarquons néanmoins que cette objection ne constitue pas véritablement une réfutation de la conclusion tirée par Black. Hacking ne fait que proposer une nouvelle description du monde de Black. Ce dernier pourrait certes avoir une géométrie courbe et ne contenir qu'une seule sphère. Mais sous une autre description, il pourrait aussi obéir à une géométrie euclidienne et contenir deux sphères. Ainsi, la défense par l'identité de la version faible du PIdI semble échouer. Il reste cependant une dernière voie à explorer.

Selon Hacking, le monde des deux sphères est l'effet d'une description illusoire : « This is a description of a world occupied by one water drop and a soul that sees double" (HACKING, 1975, p. 252). Nous n'aurions affaire qu'à une seule sphère dans un espace courbe. Robert Adams (ADAMS, 1979) a confronté cette objection d'unicité en proposant l'expérience de pensée suivante. Imaginons une répartition infinie de sphères alignées et réparties à distances égales. Admettons en outre que ces sphères soient absolument indiscernables les unes des autres à l'exception d'une seule d'entre elles, marquée d'une petite griffure. Nous aurions ainsi des répliques pures dont chacune serait discernable par le moyen d'une relation impure, à savoir, sa position par rapport à l'unique sphère marquée. Imaginons maintenant que cette marque disparaisse. Dès lors, le seul moyen de discerner les sphères viendrait à manquer. Devrions-nous alors considérer qu'il n'y a plus, maintenant, qu'une seule sphère ? Notre intuition semble réfractaire à l'idée que la présence de différentes sphères soient ainsi ramenée, par l'effacement d'une marque, à l'unicité d'une sphère. Cela reviendrait ni plus ni moins à supposer que choisir de marquer ou non une sphère modifie la géométrie de l'espace, ce qui semble assez absurde [52].

L'argument peut s'employer de manière « sorite », en s'appuyant sur nos intuitions concernant la continuité de constitution. On peut en effet imaginer que l'effacement de la marque se fasse progressivement, par petites altérations successives. Si deux sphères étaient *presque*, mais pas tout à fait indiscernables (voir la figure 2.7), leur distinction numérique ne devrait pas faire discussion. Mais au bout de quelques altérations, ces sphères deviendraient absolument indiscernables. Par application de la transitivité de l'identité, cela impliquerait leur diversité numérique.

Enfin, l'argument d'Adams peut également être formulé de manière modale. Imaginons qu'une des sphères de l'alignement infini soit marquée. Il semble juste de dire qu'elle aurait pu ne pas l'être. Si on affirme que dans ce cas les sphères n'en sont qu'une seule, nous devrions alors, par application du principe de nécessité de l'identité, affirmer cette identité dans tous les mondes

[52] Mentionnons une alternative due à Alfred Ayer, qui consiste à concevoir la répétition infinie de quatre sons joués à intervalles réguliers, de même hauteur, de même timbre et de même durée. Chaque son de la série serait indiscernable des autres, cela, à la fois de manière qualitative mais aussi de manière relationnelle : « for each is preceded and succeeded by the same number of B C Ds, at the same intervals and in the same order. » (AYER, 1954, p. 32).

possibles, ce qui contredit notre hypothèse de départ.

Constatons d'ailleurs qu'il n'est pas besoin de mentionner une infinité de sphères pour parvenir à des résultats similaires. Trois suffisent amplement. En effet, dans un monde qui serait constitué de trois sphères qualitativement indiscernables dont l'une serait située au milieu des deux autres, il n'y aurait pas moyen de discerner de manière relationnelle les sphères situées aux deux extrémités. Dans ce cas, l'objection de Hacking convie à conclure qu'il n'y a en fait non pas trois mais seulement deux sphères. On se retrouve à nouveau dans la situation exposée par Black, à savoir, on a deux sphères indiscernables. Mais si on adhère à nouveau à l'argument d'Hacking, il faut considérer qu'il n'y a en réalité qu'une seule sphère. Mais cela contredit notre hypothèse de départ. En effet, rappelons-nous que nous avions imaginé une sphère discernable par ses relations aux deux autres, à savoir, celle qui était au milieu.

L'argument précédent se généralise en réalité à tout cas fini de sphères. Mais il est même possible de récuser l'objection de Hacking en n'imaginant qu'une seule sphère. En effet, David Wiggins (WIGGINS, 2012, p. 6) a montré que le principe d'identité des indiscernables semble impliquer l'impossibilité d'objets parfaitement symétriques. Lorsqu'on considère le cas d'une sphère parfaite, pour tout couple de demi-sphères, à supposer que l'espace et son orientation soient relatifs, le principe d'identité des indiscernables impliquerait manifestement l'identité des demi-sphères, ce que semble, une fois encore, plutôt absurde. Mais on peut encore pousser le paradoxe plus loin, puisque nous serions successivement conduits à identifier aussi les quarts de sphère, les huitièmes de sphère, les seizièmes de sphère, *ad libitum*. Au fur et à mesure de la régression, les parties symétriques se résorberaient ainsi les unes dans les autres.

Enfin, mentionnons un dernier argument d'Adams, consistant à faire l'expérience de pensée d'une personne affublée d'un *doppelgänger*. Il s'agit d'admettre en outre que la seule différence qui permet de les discerner serait une légère différence dans le contenu d'un rêve. À la différence des cas précédents, ce qui permet maintenant de discerner les individus ne sont plus directement des propriétés physiques, mais des événements mentaux. À nouveau, nous pouvons vérifier que si les rêves avaient été identiques, il serait néanmoins étrange d'affirmer que les individus auraient été numériquement identiques.

Voie 3 : défense du PIdI par non-individuation des sphères

Une dernière tentative visant à rejeter la dispersion des sphères consiste en ce que Katherine Hawley appelle « la défense par la somme » (HAWLEY, 2009). Il s'agit de nier l'individuation de chacune des sphères et à conférer le rôle d'individu à l'objet plus grand constitué par leur somme méréologique. Par comparaison avec ce qui précède, la propriété d'être une sphère était attribuée à deux objets dans le cas de la défense par la discernabilité, et à un seul objet dans le cas de la défense par l'identité. Selon la solution proposée par Hawley, il n'y a tout simplement aucun objet ayant cette propriété. Car un seul objet existe et il est sans parties sphériques. Les deux sphères n'existent pas indépendamment de l'objet total qu'elles constituent à elles deux. Sous cette conception, l'objet considéré existe de manière « éclatée ». Il est fragmenté dans des régions non-connexes de l'espace, bien que ces régions n'en contiennent pas des parties propres. Mais une telle hypothèse demande d'adhérer à une conception extrêmement contre-intuitive des particuliers et implique de revoir complètement nos décomptes ordinaires, ce qui est certainement un coût trop

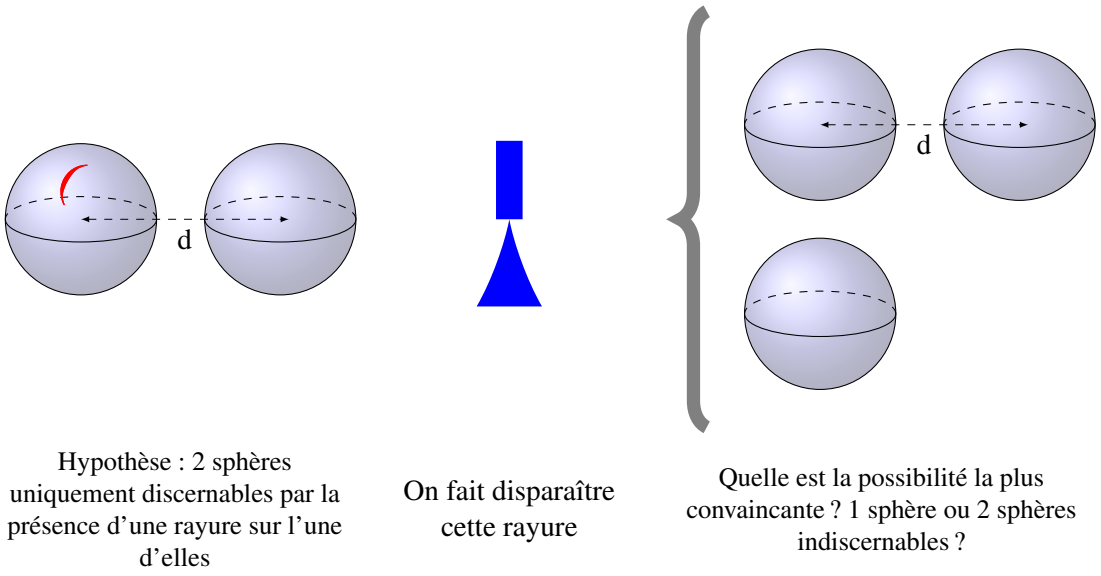


FIGURE 2.7 : Argument modal des sphères quasi-parfaites

important.

Les tentatives de défense du PIdI contre l'expérience de pensée des deux sphères s'avèrent ainsi peu concluantes. En outre, cela indique que toute répartition symétrique d'objets pourrait contrevenir à la version modérée du PIdI. Dans son article, Black souligne que les arguments en faveur des sphères indiscernables sont également valables si l'on considère un nombre plus grand d'objets, voire à l'échelle de mondes. On peut par exemple imaginer un univers composé de deux mondes symétriques séparés par « un plan flottant dans l'espace, avec ceci que tout ce qui se passe sur un de ses côtés soit toujours exactement reproduit à égale distance de l'autre côté. » (BLACK, 2012, p. 128). De tels mondes constitueraient bien des doubles purs.

Dans l'expérience de pensée décrite par Black, la différence numérique des sphères est un fait brut qui contredit d'emblée le principe. Que penser de cette identité brute ? Doit-on convenir avec Russell que « dire que B et C sont "réellement" deux, bien qu'ils semblent un, c'est dire quelque chose qui paraît complètement dépourvu de sens, si B et C sont totalement indiscernables » (RUSSELL, 1940, p. 117) ou doit-on au contraire convenir avec Wittgenstein que « la définition que donne Russell de "=" ne suffit pas ; car on ne peut, selon elle, dire que deux objets ont en commun toutes leurs propriétés. (Même si cette proposition est incorrecte, elle a pourtant un sens) » (WITTGENSTEIN, 1993, aph. 5.5302) ? Les mondes décrits par Black permettent sans doute de conclure en la possibilité logique de doubles purs. La question est finalement de comprendre ce qu'implique, en termes de décomptes, un arrangement qualitatif. La pure concevabilité, désancrée des théories physiques, est certainement un guide précieux pour la possibilité logique de faits purement qualitatifs, mais elle pourrait s'avérer inopérante pour juger de la possibilité métaphysique de faits d'individuation.

Le cas des particules quantiques

Les particules quantiques pourraient bien contredire la version modérée du PIdI. Comme le constate Steven French, dans ce cas, le PIdI ne serait donc pas même vrai de manière contingente (FRENCH, 1989). Par comparaison avec la physique classique, en physique quantique, il reste vrai que deux particules de la même espèce possèdent les mêmes qualités intrinsèques (charge, masse, moment magnétique, spin). Mais contrairement aux particules classiques, ces particules ne sont pas localisables et n'obéissent pas au principe de non-pénétrabilité. Il n'est donc pas possible de les « étiqueter » selon les trajectoires qu'elles parcourraient. De telles entités pourraient donc s'avérer être de bons candidats au rôle de doubles purs.

Afin d'en discuter, rappelons d'abord plus précisément le genre de problème d'individuation que posent les particules quantiques. Partons d'abord du paradoxe que Gibbs (GIBBS, 1902) expose en entropie statistique. Admettons que des particules de gaz se répartissent entre deux boites communicantes. Comparons alors la distribution des particules quantiques à celle prévue, en mécanique classique, par les statistiques de Maxwell-Boltzmann (DALIBARD, 2016). Pour simplifier, considérons qu'on dispose de 2 boites discernables et qu'on y répartisse aléatoirement 2 particules classiques. De manière générale, le nombre de répartitions possibles de n atomes discernables dans k boites discernables est égal à k^n, puisque chaque atome peut être rangé de k façons différentes. Dans l'exemple qui nous occupe, $n = k = 2$ et on a donc $2^2 = 4$ issues équiprobables différentes comme illustré sur la figure 2.8.

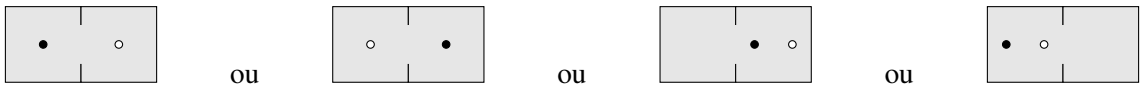


FIGURE 2.8 : 4 permutations possibles de 2 éléments discernables placés dans 2 boites discernables.

Mais l'expérimentation montre que les particules quantiques ne se distribuent pas du tout selon le modèle précédent. La loi de probabilité de distribution des bosons n'est pas la même. En effet, pour n particules et k boites, les statistiques de Bose-Einstein (DALIBARD, 2016) prévoient que le nombre d'issues possibles corresponde aux combinaisons $\binom{n+k-1}{k-1}$. Or ce nombre ne correspond plus aux manières de répartir n particules *discernables* dans k boites discernables, mais aux manières de répartir n particules *indiscernables* dans k boites discernables [53]. Dans le cas qui

[53] Pour établir cette formule, on peut représenter les particules par des « o » et repérer la localisation des boites par des cloisons. Par exemple, pour 6 particules indiscernables rangées dans 4 boites discernables, « |oo|oooo » représente la situation particulière de 0 particule dans la première boite, 2 particules dans la seconde boite, 0 particule dans la troisième boite et 4 particules dans la dernière boite. L'idée est alors de déterminer toutes les possibilités de cloisonner les particules, sachant qu'il y a $k-1$ cloisons à placer sur $n+k-1$ positions (la première position étant implicitement déjà occupée par une barre de cloison). Or, dénombrer les rangements possibles des $k-1$ symboles « | » sur les $n+k-1$ emplacements disponibles revient à déterminer le nombre de combinaisons de $k-1$ éléments choisis parmi $n+k-1$, c'est-à-dire $\binom{n+k-1}{k-1}$.

nous intéresse, on est donc ramené maintenant non plus à 4, mais à $\binom{2+2-1}{2-1} = \binom{3}{1} = 3$ issues équiprobables, comme illustré sur la figure 2.9.

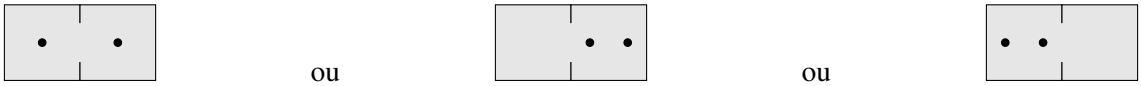


FIGURE 2.9 : 3 permutations possibles de 2 éléments indiscernables placés dans 2 boites discernables

Ce qu'on peut en conclure, c'est que la fonction de probabilité de distribution est insensible à toute permutation de deux bosons. On peut alors interpréter les résultats mathématiques de la théorie quantique de deux manières. Soit on se contente de dire qu'il est, épistémiquement parlant, impossible de discerner un boson d'un autre. Soit on va plus loin, et on admet de manière ontologique que de telles particules ne sont pas réellement des objets, parce qu'elles ne sont pas individuées. Cette dernière hypothèse est retenue par Hermann Weyl qui la commente ainsi :

> La possibilité qu'un des jumeaux Mike et Ike soit dans l'état quantique E1 et l'autre dans l'état quantique E2 ne comporte pas deux cas distinguables qui seraient permutés en permutant Mike et Ike : il est impossible pour chacun de ces individus de conserver son identité de telle manière que l'un d'entre eux soit toujours capable d'affirmer « je suis Mike » et l'autre « Je suis Ike ». On ne peut demander un alibi à un électron, même en principe ! De ce point de vue, le principe leibnizien de *coincidentia indiscernibilium* est bien valide en mécanique quantique. (WEYL, 1931, p. 241)

En un sens, le terme de « particule » serait impropre parce que la question de leur individuation serait tout simplement un non-sens. Les bosons échoueraient, pour cette raison, à constituer des exemples de doubles purs.

Mais plutôt que d'abandonner les particules quantiques au sort de non-individuation, on peut encore tenter d'utiliser la version faible du PIdI. Il se trouve que les fermions suivent une loi de probabilité différente de celle des bosons. Les statistiques de Fermi-Dirac (DALIBARD, 2016) prévoient en effet une unique issue possible pour la distribution des fermions parmi deux boites, à savoir, une particule par boite. Car selon le principe d'exclusion de Pauli, deux fermions ne peuvent se trouver au même emplacement dans le même état quantique. Précisément, si deux fermions sont réunis, leurs spins auront alors des valeurs opposées. C'est pourquoi Simon Saunders (SAUNDERS, 2006) propose d'utiliser la propriété anti-réflexive exprimée par le prédicat « avoir un spin opposé » afin de faiblement discerner ces particules. Si c'est juste, cela signifie alors que les fermions pourraient être des répliques pures, mais pas des doubles purs.

Mais outre le fait que d'autres particules, comme les bosons élémentaires, semblent demeurer réfractaires à la faible discrimination, Steven French et Décio Krause (FRENCH & KRAUSE, 2006, p. 170) objectent que l'utilisation d'un critère de discrimination faible mène vraisemblablement à une circularité vicieuse pour ceux qui veulent fonder les faits d'identité sur des faits qualitatifs. Car affirmer de x qu'il instancie avec y une relation anti-réflexive consiste à présupposer d'emblée que x et y sont numériquement distincts.

Il ressort donc de ce qui précède qu'il est concevable et sans doute physiquement possible que les particules quantiques soient des doubles purs.

2.3.4 Coïncidents purs

Venons-en maintenant au cas des coïncidents purs. De telles entités font l'enjeu d'un certain nombre de problèmes lorsque l'identité d'un individu est ancrée dans sa constitution matérielle. On peut mentionner par exemple le paradoxe du chat Tibbles imaginé par Peter Geach (P. T. GEACH, 1962) et ceux ayant pour résultante la différenciation entre une statue et le bloc de matière qui la compose. Je commence par les rappeler en les illustrant de manière adaptée.

Le paradoxe de Peter Geach

Le premier problème consiste à affirmer qu'un objet matériel quelconque coïncide parfaitement avec une infinité d'autres. J'adapetrai dans ce qui suit le paradoxe de Tibbles au cas d'une statue. Prenons par exemple le cas de la *Pietá* du Vatican. Avant 1972, une unique statue, taillée par Michel-Ange dans un unique bloc de marbre, reposait sur le sol de la basilique Saint-Pierre. Le 21 mai 1972, la statue fut victime d'un acte de vandalisme au cours duquel la partie représentant le nez de la vierge fut détruit. Admettons pour les besoins de notre discussion qu'elle soit restée ainsi [54].

Considérons sous le nom de *Pietà*$^-$, la partie de statue qui n'a pas été endommagée. L'argument de Geach consiste d'abord à montrer que *Pietà*$^-$ n'est pas seulement une « partie détachée » de statue, mais bien une statue. Pour cela, constatons que depuis le 21 mai 1972, l'entité matérielle *Pietà* est parfaitement indiscernable de l'entité matérielle *Pietà*$^-$. De quelle sorte d'objet s'agit-il alors ? Puisque les deux objets ont, après le 21 mai, exactement les mêmes propriétés matérielles, nous n'avons aucune raison de croire que *Pietà*$^-$ n'a pas la même nature que *Pietà*. *Pietà*$^-$ est donc, après 1972, une statue. Mais nous n'avons pas de raison de croire qu'elle ne le fut pas avant cette date. En effet, puisque rien n'est arrivé à *Pietà*$^-$, il n'y a pas de raison de penser que sa nature ait été différente avant 1972.

Sommes-nous en présence de deux statues, *Pietà* et *Pietà*$^-$? Ou s'agit-il d'une unique statue désignée par deux noms différents ? Constatons qu'avant 1972, *Pietà* et *Pietà*$^-$ exemplifiaient des différences : la vierge de *Pietà* avait un nez, mais pas celle de *Pietà*$^-$. Nous pouvons les discerner de manière historique. Dès lors, l'application de la contraposée de l'indiscernabilité des identiques aux objets auxquels réfèrent ces noms implique qu'on a bien affaire à *deux* statues différentes. Plus précisément, à deux statues qui se chevauchent avant le 12 mai 1972 et qui, après cette date, coïncident complètement. En généralisant le mécanisme du paradoxe à n'importe quel objet et à n'importe laquelle de ses parties propres, nous sommes amenés à constater que chaque chose recouvre en réalité une infinité d'autres choses de la même sorte.

[54] Deux alternatives s'offrirent aux restaurateurs. La première était de ne conserver que les éléments d'origine de la statue, et donc de laisser brèches et fêlures en l'état. La seconde était de colmater les parties détériorées par des matériaux de substitution. C'est cette seconde possibilité qui fut choisie. Le Professeur de Campos, chargé de la restauration, jugea qu'il valait mieux que l'œuvre fût reconstruite plutôt que présentée en l'état. Les parties brisées du bloc de marbre furent donc remplacées par des amalgames.

Selon Geach, la solution qu'il convient de donner à ce paradoxe passe par le rejet de l'indiscernabilité des identiques, en abandonnant l'identité absolue au profit d'une identité relative. Dès lors, une chose x peut être identique-en-tant-que-F à une chose y tout en n'étant pas identique-en-tant-que-G à y. Avant 1972, *Pietà*$^-$ est bien la même-statue que *Pietà*, mais pas le même-morceau-de-matière. Il n'y a donc qu'une seule statue et nos comptes sont saufs. Mais, comme nous l'avons déjà dit, l'opération de sauvetage aura pour conséquence de relativiser l'identité. Cette stratégie malmène le principe d'indiscernabilité des identiques et c'est pourquoi peu de philosophes sont prêts à l'accepter.

D'autres solutions ont été proposées. Peter van Inwagen (P. INWAGEN, 1981) rejette l'existence de « parties arbitrairement détachées ». Ce que nous appelons « *Pietà*$^-$ » ne correspondrait à rien avant 1972, et serait identique à *Pietà* après cette date. Citons aussi la solution de David Lewis, consistant à abandonner une vision endurantiste stipulant que les individus persistent dans le temps en étant entièrement présents à chaque moment au profit d'une vision perdurantiste affirmant que les individus sont constitués par des parties temporelles dont ils représentent la somme méréologique. Dès lors, il n'y aurait plus de paradoxe à considérer que *Pietà* et *Pietà*$^-$ partageraient certaines parties temporelles tout en étant numériquement distinctes. Mais cela n'est évidemment possible que si ces individus ne partagent pas *toutes* leurs parties temporelles, sans quoi ils seraient identiques.

C'est justement cette situation que fait intervenir un autre paradoxe de constitution matérielle où la coïncidence n'est pas seulement temporaire mais permanente. Admettons qu'une reproduction *kitsch* de la *Pietà* (désormais *PietàK*) soit produite par injection instantanée de plastique dans un moule. Admettons aussi que quelques instants plus tard, l'objet soit détruit par dissolution dans un bain d'acétone. Ces deux hypothèses ont pour effet de garantir que la statue et le bloc de plastique aient une carrière temporellement coextensive. Donnons le nom « *PlastiK* » au bloc de plastique qui constitue *PietàK*. Constatons que contrairement au cas précédent, aucune différence historique ne permet maintenant de différencier l'une de l'autre. S'agit-il, alors, cette fois, d'une seule et même chose ? Là encore, non. Car en effet, deux intuitions modales semblent parler en faveur de leur distinction :

- d'une part, *PietàK* pourrait persister à la perte de son nez ;
- d'autre part, *PlastiK* ne pourrait pas persister à la perte de son nez.

Autrement dit, le prédicat « persister à la perte de son nez » semble s'appliquer à l'une des entités, mais pas à l'autre. D'après la contraposée du principe de nécessité de l'identité, si deux objets sont possiblement distincts, ils sont alors actuellement distincts [55]. Or *PietàK* et *PlastiK* étant des désignateurs rigides, ils dénotent les mêmes entités dans tous les mondes possibles. Or, d'après le principe d'indiscernabilité des identiques appliqué aux deux hypothèses précédentes, il y a un monde possible dans lequel cette identité n'est pas réalisée. Les deux noms dénotent donc deux entités parfaitement coïncidentes, et cela de manière permanente.

[55] Pour des désignateurs rigides a et b, le principe de nécessité de l'identité affirme que $(a = b) \rightarrow \Box\,(a = b)$, où « $\Box$ » représente l'opérateur modal de nécessité. De manière contraposée, ce principe stipule que $\Diamond\,(a \neq b) \rightarrow (a \neq b)$ où « $\Diamond$ » représente l'opérateur modal de possibilité.

Comme nous l'avons déjà remarqué, certains philosophes boivent le vin du paradoxe jusqu'à la lie en admettant qu'il y a bien des coïncidents permanents, à condition qu'ils ne soient pas la même sorte de choses (WIGGINS, 1968). De ce point de vue, les coïncidents sont finalement numériquement différents mais néanmoins discernables de manière historique ou modale. Cela nécessite de désolidariser l'identité de la constitution matérielle (JOHNSTON, 1992), en concevant deux objets se chevauchant : une statue et un bloc de matière. Nous aurions alors un argument pour affirmer l'existence de coïncidents purs, argument qui prendrait la forme du raisonnement par l'absurde suivant :

(1)	*PietàK* est identique à *PlastiK*	hypothèse d'identité
(2)	*PietàK* est telle qu'elle pourrait persister à la perte de son nez ;	hypothèse de tolérance constitutive
(3)	*PlastiK* est telle qu'elle ne pourrait pas persister à la perte de son nez ;	hypothèse de restriction constitutive
(4)	*PietàK* est telle qu'elle pourrait et ne pourrait pas persister à la perte de son nez.	de (1), (2) et (3).

Pour les partisans de la désolidatisation de l'identité et de la composition matérielle, la conclusion contradictoire de l'argument indiquerait manifestement la fausseté de l'hypothèse d'identité. Mais comment éviter cela ?

La conception quasi-rigide de Lewis

David Lewis, nous l'avons vu, admet bien la possibilité de fissions dans les cas d'une coïncidence qui n'est que temporaire entre deux entités [56]. Mais dans le cas d'une coïncidence permanente, Lewis rejette la solution consistant à admettre un dualisme distinguant ontologiquement les individus de leur support matériel. Lewis soutient que si les parties spatio-temporelles d'un individu a sont exactement les mêmes que celles d'un individu b, il faut alors admettre leur identité. Lewis rejette donc l'hypothèse de coïncidents purs à l'aide de relations de contreparties contextuels pouvant mobiliser simultanément plusieurs aspects comparatifs. Selon le principe de différence des discernables, une différence de propriété modale implique bien une différence numérique. Mais la stratégie de Lewis consiste à objecter qu'on ne prédique pas une même propriété modale à deux choses différentes, mais deux propriétés modales à une seule et même chose. En effet, la théorie des contreparties permet de réinterpréter ainsi les expressions précédentes :

1. (2') *PietàK* est telle qu'il manque le nez à l'une de ses contreparties.
2. (3') *PlastiK* est telle qu'à aucune de ses contreparties il ne manque le nez.

Si l'on en reste là, il semble que la conclusion de distinction numérique continue à s'imposer. Puisqu'elles diffèrent par leurs contreparties respectives, *PietàK* et *PlastiK* sont bien deux

[56] Cette section s'inspire largement de l'article de Filipe Drapeau Contim intitulé *L'essence en contexte* (DRAPEAU VIEIRA CONTIM, 2012).

choses différentes. Cependant, Lewis souligne que la contrepartie d'un individu est toujours visée de manière contextuelle. Il faut donc considérer que l'interprétation donnée ci-dessus est encore incomplète. Il faut rendre plus explicite le type de relation qui permet d'accéder à la contrepartie concernée, par exemple de la manière suivante :

1. (2") *PietàK* est telle qu'il manque le nez à l'une de ses contreparties-en-tant-que-statue.
2. (3") *PlastiK* est telle qu'il ne manque le nez à aucune de ses contreparties-en-tant-que-bloc-de-plastique.

De fait, cette précaution syntaxique permet d'éviter la conclusion de distinction numérique de *PietàK* et *PlastiK*. En effet, tirer la distinction numérique de *PietàK* et *PlastiK* à partir des deux prémisses précédentes relève d'un sophisme du même type que celui-ci, plus manifeste :

- (A) Alice Cooper [57] ne ressemble pas à Emmanuel Macron (par exemple par la profession).
- (B) Vincent Furnier ressemble à Emmanuel Macron (par exemple par les origines françaises).
- (C) Donc Alice Cooper n'est pas Vincent Furnier.

Il n'y a pas plus de raison d'inférer de (2") et (3") que *PietáK* est numériquement distincte de *PlastiK* qu'il n'y a de raison de conclure de (A) et (B) qu'Alice Cooper est numériquement distinct de Vincent Furnier. Ce que dit l'argument, c'est que, considérée comme une statue, *PietàK* pourrait bien perdre son nez, mais que considérée comme bloc de plastique, elle ne le pourrait pas. Aucune raison n'indique qu'il faille cependant en déduire la distinction numérique des entités.

Ce qui explique ceci, c'est qu'un même prédicat modal (« perdre un nez ») peut exprimer des propriétés différentes en fonction du contexte, autrement dit, avoir des contreparties différentes selon son contexte d'énonciation. Refuser le prédicat « n'avoir possiblement pas de nez » à *PlastiK* est contextualisé par la propriété modale de n'avoir aucune contrepartie-en-tant-que-bloc-de-plastique auquel il manque le nez. Quand le même prédicat « n'avoir possiblement pas de nez » est attribué à *PietàK*, il crée un nouveau contexte. La différence de prédicat n'induit pas une différence réelle entre les deux entités mais une différence de représentation dans des mondes possibles. Considéré sous l'aspect d'une statue, il est est à la fois vrai de *PietàK* et de *PlastiK* qu'elles pourraient survivre à la perte du nez : cela signifiant que les contreparties de *PietàK*-en-tant-que-statue peuvent subir ce dommage. Et de même, considérées sous l'aspect de blocs de plastique, il est tout aussi vrai de *PietàK* que de *PlastiK* qu'elles n'auraient pas pu persister à cette perte : il n'existe pas de contrepartie de *PietàK*-en-tant-que-bloc-de-plastique à laquelle il manquerait le nez.

Revenons, avec ces précisions, à l'argument de départ :

- *PietàK* est identique à *PlastiK* ;

[57] « Alice Cooper » est le pseudonyme de scène de l'auteur, compositeur et interprète Vincent Furnier.

- *PietàK*-en-tant-que-statue est telle qu'elle pourrait persister à la perte de son nez ;
- *PlastiK*-en-tant-que-bloc-de-plastique est telle qu'elle ne pourrait pas persister à la perte de son nez ;

Nous n'avons plus aucune raison maintenant d'inférer l'identité numérique de *PietàK* et *PlastiK*. La théorie des contreparties permet de répondre au paradoxe de constitution en ce que l'apparente distinction de propriétés modales n'est en fait qu'une différence de description d'une seule et même chose.

La contextualisation des représentations associées à un même individu pourrait ainsi régler notre problème, s'il n'était pas aggravé par l'admission du principe de nécessité de l'identité. Si *PietàK* et *PlastiK* sont actuellement identiques, il est cependant vraisemblable qu'elles auraient pu s'avérer distinctes. La coïncidence est contingente : *PietàK* pourrait être refondue en une autre forme qui préserverait l'intégrité de *PlastiK* sans préserver celle de *PietàK*. Même en faisant l'hypothèse d'une coïncidence actuelle permanente, il y a quand même des scénarios dans lesquels cette coïncidence s'avère seulement temporaire. Or nous avons vu que dans ce cas, nous aurions bien affaire alors à deux choses différentes. Cela vient à admettre que, actuellement identique à *PlastiK*, *PietàK* aurait toutefois pu être distincte de *PlastiK*. Mais cela constitue une entorse évidente au principe de nécessité de l'identité, dont la contraposée affirme que si deux objets diffèrent possiblement, alors ils diffèrent *actuellement*.

L'argument est donc le suivant :

(1)	*Pietà* et *PlastiK* sont telles qu'elles auraient pu être différentes	hypothèse
(2)	*Pietà* est identique à *PlastiK*	hypothèse
(3)	*PlastiK* et *PlastiK* sont telles qu'elles auraient pu être différentes	(1) et (2) et substitutivité des identiques *salva veritate*

Pour éviter la contradiction, Filipe Drapeau Contim (DRAPEAU VIEIRA CONTIM, 2012, p. 126) distingue trois possibilités :

- Rejeter l'identité *Pietà* = *PlastiK* en affirmant par exemple que notre intuition d'identité est l'effet d'une illusion modale ;
- Accepter l'identité *Pietà* = *PlastiK* en rejetant l'affirmation que « *Pietà* » ou/et « *PlastiK* » sont des désignateurs rigides ;
- Accepter l'identité *Pietà* = *PlastiK* mais rejeter la validité du principe de nécessité de l'identité ;

Lewis choisit la troisième possibilité en rejetant le principe de nécessité de l'identité. Il accepte donc l'identité sans protester contre l'idée que les noms sont des désignateurs rigides. Contre

Kripke, Lewis affirme bien qu'on peut soutenir à la fois la thèse des désignateurs rigides et à la fois la thèse de l'identité contingente. Cette voie pourrait pourtant paraître d'emblée impraticable pour Lewis en ce que, dans sa théorie des contreparties, le même individu ne peut pas exister dans deux mondes distincts [58]. Mais de fait, dès sa première théorie des contreparties, Lewis insiste bien sur le caractère rigide des désignateurs dans les modalités *de re* :

> In my original counterpart theory, any *de re* modal predication is referentially transparent. Something has the same counterparts how ever we may choose to refer to it. Given a *de re* modal predication, we find the thing denoted by the subject term in the actual world; then we consider what befalls that thing-or rather, its counterparts in other worlds. Only the denotation of the subject term matters. We can substitute another subject term with the same denotation but different sense, and the truth value of the modal predication will not change. (LEWIS, 1971, p. 211)

Dans la seconde version de sa théorie, Lewis est amené à assouplir cette rigidité. La référence d'un terme dans un contexte modal n'est plus directe, mais médiatisée par le contexte dans lequel elle est proférée. N'en déplaise à Kripke, Lewis ré-introduit donc une nouvelle sorte de contexte opaque dans les énoncés modaux qui tient tête à la théorie des désignateurs rigides. Cela semble donc marquer un renforcement à une forme de descriptivisme :

> *de re* modal predications are not in general transparent. Not only the denotation of the subject term matters, but also the counterpart relation it selects. If we substitute another subject term with the same denotation but different sense, it may change the truth value of the modal predication by selecting a different counterpart relation. Then even though the denoted thing here in our world remains the same, we have a different way of following the fortunes of that thing in other worlds. (LEWIS, 1971, p. 211)

Dans les dernières pages de *De la pluralité des mondes*, Lewis qualifie cette sorte de rigidité de quasi-rigidité :

> [...] sans le recouvrement des mondes, nous n'attendrions pas du nom propre ordinaire d'une personne ou d'une chose [...] qu'il soit strictement rigide. Néanmoins, un nom propre ordinaire pourrait bien être *quasi-rigide* : c'est-à-dire qu'il pourrait nommer, là bas, dans un autre monde, la contrepartie de ce qu'il nomme ici. (LEWIS, 2007, p. 389)

L'intérêt de la quasi-rigidité est que différentes relations de contrepartie peuvent intervenir dans le même énoncé. Lewis en donne l'exemple suivant :

> Pour une illustration familière du besoin de relations de contrepartie exploitant différents aspects de comparaison, considérez les énoncés du type « si j'étais vous... ». Les mondes qui rendent vrai l'antécédent sont des mondes dans lesquels vous et moi sommes représentés comme étant identiques ; autrement dit, nous avons une contrepartie en commun. Mais nous voulons que celle-ci soit dans votre situation avec mes idées, et non l'inverse. Elle devra donc

[58] Filipe Drapeau Contim le souligne d'ailleurs lui-même (DRAPEAU VIEIRA CONTIM, 2012, p. 126).

> être votre contrepartie sous une relation de contrepartie attachée à la similarité de situation, et la mienne sous une relation de contrepartie différente, liée à la similarité d'idées. (LEWIS, 1973, p. 43)

Selon Lewis, le cas inverse peut aussi se produire. C'est-à-dire qu'il est possible qu'un seul et même individu ait deux contreparties par deux aspects différents sélectionnés dans un même contexte d'énonciation. Les relations de contrepartie ne sont donc pas seulement contextuelles, elles sont aussi plurivoques. Il s'agit ainsi, écrit Lewis d' « utiliser une multiplicité inoffensive de relations de contrepartie afin d'éviter la multiplication indésirable des entités » (LEWIS, 2007, p. 43). Comme le montre Filipe Drapeau Contim (DRAPEAU VIEIRA CONTIM, 2012, p. 132), c'est précisément ce qui se passe dans le cas de *PietáK* et de *PlastiK*. Même s'il n'y a actuellement qu'un seul objet, nous pouvons dans un même contexte le considérer à la fois comme une statue et comme un morceau de plastique. On peut donc sélectionner simultanément deux relations de contrepartie. Admettons qu'on ait un monde possible dans lequel une statue de plastique *PietàK'* représente *PietàK*, et qu'elle ait perdu un nez, ce qui a détruit le bloc de plastique *PlastiK* dont elle était originellement constituée, et engendré un autre bloc *Plastik'*. On a bien affaire, dans ce monde possible, à deux individus *PietàK'* et *Plastik'* distincts (quoique partageant entre elles des parties temporelles communes). Comme l'indique Contim, on peut alors ré-interpréter la phrase « *PietàK* aurait pu être distincte de *PlastiK* » de la manière suivante : « Il y a un monde possible dans lequel une contrepartie$_{statue}$ de PietàK (autrement dit PietàK') est distincte d'une contrepartie$_{plastique}$ de PietàK (autrement dit PlastiK'). » Il n'y a ainsi, constate Contim, plus de contradiction à n'admettre actuellement qu'un seul objet, considéré de deux manières différentes, bien qu'il aurait pu y en avoir deux. Comme l'écrit Lewis,

> [...] ces références différentes tendent à évoquer deux façons différentes de représenter, de sorte qu'un même monde peut donner lieu à deux représentations *de re* conflictuelles d'une même chose. Vous pouvez dédoubler, ici dans un même contexte, les façons dont un monde représente un unique objet en plastique et en forme de bassine. Ou bien vous pouvez dédoubler les entités elles-mêmes, en distinguant le morceau de plastique de la bassine. Je soutiens que le premier dédoublement est plus crédible que le second. (LEWIS, 2007, p. 253)

Il est clair que Lewis détourne ainsi la thèse de nécessité de l'identité. Mais contrairement à ce que soutient Contim, et quoi que puisse en dire Lewis, il me semble que la notion de quasi-rigidité est en désaccord complet avec la théorie des désignateurs rigides. La quasi-rigidité n'est pas une rigidité assouplie : elle n'est pas du tout rigide, elle ne peut l'être puisque pour cela, il faudrait que la représentation dans un monde le soit. Or tout peut représenter n'importe quoi.

Ce qui motive Contim dans l'explication qu'il donne de la stratégie de Lewis apparaît sans doute dans ce commentaire : « [...] assimiler ou pas les noms à des descriptions, les tenir pour rigides ou au contraire non rigides, est une thèse purement sémantique qui est indépendante des vues métaphysiques qu'on peut adopter vis-à-vis de la nature des individus et des mondes possibles » (DRAPEAU VIEIRA CONTIM, 2012, p. 126). A mon sens, la thèse de rigidité des noms propres n'est pas seulement sémantique, elle est imbibée de toute part par un fort principe métaphysique en ce qu'elle implique l'idée d'individus qui sont dénotés indépendamment de nos manières de nous les représenter. Elle est basée sur le constat d'un domaine objectif d'individus réalisés par le

monde actuel, et elle est inséparable de la thèse de la nécessité de l'identité, qui en découle directement, et qu'il est difficile de ne pas reconnaître comme une thèse métaphysique. Il semble donc improbable de désolidariser, comme le souhaite Contim (DRAPEAU VIEIRA CONTIM, 2012, n. 29 p. 129), une thèse métaphysiquement neutralisée de rigidité de la thèse de nécessité de l'identité.

Lorsque Contim montre que Lewis utilise la quasi-rigidité non pour référer, par la médiation d'un nom propre, à un individu existant dans un autre monde possible, mais pour le rapporter à tout un ensemble de représentants dans une multiplicité d'autres mondes possibles, cela témoigne plus probablement d'un amendement descriptiviste de la théorie initiale de Lewis, par une sorte d'adaptation de la notion de ligne de mondes à l'intérieur de la théorie des contreparties. La quasi-rigidité est en effet définie comme un rapport aux individus qui n'est plus basée sur une individuation métaphysique qui pourrait être, d'une certaine façon, rigidifiée par un jeu de ressemblances objectives, mais qui est dépendante d'identifications relatives à une multiplicité de représentations. Les présupposés métaphysiques qui interviennent dans la relation de rigidité ne sont pas neutralisés par la théorie des contreparties : que Lewis l'admette ou non, c'est vraisemblablement la thèse de rigidité qu'il abandonne, en rapprochant la théorie des contreparties d'une conception perspectiviste et scénaristique des mondes possibles comme celle employée par Hintikka [59].

Même si des solutions existaient pour éviter de stipuler des coïncidents permanents, de tels paradoxes sont loin d'être les seuls à faire intervenir des coïncidents purs. Le rejet de la version modérée du PIdI implique inévitablement leur possible prolifération. Car l'hypothèse que l'identité est un fait « brut » jointe au rejet du principe d'identité des indiscernables implique une conséquence apparemment néfaste, à savoir, la possibilité que chaque chose coïncide avec une infinité d'autres. La difficulté est soulignée par (DELLA ROCCA, 2005), et déjà pointée dans l'article de Black :

> Si le principe était faux, le fait que je puisse voir seulement deux de tes mains ne serait pas une preuve que tu n'en as que deux. (...) Tu pourrais avoir neuf mains - différentes les unes des autres et toutes indiscernables de ta main gauche - et neuf mains de plus - toutes différentes entre elles mais indiscernables de ta main droite. (BLACK, 2012, p. 123)

De telles entités seraient absolument indiscernables, non seulement de manière intrinsèque et extrinsèque, mais également de manière historique et modale. Et elles le seraient de manière permanente. En cela, ces coïncidents purs entretiendraient entre eux des relations encore plus serrées que celles des coïncidents mis en scène dans les paradoxes précédents. Admettre leur possibilité serait admettre des entités aussi étranges qu'absolument indétectables.

On pourrait spontanément rejeter les coïncidents purs en affirmant qu'aucun plancher ne résisterait à la masse d'une infinité d'objets coïncidents. Mais à bien y réfléchir, chaque objet étant constitué exactement de la même matière que l'infinité de ceux avec lesquels il coïncide, la masse ne devrait finalement pas être excédentaire, et les planchers pourraient bien résister. Les coïncidents purs ne semblent ni logiquement contradictoires, ni même physiquement contradictoires puisque les théories physiques demeurent sous-déterminées à leur sujet. Mais le moins qu'on puisse dire, c'est qu'ils s'avèrent métaphysiquement gênants.

[59] A ce sujet, voir aussi l'article *How to Lewis a Kripke-Hintikka* d'Alessandro Torza (TORZA, 2013).

2.3.5 Mondes haeccéitiques

Revenons maintenant à la question initiale des mondes similaires. Nous avons dit qu'en admettant que les atomes de la physique classique ou que les particules élémentaires de la physique quantique relèvent respectivement du cas de répliques pures et de doubles purs, nous devrions admettre ceci également à l'échelle des mondes. Nous serions alors engagés à admettre l'existence possible, sinon actuelle, de mondes strictement similaires par les qualités. Dans ce qui suit, j'appellerai « mondes haeccéitiques » de tels mondes qualitativement indiscernables quoique numériquement distincts.

L'hypothèse qu'un certain monde possible soit un double qualitatif d'un autre monde tout en étant numériquement distinct de lui implique que l'individuation n'est pas entièrement réductible aux seules propriétés qualitatives. Les faits d'identité ne surviennent pas sur les qualités. En revanche, pour la thèse anti-haeccéitiste, il n'est pas possible pour deux choses d'être qualitativement similaires sans être aussi numériquement identiques. La thèse d'anti-haeccéitisme semble d'emblée ainsi solidaire du principe d'identité des indiscernables, sinon qu'elle est appliquée de manière transmondaine. Mais c'est peut-être aller un peu vite. Quelles sont au juste les relations entre le principe d'identité des indiscernables et l'anti-haeccéitisme ? Constatons que le second principe n'implique pas le premier : par exemple, Lewis défend l'anti-haeccéitisme mais reste neutre quant à l'identité des indiscernables.

Il est nécessaire de bien distinguer les trois notions différentes que sont le point de vue *haeccéitique*, le recours aux *haeccéités* et les particuliers nus. À la différence du PIdI et de la croyance aux haeccéités, l'haeccéitisme est une thèse modale. En conséquence, cette thèse est tributaire des préjugés métaphysiques qui structurent les conceptions modales qui en font l'enjeu. En dépit de ce que pourrait encourager la terminologie employée, la théorie haeccéitique n'est pas la thèse d'haeccéité, car si la seconde implique la première, en revanche, il n'est pas le cas que la première implique la seconde. On peut souscrire à une forme d'haeccéitisme des mondes sans recourir pour cela à des essences individuelles.

Chez Lewis, la relation de contrepartie étant basée sur la similarité, les mondes haeccéitiques n'ont absolument aucun sens. Deux mondes ne peuvent être indiscernables. En ce sens, on peut donc parler d'un anti-haeccéitisme de Lewis. En revanche, Lewis ne rejette pas les possibilités maximales haeccéitiques. Il constate en effet que certaines possibilités haeccéitiques sont difficiles à rejeter. Par exemple, il considère le cas d'un univers constitué de mondes qui seraient à la fois répétés temporellement et répétés spatialement. Plus précisément, il envisage l'hypothèse de trois mondes indiscernables alignés dans trois régions différentes de l'espace, chacun d'eux étant répété à l'infini de manière cyclique, cycle ayant toutefois une origine permettant de localiser chacun d'entre eux dans le temps. Considérons un individu, $i_{\mathcal{R}_2,17}$ habitant la seconde région $\mathcal{R}_2$ du 17e cycle. Un tel individu aurait bien sûr des répliques pures à la fois dans les deux autres régions spatiales mais aussi dans les différents cycles temporels. Considérons alors l'hypothèse suivante : $i_{\mathcal{R}_2,17}$ aurait pu être l'actuel habitant de la région $\mathcal{R}_3$ du 18e cycle qui est actuellement sa réplique (disons $i_{\mathcal{R}_3,18}$). Admettre cette possibilité implique que nous fassions référence non seulement à des individus qui ne différeraient pas qualitativement, mais plus encore à des possibilités qui ne différeraient pas qualitativement.

Étant donné que Lewis n'accepte pas les mondes possibles indiscernables, il s'agit pour lui

d'une véritable difficulté. La solution qu'il donne consiste à admettre qu'il n'y a véritablement qu'un seul monde possible qui représente deux possibilités. La possibilité exprimée ne nous guide pas vers un nouveau monde, mais nous fait retourner dans le même monde que celui de départ. Or si Lewis soutient qu'il n'y a ici qu'un seul monde et non deux, il lui faut alors reconnaître que $i_{\mathcal{R}_2,17}$ a une contrepartie différente de lui-même (à savoir, $i_{\mathcal{R}_3,18}$) *dans son propre monde*. Or ceci était, rappelons-le, interdit par l'axiome 5 de la première version de la théorie des contreparties. C'est la raison pour laquelle Lewis a révisé sa théorie en rejetant finalement cet axiome : « Un même monde peut donner nombre de possibilités, puisque nombreux sont les individus qui y habitent » (LEWIS, 2007, p. 352)

On voit ici se disjoindre les questions épistémique et métaphysique. Selon Lewis, il n'y a pas de mondes possibles haeccéitiques, autrement dit, il n'y a pas de mondes possibles qui soient des doubles purs. En revanche, il y a des possibilités haeccéitiques, entendues comme des représentations *de re* différentes. C'est ce que Lewis appelle l'« haécceitisme économe (*cheap haecceitism*) »

> L'haeccéitiste dit : deux possibilités, deux mondes. Ils *semblent* être exactement semblables, mais il faut bien qu'ils diffèrent de quelque façon. Ils diffèrent par "l'identification transversale", c'est-à-dire par ce qu'ils représentent *de re* à propos de quelqu'un. (LEWIS, 2007, p. 352)

La différence n'est donc pas numérique, elle n'est pas liée à l'identité, elle est représentative, liée à l'identification. Il ne s'agit donc pas d'une différence réelle, mais nominale. Toute possibilité témoigne certes d'un monde possible, c'est-à-dire qu'à chaque possibilité correspond exactement un seul monde possible, mais en revanche, à chaque monde possible peuvent correspondre différentes possibilités *de re* à propos de ce monde.

Considérons encore un autre exemple de Lewis. Soit l'affirmation « j'aurais pu être l'un de deux jumeaux : l'aîné ou le puîné ». En traduisant dans le langage des mondes possibles, nous dirions qu'il y a un monde M_1 dans lequel je suis l'aîné ; et qu'il y a un monde M_2 dans lequel je suis l'autre, le puîné. En quoi différeraient ces deux mondes ? Notons d'abord que la situation est certainement complexifiée, par rapport au cas précédent des mondes répétés, parce qu'elle recourt à un phénomène d'auto-référence, dénoté par l'usage du pronom personnel « je ». Dans le monde actuel, « je » renvoie à la localisation d'un certain corps dans l'espace, approprié comme étant le mien [60]. Mais comment ce « je » pourrait-il encore s'appliquer de manière transversale ? Si l'on considère que le pronom « je » fonctionne comme un désignateur rigide, sans céder à la superstition de l'« âme », on est alors amené à dire que la totalité du flux de pensée produit par tel corps aurait pu prendre siège dans un autre corps. Je laisse pour l'instant de côté la question de l'identité personnelle qui fera l'objet du dernier chapitre. Contentons-nous pour l'instant de suivre la démarche de Lewis, qui constate que nous avons là encore *deux* possibilités indiscernables. A nouveau, Lewis déclare que ces deux possibilités ne constituent pas *deux* mondes différents. Elles diffèrent par la représentation *de re*, mais pas par les mondes possibles qu'elles représentent :

> Je dis : deux possibilités, et je le confirme. Elles diffèrent bien en représentation *de re*. Selon l'une, je suis le premier né des jumeaux ; selon l'autre, je suis le second. Mais ces possibilités

[60] Bien sûr, des erreurs peuvent se produire, comme des cas indus d'autoscopie ou une absence erronée d'identification : nous verrons cela dans le dernier chapitre de ce travail.

> ne constituent pas deux mondes. Ce sont des possibilités dans un même monde. (LEWIS, 2007, p. 352)

Autrement dit, il y a deux échelles différentes : pour chacun des individus, on aurait certes différentes possibilités. En revanche, pour le tout qu'est le monde, il n'y aurait qu'une seule possibilité :

> Chaque jumeau est une manière possible d'être pour une personne, et elle est de fait une manière possible d'être pour moi-même. J'aurais pu être l'une ou l'autre. Il y a pour moi deux possibilités distinctes. Mais pour le monde, elles n'impliquent qu'une seule possibilité ; il aurait pu être le monde habité par ces deux jumeaux. (LEWIS, 2007, p. 352)

Il n'y a donc pas de relation un-à-un entre les possibilités et les mondes possibles. Il est vrai que les mondes purement qualitatifs ne rendent pas compte de toutes les possibilités. Mais il ne s'agit pas d'en déduire qu'ils sont insuffisants et qu'il faille encore postuler, à côté d'eux, d'autres mondes, qui ne sont pas purement qualitatifs. Il n'y a, souligne Lewis, que des mondes purement qualitatifs. Mais ce qu'il faut en déduire, c'est que ces mondes nous offrent un panel trop restreint de possibilités, et qu'il faut alors se tourner vers les parties de mondes.

Lewis considère encore un troisième exemple, lui aussi lié aux questions d'identité personnelle : admettons que j'aurais pu être ce pauvre défunt de Fred. En disant cela, je ne contemple pas un monde qui serait qualitativement différent de celui-ci. On peut avoir l'idée de vivre la vie de ce pauvre Fred tout en demeurant le même, ce qui indique une différence de représentation *de re* qui serait non qualitative. Mais on peut aussi considérer être ce pauvre Fred en plus d'en mener la vie. Pour Lewis, une telle représentation ne consisterait pas en celle d'un monde qui n'aurait aucune différence qualitative avec le nôtre. Car il ne s'agit pas de considérer un individu possible, Fred, qui serait ma contrepartie dans un autre monde. Il s'agit de considérer que l'individu actuel qu'est Fred est ma contrepartie dans notre monde commun.

> Comme n'importe quelle autre personne possible, il est une manière possible d'être pour une personne. En un sens, il est même un manière possible d'être pour moi. Il est ma contrepartie sous une relation de contrepartie extraordinairement généreuse, une relation qui n'exige rien des contreparties, sinon d'être des choses du même genre. (LEWIS, 2007, p. 353)

De manière plus générale, une proposition comme celle exprimée par « Si j'étais vous » semble indiquer une idée comme « à votre place », autrement dit, l'idée de vivre la vie d'un autre en conservant cependant sa propre personnalité et ses propres pensées. De fait, on peut se demander si une telle idée n'est pas absurde. La désolidarisation entre possibilités et mondes qui est utilisée par Lewis semble impliquer une forme de mise en forme métaphysique forcée d'un problème qui est sans doute seulement épistémique. Avoir une double vision par l'effet d'un strabisme est une chose. Prétendre que cette double image se rapporte à une double réalité est autre chose. Or chez Lewis, les possibilités sont d'emblée profondément ancrées sur des réalités, celles des mondes possibles.

Il semble qu'en défendant l'idée que, en raison de la nature des représentation *de re*, il y ait plus de possibilités que de mondes possibles, Lewis apporte *pace Lewis* des arguments pour une

conception anti-réaliste des mondes possibles. Car il appert ici une séparation nette entre la ligne épistémique et la ligne métaphysique des mondes possibles [61].

Si l'on en reste à une conception non métaphysique des mondes possibles, les mondes indiscernables peuvent s'expliquer par des propriétés non-qualitatives. D'un strict point de vue sémantique, deux mondes pourraient bien être identiques par leurs qualités, mais différents par leurs propriétés modales. Dans un monde donné, nous aurions des objets dont les qualités seraient possédées de manière nécessaire, et dans tel autre monde qualitativement identique au premier, certaines propriétés de ces objets ne seraient possédées que de façon contingente. Aussi, la différence entre ces deux mondes ne résiderait pas dans ce qui y est contenu, mais dans ce qui est ou non accessible depuis l'un ou l'autre. Les ensembles d'énoncés non modaux décrivant respectivement ces mondes pourraient donc être complètement identiques, sans admettre pour autant que ces deux mondes coïncident sur l'ensemble de leurs énoncés modaux.

2.4 Conclusion

Il est temps de synthétiser ce parcours à travers la métaphysique du double. Ce qui en ressort, c'est qu'il y a bien un ensemble de problèmes liés aux doubles en philosophie. Ce que nous avons indiqué au passage, c'est que ces problèmes peuvent varier selon différents types, d'une part les doubles-mêmes produits par l'identité croisée dans la théorie essentialiste des mondes possibles de Kripke, d'autre part, les indiscernables que sont les répliques pures, les doubles purs et les coïncidents purs. Parmi ces cas de doubles, nous avons vu que les coïncidents purs manifestent une catégorie particulièrement tératologique qui semble miner nos décomptent ordinaires.

En outre, nous avons vu que, sous une perspective réaliste admettant de quantifier sur des propriétés, les atomes de la physique classique et les particules élémentaires de la physique quantique relèvent respectivement du cas des répliques pures et des doubles purs. Or, nous avons vu qu'en admettant cela à l'échelle des particules, nous sommes manifestement conduits à l'admettre également à l'échelle des mondes. Notre monde serait ainsi dédoublé par une infinité d'autres. Mais nous devrions être attentifs que ce que nous appelons atome et particule élémentaire sont d'abord des êtres de langage plutôt que des réalités métaphysiques.

Ce sont des présupposés métaphysiques, dont principalement le réalisme des propriétés et l'essentialisme qui engendrent de telles pathologies métaphysiques. Elles ont pour origine l'idée que le processus d'identification des individus serait complètement indépendant de leur mode de représentation. Or l'aspect sous lequel nous connaissons une chose joue inévitablement un rôle cognitif dans notre compréhension du monde. Cet aspect pouvant varier avec un certain degré de liberté, la réification des représentations d'un individu en une réalité métaphysique a pour conséquence inévitable la production de doubles. C'est directement lié à notre tendance naturelle, dans nos processus cognitifs, à identifier des choses manifestant entre elles des similarités. Mais il faut garder à l'esprit que ces similarités sont contextuelles, relatives à nos intérêts. Dès que nous oublions ce fait, des doubles ontologiques apparaissent.

[61] C'est d'ailleurs ce que remarque aussi Robert Stalnaker qui écrit que : « [...] I think his way of characterizing the doctrine blurs the difference between metaphysical questions about what kinds of distinction there are to be made between the possibilies and questions about how we can do talk and think about the possibilities » (STALNAKER, 2012, p. 59).

En adoptant le point de vue du réaliste des universaux, nous avons conclu en faveur d'une possibilité logique et physique des répliques pures et des doubles purs. Or, comme nous l'avons vu, il semble alors qu'il faille accepter la possibilité de coïncidents purs. Pourtant, comme l'indique d'ailleurs la réponse laconique de Black dans son article, il paraît très vraisemblable que nous sachions compter des mains, en particulier, lorsqu'il n'y en a que deux. Contentons-nous donc de continuer à compter exactement deux mains au bout des bras, et non dix-huit.

Il est vraisemblable que la putative menace occasionnée sur nos décomptes par les coïncidents purs n'en soit une que pour celui qui fait reposer l'individuation des particuliers sur des principes métaphysiques. C'est sans doute la croyance que les choses sont individuées indépendamment de la façon dont nous y faisons référence qui est responsable des paradoxes qui précèdent. En ce sens, la question qui sous-tend la critique d'Hacking, qui est celle de décrire correctement nos expériences de pensée, reste pertinente. Si l'on évite la question métaphysique de l'individuation au profit de la question épistémique de l'identification, le problème des coïncidents purs semble se dissiper de lui-même. Souscrire à un tel point de vue, c'est bien sûr rejoindre un anti-réalisme qui sape, du même coup, les bases réalistes du débat concernant le PIdI.

Le problème des doubles ontologiques réside dans l'idée même de l'individuation comme principe métaphysique. Le recours au PIdI, qui prétend fournir un tel principe, n'est certainement pas la seule solution face aux coïncidents purs. Nous pouvons aussi suivre une règle de parcimonie prescrivant qu'il est sans intérêt de compter plus de choses que nécessaire. Si nous avons le choix entre deux modèles représentant une même situation, et s'il n'y a pas de différence qualitative entre les deux modèles, nous devons compter en choisissant le modèle qui prescrit le plus petit nombre d'entités. Constatons que dans l'expérience de Black, il y a une différence qualitative entre le modèle à une sphère et son concurrent à deux sphères : par exemple par le fait que les masses et les volumes exemplifiés doublent d'un modèle à l'autre. En revanche, un modèle faisant l'hypothèse de coïncidents purs ne différerait aucunement d'un modèle faisant mention d'une unique chose. Et dans ce cas, nous devrions arrimer la pertinence au deuxième modèle plutôt qu'au premier. Or si l'on admet que la situation décrite par Black contredit bien le PIdI, la règle de parcimonie en est donc indépendante. Et cela me paraît simplement montrer que nos manières d'individuer les choses sont dépendantes de nos manières de les compter, et non l'inverse.

En particulier, les discussions qui précèdent auront montré qu'une tension se manifeste entre la théorie de la référence et la métaphysique des modalités. Cette tension prend son point culminant chez Lewis avec l'insertion, dans sa seconde version de la théorie des contreparties, d'une relation d'identification par représentation, autrement dit, d'une relation qui n'est plus basée sur une métrique objective de ressemblances qualitatives, mais sur une représentation contextuelle. L'idée générale de Lewis est de remplacer la relation d'identité transmondaine par une relation de représentation qui est dépendante du contexte et qui peut générer plusieurs aspects circonstanciels dans la même interprétation.

Nous avons vu que la stratégie déployée par Lewis permettait d'évincer les coïncidents purs. S'agit-il d'une façon toute humienne, *pace* Kripke, de rabattre les modalités aléthiques sur les modalités épistémiques, autrement dit, de rabattre la possibilité métaphysique sur le concevable ? S'agit-il d'une allégeance à la formule de Hume, basculant l'essentialisme par dessus bord dans sa célèbre formule : « La nécessité est quelque chose qui existe dans l'esprit, pas dans les objets, et il ne nous est jamais possible d'en former l'idée la plus lointaine si nous la considérons comme

une qualité qui se trouve dans les corps » (HUME, 1968b, p. 1.3.14). Quelle que soit la réponse à cette question, l'analyse de la théorie des contreparties aura montré qu'une manière d'échapper au dualisme identité/composition, et plus généralement aux coïncidents purs, passe par l'abandon de la thèse de désignation rigide. Si l'on s'en tient à une théorie de l'identification par représentation, les doubles deviennent alors ontologiquement inoffensifs.

Tout ceci tient à ce que l'identité n'est pas une réalité métaphysique, mais relève d'une aptitude cognitive. Si nous adoptions un langage suffisamment précis, nous ne devrions jamais dire que deux objets sont identiques, mais que nous connaissons que deux objets sont les mêmes. De ce point de vue, l'identité est de part en part épistémique. Comme l'écrit Luis Jorge Prieto,

> telle que nous la définissons, l'identité qu'on reconnaît entre deux objets se situe, non pas dans ces objets eux-mêmes, mais dans la connaissance que l'on a d'eux. L'identité qu'on reconnaît à un objet, en effet, n'est en définitive que la façon dont on le connaît, et deux objets identiques entre eux ne sont par conséquent que deux objets que l'on connaît de façon identique. En toute rigueur, donc, il faudrait dire, non pas que deux objets sont identiques, mais qu'ils *apparaissent* comme tels à la connaissance. (PRIETO, 1975, p. 83)

Les expériences de pensée qui jalonnent ce chapitre auront amplement montré que, indépendamment du fait d'accepter dans l'ontologie de telles entités, c'est un fait que nous pouvons concevoir des doubles. Le double n'est pas logiquement impossible. Aussi, la question que je souhaite aborder maintenant se rapporte aux enjeux proprement intentionnels en analysant la question des relations entre concevabilité et possibilités. Les fictions, nombreuses, qui nous proposent de concevoir des doubles, serviront de guide dans cette entreprise, qui constituera la seconde partie de ce mémoire. En outre, du point de vue du réaliste, l'individuation des personnages de fiction pose de nouveaux problèmes : la discernabilité spatiale devient inopérante, et le principe de non-pénétrabilité n'a évidemment plus cours. Les problèmes de doubles en deviennent d'autant plus tenaces.

Conclusion de la première partie

Quel bilan tirer de cette première partie ?

Le rapport problématique au double relève d'une forme d'hésitation ou d'erreur dans le dénombrement d'individus. L'exercice d'un décompte nécessite de ne pas comptabiliser deux fois la même chose, autrement dit, de saisir non seulement l'unité mais aussi l'unicité d'une chose. Nous avions en outre distingué deux questions propres à l'individuation métaphysique, celle de l'unicité, celle de l'unité.

Si l'identité est certainement efficace pour le dénombrement dans un monde figé, qu'en est-il de choses temporelles qui paraissent toujours différentes d'elles-mêmes parce que soumises à un incessant changement ? Il pourrait être prétendu que tous les existants sont des doubles d'eux-mêmes en ce qu'ils sont des *devenants* plutôt que des *demeurants*. De surcroît, les choses étant composées de parties, elles peuvent apparaître à la fois comme des unités et à la fois comme des multiplicités, ce qui risque de compromettre la possibilité d'un dénombrement objectif. Nous avons vu cependant que le changement n'est pas contraire à l'identité et que le nombre n'est pas une propriété des choses. Un premier résultat important est donc que le double n'est donc pas nécessairement entaché d'un rapport intrinsèque à la contradiction.

Nous avons vu plus précisément que le dénombrement d'individus implique principalement deux réquisits. Le premier est de saisir une chose comme un tout saillant sur le fond de son environnement, autrement dit, en tant qu'unité, comme quelque chose qui compte pour un. La seconde condition est de saisir la chose en question comme différente de toutes les autres choses de la même sorte, d'appréhender son unicité. Mais nous avons deux manières de satisfaire de tels réquisits : soit en considérant qu'ils doivent l'être de manière absolu, parce qu'ils relèvent de conditions métaphysiques, soit en considérant qu'ils relèvent de conditions pratiques d'identification seulement contextuelles.

Ainsi, nous avons distingué, dans cette première partie, deux conceptions de l'identité, l'une objectuelle, basée sur la recherche de critères d'identité intrinsèques aux objets, et l'autre conventionnaliste. Là où la première admet que l'identité s'applique de manière absolue aux objets, la seconde affirme que c'est le contexte qui détermine son application. Or, après avoir explicité un certain nombre de paradoxes de l'identité, nous avons conclu que ce n'est pas l'identité logique qui doit être relativisée aux sortes, mais que c'est l'application que nous faisons de l'identité logique qui doit être relativisée aux contextes, à nos intentions. Nous sommes ainsi passés de la question métaphysique d'individuation à celle, pragmatique, du jugement d'identité. Nous avons finalement souscrit à une conception conventionnaliste modérée de l'identité, en admettant que la

notion d'unité n'est pas seulement un fait linguistique, mais aussi psychologique.

D'autre part, creusant plus profondément le sillon des doubles ontologiques, nous avons vu que les paradoxes qui découlent de l'identité objectuelle nous confrontent à des doubles ontologiques. La longue tradition du réalisme des universaux semble s'échiner vainement à apprivoiser, par d'ingénieux critères d'individuation, les turbulents doubles ontologiques qu'elle génère d'elle-même. Nous avons précisé la nature de ces doubles, qui appartiennent à deux catégories, d'une part les doubles-mêmes, qui relèvent des théories modales de l'identité croisée, et d'autre part les indiscernables, ou ressemblants parfaits, qui relèvent des théorie qui individuent les particuliers par des propriétés vues comme des universaux. Et parmi les indiscernables, nous avons encore établi différents types : les répliques pures, les doubles purs, les coïncidents purs.

En particulier, nous avons vu que les problèmes de doubles apparaissent principalement en raison de l'idée que la référence d'une expression est indépendante de son mode de présentation, que le langage est susceptible de décrire les choses indépendamment du filtre de la représentation, qu'un nom relève de l'imitation objective d'une chose plutôt que de sa représentation contextuelle. C'est du moins ce qu'il ressort des philosophies qui rejettent la médiation des représentations dans l'accès à la référence, philosophies qu'on peut regrouper sous la dénomination générale de théories picturales du langage.

La discussion sur l'existence de mondes répétés nous a mené à distinguer les notions d'authenticité (axiologique) et d'originalité (temporelle), puis à considérer l'idée d'univers indiscernables, et de là, à celle d'individus intermondains indiscernables. Les solutions apportées au paradoxe modal de Chisholm par les philosophes qui souscrivent aux identités croisées passent par soit par la stipulation d'entités étranges, comme les particulier nus, des substances porteuses de propriétés sans en instancier elles-mêmes, ou soit par la stipulation d'haeccéités, propriétés individuantes. Ces solutions ne sont donc pas convaincantes, car elles ne semblent qu'ajouter des doubles fantomatiques, comme le sont les particuliers nus, aux doubles qui émanent des paradoxes de l'identité objectuelle.

Lorsque les philosophes distinguent sortalement la statue du bloc de marbre sous le prétexte que l'œuvre n'a pas les mêmes conditions de persistance que son constituant matériel, ils postulent une duplication ontologique des choses matérielles par des doubles de superposition. Un autre point de vue consiste à admettre, avec Hintikka, que la référence n'est jamais rigide, mais toujours relative à une modalité particulière. Dans ce cas, les doubles ne sont plus des entités ontologiques perturbantes, mais des représentations distinctes, des doubles épistémiques. La sémantique des mondes épistémiques d'Hintikka permet alors de rendre compte des doubles sans hypostase, tout en retenant l'idée d'un monde singulier, unique, d'une réalité objective.

Notre principale conclusion peut être résumée ainsi : d'une part, le double n'est pas logiquement impossible. D'autre part, les présumés doubles ontologiques ne sont finalement que des émanations de conceptions métaphysiques qui témoignent d'ailleurs de cette possibilité logique. Mais les doubles n'ont aucun poids dans l'ontologie, ils ne sont que des médiations épistémiques, des instances de représentations.

Dans ce qui suit, nous envisageons d'utiliser ces résultats dans le contexte de la fiction. Qu'en est-il des doubles en fiction ? Y a-t-il des entités fictionnelles indiscernables ? Quels sont les paradoxes du double qui sont propres aux fictions ? La seconde partie sera l'occasion d'évaluer

l'importance des doubles épistémiques en rapport à nos identifications imaginaires.

Deuxième partie

Vers un fictionnalisme intentionnaliste

Introduction à la seconde partie

Dans la première partie, nous avons commenté ce que serait une métaphysique du double en déployant diverses instances, les doubles-mêmes, les contreparties, les coïncidents fantomatiques et les indiscernables. Nous allons maintenant lui consacrer une étude du point de vue de sa production humaine à travers le domaine de la fiction. Nous poserons dans cette partie essentiellement deux questions, la première étant de savoir en quoi consiste au juste la production matérielle d'un double, la seconde étant de déterminer en quoi consiste le fait de concevoir ou d'imaginer un double. En particulier, nous nous confronterons aussi aux cas de récits où le double semble atteindre une forme paroxystique incitant à l'associer au contradictoire, dont la nouvelle *Lui !* de Maupassant peut faire figure de cas symptomatique. Nous indiquerons néanmoins des moyens d'interpréter cette fiction sans employer de mondes impossibles, c'est-à-dire en maintenant le cap sur l'idée d'une concevabilité du double même dans des cas qui semblent *a priori* désespérés.

De fait, il y a plusieurs raisons de s'intéresser à la fiction dans le cadre d'une enquête sur la notion de double.

Une première raison est heuristique. Un grand nombre de nos intuitions au sujet du double sont fondées sur des expériences littéraires. L'imaginaire est un terrain privilégié d'expérimentations pour mettre à l'épreuve notre capacité à concevoir le double. Il y a donc tout lieu d'enquêter sur la manière dont les auteurs parlent du double, et dont les lecteurs les conçoivent.

Cette première raison, qui vaut déjà pour elle-même, est encore justifiée par une autre, plus générale : la fiction stimule des représentations dont le contenu n'est pas réel. Elle permet alors de considérer les rapports que le double entretient avec l'irréel. Le double, c'est ce qui n'existe pas, ce qui s'oppose invariablement au principe de réalité. Si l'on convient de rapprocher ainsi le double de l'irréel, nous pouvons alors poser la question de savoir si nos représentations fictionnelles ne ressortiraient pas systématiquement au double en ce qu'elles seraient des reproductions mimétiques de la réalité.

Dans l'introduction de l'ouvrage de critique littéraire *Visage du double* qu'ils consacrent au double en 1996, Pierre Jourde et Paolo Tortonese (JOURDE & TORTONESE, 1996) commencent par constater que lorsqu'on évoque le double en littérature, il est difficile de savoir exactement de quoi l'on parle. Si la notion a été rebattue par l'exégèse littéraire, elle reste néanmoins difficile à circonscrire, et les auteurs ne s'accordent ni sur la question de savoir s'il s'agit d'une figure, d'une thématique ou d'un genre, ni sur le corpus qui en constitue l'extension. Il existe cependant, au sein de la critique spécialisée, un consensus : en littérature, la figure du double s'émancipe essentiellement à partir du XIXe siècle. Le succès que connurent les récits de doubles se renouvelle

d'ailleurs aujourd'hui, notamment à travers leurs formes cinématographiques [62].

Présente autrefois comme une figure encore secondaire, dont le motif sert une intrigue sans s'y réduire, le double prend au XIXe siècle, sous l'impulsion de Jean-Paul (PAUL, 1963) et de Gogol (GOGOL, 1966), la forme d'un genre romanesque autonome présentant respectivement la figure du *Doppelgänger* (double matériel du personnage) et celle de l'*Alter Ego* (double psychologique du personnage). Le genre en question est trouble, il suscite en effet des formes pathologiques de protagonistes, soit que le récit accrédite l'existence de personnages indiscernables quoique numériquement distincts, soit qu'il mentionne un personnage dont l'identité psychologique est scindée en courants multiples, voire qu'il mette à contribution l'auteur homodiégétique à l'intérieur même de son récit dans une confusion embrouillant les frontières entre réalité et fiction, forme contemporaine de l'autofiction [63].

On peut ici avancer une hypothèse historique orthogonale au présent travail : le succès de tels récits à partir du XIXe siècle est peut-être annonciateur de la crise naissante de la *mimesis*, dont Nelson Goodman (GOODMAN, 1972) s'est fait le partisan en rompant les amarres philosophiques entre les notions de représentation et de ressemblance. Constatons que déjà, dans le roman *La Méprise*, Vladimir Nabokov affranchissait en 1934 le double de toute similarité. Le personnage principal du récit, Hermann s'identifie en effet à un autre personnage, Félix, sans que ne se tienne entre eux aucune ressemblance frappante. Nabokov nous convie ainsi à admettre que l'identification à un double n'est pas nécessairement assujettie à une quelconque ressemblance qualitative. Voilà donc une piste nous guidant vers une caractérisation non mimétique du double.

Ces remarques préliminaires nous permettent de dégager trois façons différentes de poser la question du double dans son rapport à la fiction. Premièrement, on peut analyser la manière dont nos engagements intentionnels en fiction nous demandent de dupliquer la réalité en lui superposant, en la confondant, ou en lui substituant autre chose. Puisqu'il s'agit de représentations, la fiction permet de rendre compte de la façon dont celles-ci imitent ou non la réalité, autrement dit, du caractère mimétique du langage. Nos représentations mentales sont-elles des sortes de duplication d'objets réels par une forme quelconque d'imitation ?

Deuxièmement, on peut considérer la question de la nature des œuvres textuelles en rapport à la reproduction de leurs supports. C'est la question de l'authentification des copies qui survient ici, et la place de l'œuvre de fiction en rapport aux textes matériels. Enfin, troisièmement, on peut aussi évaluer la manière dont nous concevons des doubles à l'intérieur des récits, autrement dit,

[62] Parmi les films de doubles psychologiques, on peut mentionner les nombreuses adaptations cinématographiques du roman de Stevenson *Strange Case of Dr Jekyll and Mr Hyde*, de Victor Fleming (1941) à John Carl Buechler (2006), mais aussi des scénarios originaux comme ceux de *Two-Faced Woman* (1941) de George Cukor, de *Fight Club* (1999) de David Fincher, de *La Double Vie de Véronique* (1991) de Krzysztof Kieslowski, de Volte/Face (1997) de John Woo, de *Lost Highway* (1997) de David Lynch, de *Birdman* (2009) d'Alejandro González Iñárritu, de *Shutter Island* (2010) de Martin Scorsese, de *Split* (2016) de Night Shyamalan, de *Us* (2019) de Jordan Peele et Will Scarbrough. Parmi les récits cinématographiques en rapport à la ressemblance fortuite, à la gémellité ou au clonage, citons les trois films *The Wrong Man* (1956), *Vertigo* (1958) et *Psychose* (1960) d'Alfred Hitchcock, mais encore *Plein Soleil* (1960) de René Clément, *Kagemusha* (1980) d'Akira Kurosawa, *Dead Ringers* (1988) de David Cronenberg, *The Dark Half* (1993) de George A. Romero, *Single White Female* (1992) de Barbet Schroeder, *Grosse fatigue* (1994) de Michel Blanc, *L'adversaire* et *Les Jolies choses* (2001) de Gilles Paquet-Brenner, *Mulholland Drive* (2001) et la série *Twin Peaks* de David Lynch, *The Island* (2005) de Michael Bay, *Enemy* (2013) de Denis Villeneuve, ou encore *The Double* (2013), adaptation du roman de Dostoïevski par Richard Ayoade.

[63] A ce sujet, voir par exemple les études de (DOUBROVSKY, 2001 [1977]), (COLONNA, 2004) et (SCHMITT, 2010).

les phénomènes afférents à l'imaginaire du double, à l'identification des personnages.

Ces trois questions feront successivement l'objet des trois chapitres que contient cette seconde partie du mémoire.

Chapitre 3

La fiction comme activité de duplication du réel

Toute ressemblance avec des faits ou des personnes réels n'est (...) que pure ressemblance, ou obstination de la réalité. Dans ce roman, personne, de fait, ne doit se sentir visé. Personne, non plus, ne doit s'en sentir exclu, si d'une quelconque façon il y est impliqué.

Leonardo Padura, *Passé parfait*

Ce livre est une œuvre de fiction. Toute ressemblance avec la réalité est à imputer à cette dernière.

Jorge Volpi, *La Fin de la folie*

L'imagination n'est pas un état, c'est l'existence humaine elle-même.

William Blake, *Second livre prophétique*, Berger, p. 143

Certaines de mes œuvres de fiction sont vraies, au sens littéral du terme.

Philip K. Dick, conférence donnée lors du 2e festival de science-fiction à Metz, le 24 septembre 1977 (https://images.cnrs.fr/video/2267).

Les fantômes existent : la preuve, Pacman les mange.

Atari, publicité pour *Pacman*

L'objet de ce chapitre est de considérer le rapport de duplication que la fiction entretient avec la réalité. La fiction est-elle un double du réel ?

L'enjeu global est d'interroger la démarcation entre réalité, irréel et illusoire. Précisément, nous adopterons une stratégie visant à désolidariser le contenu sémantique des fictions de quelconques engagements ontologiques. Nous indiquerons comment la sémantique des mondes possibles peut servir à cette fin.

Le plan de chapitre suivra une progression allant des thèses de discrimination sémantique et ontologique aux thèses niant la discrimination ontologique pour ensuite proposer une perspective admettant la discrimination ontologique sans la discrimination sémantique. Ainsi, après avoir discuté, dans une première section, du problème de l'intentionnalité des objets fictionnels et de son ancrage dans la théorie mimétique du langage, je présente et critique, dans la section suivante, les diverses théories des objets exotiques et des théories constructionnistes. Dans la dernière section, je soutiens une autre orientation possible, favorisant une désolidarisation entre signification et référence.

3.1 Les théories ontologiquement démarquantes

3.1.1 La représentation est-elle un double spéculaire du réel ?

Si la littérature, et plus généralement la fiction, manifestent, depuis l'antiquité, tant d'accointances avec la figure du double, c'est sans doute parce que cette dernière symbolise la production artistique, dans la volonté de reproduire une partie de la réalité, en la représentant, c'est-à-dire, en la dupliquant. Car c'est là tout l'enjeu de la théorie mimétique de la représentation. Comme le souligne Wladimir Troubetzkoy, le double n'est jamais loin des mythes de Pygmalion et de Prométhée :

> Produit d'une imitation, d'une *mimésis*, le double naît ainsi de la même activité de l'esprit que l'œuvre d'art : le double est l'horizon de la *mimésis*, laquelle tend vers un idéal impossible, contradictoire et toujours fuyant, la production d'un double parfait de la réalité. Derrière le double se profilent ainsi les mythes de Pygmalion et de Narcisse, mythes de la confusion du réel et de l'imaginaire, de l'effacement de la frontière entre le sensible et l'intelligible, la vie et le marbre, ou la toile peinte, et en tant qu'entreprise orgueilleuse de substitution d'une réalité forgée de main d'homme à une réalité « divine », les mythes de Prométhée ou de Faust : l'artisan humain prétend rivaliser avec l'artisan divin. (TROUBETZKOY, 1996, p. 3)

Quels rapports la fiction entretient-elle au réel ? La fiction est-elle effectivement une sorte de duplication mimétique ? Faut-il considérer, au contraire, que les fictions ne nous apprennent rien de la réalité, qu'il ne s'agit que de fabulations trompeuses ? A moins qu'il ne s'agisse de voir nos discours factuels comme des sortes de fictions ? Plus généralement, quelle est la démarcation entre fictions, vérités et réalité ?

Parler de frontières implique inévitablement de discuter de la possibilité de leurs affranchissements. Selon Jean Bessière, la thématique du double est traditionnellement traitée par la distinction

entre monde nomologique et monde fantaisiste. Le double figure alors les passages, les transactions que ces mondes occasionnent, soit qu'un personnage de fiction visite un autre monde, comme Alice passant de l'autre côté du miroir, soit que le double apparaisse dans le monde propre à ce personnage. La lecture du double devient alors essentiellement celle d'une traversée.

> L'impossibilité usuellement reconnue au double fait entendre, lorsque le récit invite à reconnaître le double, que l'on a alors affaire avec quelque au-delà, ou avec quelque imaginaire qui distingue nettement deux types d'ordre - celui d'un monde, celui d'un autre monde où il y a des doubles, et qui suggère que l'on peut aller d'un monde à l'autre, ou que le double peut venir visiter ce monde. [...] Dans ce que l'on a coutume d'appeler l'inquiétante étrangeté du fantastique, le double inquiète moins par le redoublement qu'il expose que par ce passage et cette hiérarchie qu'il fait supposer. (BESSIÈRE, 2018, p. 19)

Le passage dans la fiction est-il une sorte d'abrogation de la réalité ? Gaston Bachelard constate, dans *L'Air et les songes*, que « percevoir et imaginer sont aussi antithétiques que présence et absence » (BACHELARD, 1943, p. 10). Est-ce à dire que l'imaginaire est inévitablement opposé au réel ? C'est bien l'opinion de Jean-Paul Sartre, conférant à l'imaginaire une fonction *irréalisante*. Dans un passage bien connu des dernières pages de l'*Imaginaire*, Sartre analyse en effet la forme rouge qui dépeint un tapis dans un tableau de Matisse de la manière suivante :

> Certains rouges de Matisse, par exemple, provoquent une jouissance sensuelle chez celui qui les voit. Mais il faut nous entendre : cette jouissance sensuelle, si on la considère isolément - par exemple, si elle est provoquée par un rouge donné en fait dans la nature - n'a rien d'esthétique. C'est purement et simplement un plaisir des sens. Lorsqu'on saisit, au contraire, le rouge sur le tableau, on le saisit, malgré tout comme faisant partie d'un ensemble irréel et c'est dans cet ensemble qu'il est beau. (...) C'est donc dans l'irréel que les rapports de couleurs de formes prennent leur sens véritable. (SARTRE, 1940, p. 364–365)

Le rouge qui est apposé sur le tableau par l'empreinte du pinceau du peintre doit ainsi être dédoublé d'un autre rouge, celui du tapis, un rouge 'laineux', qui justement, n'est pas le rouge du tableau. Sartre propose ainsi une caractérisation dualiste entre d'une part un objet matériel réel (les taches de couleur du tableau physique) et un objet esthétique irréel (le tapis dénoté par ces taches). Là où d'autres célèbrent le caractère créatif de l'imaginaire, Sartre lui oppose ainsi un aspect destructeur : l'imaginaire procède en anéantissant le réel. L'imaginaire ne peut apparaître qu'en se substituant à ce qui tient ici et maintenant. Il y a donc une forme de compétition pour occuper une certaine place spatio-temporelle.

Contrairement à Sartre, Clément Rosset souligne que ce qui s'oppose au réel, ce n'est pas l'imaginaire, mais l'illusoire. Ce qui les différencie, c'est la notion de précision, ou plus exactement, la notion d'objet. L'illusoire, par opposition à l'imaginaire, manque de précision, il manque d'objet. Pour prendre un exemple, cette distinction, c'est tout ce qui sépare la psychologie illusoire de Madame Bovary de l'imaginaire de Don Quichotte. Chez celle-ci, les désirs sont imprécis, tout est évanescent et toujours réajourné de loin en loin. Chez celui-là, tout y prend une précision nette, quoique trompeuse. Des moulins à vent servent de supports à l'imagination avec une précision aiguë, puisqu'ils sont pris pour une armée de géants. Mais sitôt que Don Quichotte se

heurte physiquement à la réalité, il se rend tout à fait compte de sa méprise. Il l'accepte d'ailleurs comme telle, faisant peser la responsabilité sur l'enchanteur Freston, qui le trompe constamment sur l'apparence de son environnement, à la manière du démon de Descartes. L'erreur reconnue, Don Quichotte passe alors à autre chose, ce qui ne l'empêchera pas, bien sûr, de se méprendre une nouvelle fois dans d'autres circonstances. Chez Bovary, c'est tout au contraire une forme d'illusion constante, jamais terminée. Même en connaissance de cause de son erreur de jugement, la même illusion persiste, et seul le suicide permet au personnage de rompre la chaîne continue de cette erreur permanente. Car les désirs de Bovary sont sans objet particulier. Ces deux illustrations littéraires montrent que l'imaginaire n'est donc pas complètement opposé au réel, l'un et l'autre traitant d'objets singuliers. En revanche, c'est bien le cas de l'illusoire, caractérisé par la perte d'objet, ou comme l'écrit encore Rosset, par le fait de ne penser à rien :

> Quand vous écoutez [le récit du mythe d'Œdipe], vous vous dites : non, ce n'est pas possible, les choses auraient dû se passer autrement. C'est là que vous êtes dans l'illusion ! Comment voulez-vous, en conservant un scénario réaliste, qu'un homme tue son père et épouse sa mère, sinon s'il ne les reconnaît pas, parce qu'il a été élevé par d'autres ? Voilà ce que j'appelle le double oraculaire : l'événement annoncé et attendu surprend quand il survient, comme s'il prenait la place d'un autre événement que pourtant on ne peut imaginer. À mon sens, la pensée du double est évanescente : il ne s'agit pas de former dans son esprit une illusion fort bien construite mais de s'imaginer qu'on pense quelque chose, alors que, en réalité, on ne pense rien. (ROSSET & BAYARD, 2015)

Nous devrions ainsi admettre que les fictions contribuent d'une certaine manière à la réalité. En tout cas, il faut bien reconnaître qu'elles entretiennent des ressemblances avec la réalité. C'est sans doute ce qui nous permet de nous immerger dans leurs mondes comme nous sommes immergés dans la réalité. Mais jusqu'à quel point ressemblent-elles au réel ? La fiction est-elle un double du réel ? Et si tel est le cas, n'est-ce pas conflictuel avec le caractère singulier du réel ? Ne sont-elles pas, dès lors, en trop, en concurrence avec le réel ? Dans *Impressions fugitives*, Rosset constate que le double n'est pas toujours l'indice du refoulement de la réalité. Car, à l'opposé des doubles psychologiques qui sont des doubles par refus de voir la réalité, il y a aussi des doubles qui consacrent la nature singulière du réel :

> Le double est sans doute le symptôme majeur du refus du réel et le facteur principal de l'illusion ; mais il existe certains doubles qui sont au contraire des signatures du réel garantissant son authenticité : telle précisément l'ombre qui vient à manquer à la femme sans ombre, tels aussi le reflet et l'écho. Ces doubles de 'seconde espèce' se caractérisent par une proximité par rapport à la réalité - humaine, vivante ou inanimée - qu'ils suivent comme son ombre, accompagnent comme son reflet, dupliquent comme son écho. Ces doubles-ci, qu'on pourrait appeler des doubles de proximité ou doubles mineurs, comme il y a des ordres mineurs, ne sont pas des prolongements fantomatiques du réel, mais des compléments nécessaires qui sont des *attributs* obligés [...]. S'ils viennent à manquer, l'objet perd sa réalité et devient lui-même fantomatique. (ROSSET, 2004, p. 10)

Si la fiction constitue un double du réel, ce n'est donc pas nécessairement sous le mode de l'illusoire. La fiction permet d'appréhender la réalité par l'imaginaire, qui n'est pas nécessairement antithétique au réel.

Quel genre de double « de proximité » constitue alors la fiction ? Nous avons vu dans ce qui précède qu'un double peut être conçu comme un indiscernable basé sur la relation de ressemblance, comme une contrepartie basée sur une notion de représentation ou encore comme un *même* basé sur une relation transmondaine d'identité. Considérons la première option. La fiction est-elle indiscernable du réel ? Nous pourrions avoir d'emblée un premier mouvement de recul : bien sûr que non ! Les moulins ne sont pas d'authentiques géants ! Nous savons heureusement, comme d'instinct, ce qui relève de la fiction ou non, et cela contribue d'ailleurs souvent à notre survie.

Pourtant, à y regarder de plus près, cette démarcation est bien plus fragile qu'un examen trop hâtif pourrait le faire croire. Umberto Eco remarque par exemple une certaine confusion entre personnes authentiques et personnages de fiction :

> [...] selon une enquête fiable, 25% des Britanniques croient que Winston Churchill, Gandhi et Charles Dickens étaient des personnages de fiction, tandis que beaucoup d'entre eux pensent que Sherlock Holmes et Eleanor Rigby ont vraiment existé (ECO, 2010, § 62)

La réalité entretient des relations ambiguës avec la fiction. À Vérone, on peut visiter la maison de Juliette, héroïne de la pièce de Shakespeare. À Londres, on peut pénétrer dans la résidence de Sherlock Holmes. Des touristes enthousiastes qualifient d'ailleurs l'édifice de "reconstitution très bien faite" [1]. Mais comment pourrait-on « reconstituer » quelque chose qui n'a jamais été ? Dans le même ordre d'idées, Eco raconte qu'il a été assailli de lettres provenant de lecteurs affirmant avoir localisé le monastère fictif du *Nom de la Rose* [2]. Inversement, la Russie fait partie du décor de *Guerre et paix* et Napoléon figure dans les récits de Tolstoi, de Balzac, de Stendhal, pour ne citer qu'eux. En outre, les récits de fiction contiennent d'innombrables énonciations purement factuelles.

Décrivant ce phénomène dans une conférence prononcée à Metz, en 1977, l'auteur américain Philip K. Dick fit remarquer que l'activité d'écrivain, en particulier d'écrivain de science-fiction, remettait en question l'idée que c'est l'auteur qui détermine ce qui relève ou non du fictionnel dans un texte.

> Nous avons la fiction qui parodie la réalité et la réalité qui parodie la fiction. Nous avons deux choses qui se chevauchent dangereusement, qui se brouillent dangereusement l'une l'autre. (...) On ne peut pas légalement contraindre un auteur à étiqueter correctement ce qu'il produit, comme une boîte de boudin dont les ingrédients sont mentionnés sur l'étiquette (...) on ne peut pas le forcer à déclarer quelle partie est vraie et laquelle ne l'est pas s'il ne le sait pas lui-même. (DICK, 1987, p. 21)

[1] Commentaires consultables sur http://www.tripadvisor.fr/Attraction_Review-g186338-d211907-Reviews-Sherlock_Holmes_Museum-London_England.html.

[2] « Après la publication de mon roman *Le nom de la rose* beaucoup de lecteurs m'ont écrit disant qu'ils avaient découvert et visité l'abbaye où j'avais situé mon histoire. [...] Dans cette même introduction je dis que j'ai trouvé un livre anonyme d'Athanasius Kircher dans une librairie ancienne de Buenos Aires. Récemment, c'est-à-dire presque trente ans après la publication de mon roman, un camarade allemand m'a écrit qu'il venait de trouver un livre de Kircher dans une librairie ancienne de Buenos Aires, et il se demande si par hasard il ne s'agirait pas du même antiquaire et du même livre que ceux trouvés dans mon roman. Il est inutile de dire que j'ai inventé tout à la fois l'architecture et l'emplacement de mon abbaye [...] et que, tant le livre mystérieux de Kircher que l'antiquaire plus mystérieux encore, étaient affaire d'invention » (ECO, 2010, § 1)

On raconte que, lors de la dite conférence, la suite des propos de l'auteur eu pour effet de laisser pantoise l'assemblée des auditeurs. Car il expliqua qu'une grande partie de ses récits n'étaient pas du tout imaginaires, mais qu'ils décrivaient des expériences dont il avait véritablement été le protagoniste. Il est difficile de déterminer si de telles affirmations résultaient d'une maladie mentale, d'un abus de psychotropes, ou s'il s'agissait d'actes illocutoires feints, autrement dit, d'une sorte de performance destinée à créer un effet de confusion en élargissant et brouillant la portée fictionnelle des récits. Dick était-il sincère ou s'agissait-il d'un effet littéraire métaleptique, voire de marketing auctorial anticipant ce qu'on appelle aujourd'hui le marketing de l'hyperréalité [3] ?

La réponse à cette question est sans doute complexe et n'intéresse de toute façon pas directement notre propos. En revanche, pour ce qui nous concerne, il importe de relever que la difficile démarcation entre fiction et réalité a pour corollaire l'incompatibilité des représentations subjectives que nous nous faisons du monde, en particulier dans le cas des maladies mentales. Car s'il y a un monde commun, nous le voyons alors chacun d'une façon différente. Mais que vaudrait un monde sans aucune inter-subjectivité ? Que vaudrait l'objectivité d'un monde si celui-ci ne pouvait être saisi qu'à travers le filtre de la subjectivité ? La question soulevée ici n'est pas de savoir s'il existe des mondes parallèles peuplés d'hécatonchires sous-marins ou de gnomes pourpres. Il s'agit plutôt de connaître la place de la réalité dans nos représentations, parfois incompatibles, quand bien même nous sommes censés vivre dans le même monde :

> Si la réalité diffère d'un individu à l'autre, pouvons-nous parler de réalité singulière, ou devrions-nous en fait parler de réalités plurielles ? Et s'il existe des réalités plurielles, certaines sont-elles plus vraies (plus réelles) que d'autres ? Que dire de l'univers d'un schizophrène ? Peut-être est-il aussi réel que le nôtre. Peut-être nous est-il impossible d'affirmer que nous sommes en contact avec la réalité et pas lui, peut-être devrions-nous plutôt dire : sa réalité est si différente de la nôtre qu'il ne peut pas nous l'expliquer et que nous ne pouvons pas lui expliquer la nôtre. Le problème, dans ce cas, c'est que si l'on expérimente des univers subjectifs trop différents, il se produit une rupture de la communication - et c'est *là* qu'est la véritable maladie. (DICK, 1987, p. 12)

La maladie mentale ne relèverait donc pas d'une impossibilité à atteindre la réalité, sauf à admettre qu'une telle maladie soit universelle. Le diagnostic ne porte pas sur la différence des représentations, mais sur la rupture de communication qu'elle implique.

J'aimerais illustrer l'idée, encore vague, d'une disparité des mondes épistémiques par un autre exemple littéraire, celui-ci d'obédience réaliste. Dans *White*, Bret Easton Ellis mentionne l'étonnement qu'il ressentit à la découverte, selon lui aussi tardive que brusque, de l'important clivage politique divisant sa nation en deux mondes irréconciliables :

> Au milieu de la campagne, j'ai remarqué que je ne lisais plus seulement le *Times* et que je ne regardais plus seulement CNN et MSNBC ; je jetais un coup d'œil aussi à FOX et à d'autres sources d'information conservatrices [...] et je me rendais compte avec une incrédulité corrosive que nous vivions dans deux mondes complètement séparés que je ne m'étais jamais préoccupé de remarquer auparavant, dans deux mondes qui étaient loin de se recouper, et je

[3] Voir par exemple (GRAILLOT & BADOT, 2005).

> me suis senti très naïf de ne pas avoir saisi ce contraste frappant plus tôt. (ELLIS, 2019, p. 182)

De fait, poser la question « qu'est-ce que la réalité ? », ou pire, prétendre y répondre, paraît relever d'une entreprise manifestement mégalomane. On pourrait alors se contenter d'une métaontologie, en commentant ce dont il est dit que « ça existe », voire une ontologie apophatique qui ne se prononcerait que sur ce qui ne peut pas être. Dick pourrait encore servir de guide ici. Car en dépit du solipsisme tenace dont témoigne cet écrivain qui ne cesse de suggérer dans ses œuvres la possibilité d'une multiplicité de mondes, et en dépit du peu de crédit qu'il confère à l'idée d'une réalité objective, Dick est pourtant à l'origine d'une caractérisation de la réalité très efficace : « La réalité est ce qui, lorsqu'on cesse d'y croire, ne s'en va pas. » (DICK, 1987) Le Réel est obstiné, son existence, même si elle est parfois refoulée, est pourtant irrévocable et totalitaire.

Si la réalité objective « ne s'en va pas », en revanche, bien d'autres mondes, imaginaires ou illusoires, vont et viennent. Interpréter, est-ce alors dédoubler la réalité d'un fait ou le contenu d'un discours par autre chose ? La question que nous voulons poser dans la suite de ce chapitre pourrait être résumée comme un long commentaire sur la citation, sans doute polémique, de Rosset :

> Il n'y a pas de « délire d'interprétation » qui tienne, puisque toute interprétation est délire. Ce qui différencie, socialement parlant, le fou de l'homme normal est un critère purement quantitatif et proportionnel : non un usage sain ou malsain de la logique, mais la quantité de temps, et l'ampleur du champ, qui sont attribuées à l'interprétation. (ROSSET, 1993, p. 21)

Ce ne serait donc pas la manière de raisonner qui mènerait à l'illusoire ou à l'imaginaire, mais un certain type d'exploitation de la logique afin de suggérer du Réel ce qu'il n'est justement pas. La conséquence, c'est que le monde factuel, les mondes illusoires et les mondes imaginaires ont en commun l'utilisation d'une logique. Ce qui les différencie relève essentiellement d'un degré d'interprétation. Mais le Réel, comme le rappelle Rosset, est singulier, c'est-à-dire sans double.

> J'appellerai ici réel, [...] tout ce qui existe en vertu du principe d'identité qui énonce que A est A. J'appelle irréel ce qui n'existe pas selon ce même principe, c'est-à-dire non seulement tout ce qui fait parade d'existence que sous le mode de l'imaginaire ou de l'hallucination, mais aussi ce qui semble bénéficier du privilège d'existence mais se révèle illusoire à l'analyse pour ne pas répondre rigoureusement au principe d'identité. (ROSSET, 1997, p.11)

Ce que la tautologie « A est A » gagne en exactitude, elle le perd en signification. « [...] la succession des tautologies, assurant chaque fois de telle chose qu'elle est cette chose même, ne dit finalement rien d'aucune chose. » (ROSSET, 1997, p. 18) Mais que la tautologie ne dise rien est justement ce qui fait son exactitude en tant que vérité logique : elle ne dépend pas des circonstances dont sont tributaires les vérités matérielles. En revanche, l'illusion ressortit aux jugements qui diffèrent des propositions analytiques. Aussi, l'illusion est-elle « toujours marquée au coin du double, d'une duplication hallucinatoire de l'unique qui constitue précisément sa "duplicité" » (ROSSET, 1997, p. 11). Il n'y a donc, conclut Rosset, qu'une seule manière de parler du Réel, qui est tautologique. Le Réel est idiot (*idotes*), au sens étymologique de « simple », c'est-à-dire singulier. Or, « si le réel est le simple, il manquera toujours à être reconnu, puisqu'une telle reconnaissance implique, de par l'insistance de son « re », l'appoint d'un autre que sa propre définition

exclut » (ROSSET, 1979, p. 17). Autrement dit, toute autre façon que tautologique de référer au réel relève de l'interprétation, du *double* qui, justement, est la marque de l'irréel. Il n'y aurait donc que deux alternatives : « ou parler de façon tautologique et dire vrai, ou parler en s'écartant de la tautologie et dire faux » (ROSSET, 1997, p. 14). Ce que vise Rosset, c'est de souscrire à l'idée que les vérités matérielles sont moindres vis à vis des vérités logiques.

Il y a au moins deux conceptions qui s'opposent radicalement à la caractérisation totalitaire que Rosset confère au Réel. D'une part le pluralisme épistémique de Nelson Goodman, et d'autre part l'ontologie modale de David Lewis. Chez Goodman, le monde réel n'est jamais donné de manière univoque, mais dédoublé en une infinité de versions. À une autre extrémité ontologique, le réalisme modal de Lewis affirme que tout discours étant ontologiquement dense, il réfère toujours à une réalité, quitte à ce que celle-ci échappe au domaine causal que nous habitons. Chacune des trois conceptions est ainsi fondée sur une caractérisation forte du principe de réalité : chez Rosset, l'identité singulière et purement tautologique du Réel, chez Goodman, la dualité sans fin des représentations, et chez Lewis, la ressemblance imitative sous de multiples aspects.

Dans *l'Air et les songes*, Bachelard fait remarquer qu'un « être privé de la *fonction de l'irréel* est un névrosé aussi bien que l'être privé de la *fonction du réel* », invitant ainsi, un peu plus loin, à « trouver une filiation régulière du réel à l'imaginaire » (BACHELARD, 1943, p. 14). C'est cette « filiation régulière » que nous nous proposons justement de rechercher. Ce chapitre pose donc la question de savoir si nos représentations, en particulier quand elles ressortissent à la fiction, sont des doubles du monde, soit qu'elles en sont indiscernables, soit qu'elles sont produites comme des contreparties. Les théories des trois auteurs cités donneront la tonalité, parce qu'elles proposent chacune de considérer le réel d'une manière paradigmatique, comme singulier ou multiple, comme objectif, comme représentation ou comme imitation. Dans ce qui suit, je ne chercherai pas une définition de la fiction, ce qui me paraît être peine perdue. Je tenterai par contre d'en circonscrire le domaine par certaines caractéristiques fondamentales : la fiction est une activité, produite par l'imaginaire, qui relève de l'irréel mais entretenant des rapports à la vérité. Je défendrai ainsi l'idée que la fiction n'est pas un double ontologique du réel - d'où sa part d'irréalité, mais un double épistémique - d'où son intérêt en termes de connaissance du réel. Je pense qu'il faut effectivement la distinguer de l'illusoire, non en ce qu'elle se rapporterait à une certaine réalité, mais en ce que cette dernière relève toujours du faux.

Dans *L'origine de la tragédie*, Friedrich Nietzsche distingue deux sens contradictoires de la notion de représentation. Le philologue constate que lors des célébrations des mystères dionysiaques, c'est parfois le dieu lui-même qui est dit apparaître aux bacchants, et que parfois, il ne s'agit pas de sa présence directe, mais de symboles qui le signifient. Ainsi, le premier sens de « représentation » relève de l'identité d'une chose bien présente, chose qui revient, qui se présente à nouveau, qui donc est *re-présentée*. Le deuxième sens signifie au contraire que ce n'est pas la chose elle-même qui revient, mais une autre qui prend sa place, qui tient place pour l'originale, par exemple en la symbolisant, chose qui est donc présente à la place de son modèle. C'est entre ces deux sens, de l'identité à la substitution, que se place la question de la fiction dans l'ordre de la représentation.

Pour poser une démarcation entre fiction et réalité, principalement deux types d'arguments se

présentent : ontologiques et sémantiques [4].

Considérons d'abord l'argument sémantique : fiction et réalité seraient discernables par leurs rapports respectifs aux vérités. Certaines affirmations concernant les personnages de fiction paraissent cependant relever de faits objectifs. Il semble par exemple tout à fait objectif que Sherlock Holmes pratique la profession de détective, et non celle de plombier. Mais l'« objectivité » qui touche les personnages de fiction est certainement différente de l'objectivité dont il est question lorsque je désigne l'hortensia qui est actuellement sur mon balcon. Sherlock Holmes n'existe pas. L'hortensia, oui. La plante existerait même si je n'étais pas présent pour constater son existence. Et elle existerait encore lorsque personne ne serait là pour en témoigner. Par comparaison, la réalité culturelle d'un personnage fictionnel portant le nom de Sherlock Holmes ne dépend pas non plus de mes seules intentions. Son objectivité dépend néanmoins de l'existence d'un grand nombre d'autres individus, à commencer par son créateur, Arthur Conan Doyle. Aussi, si l'on reconnaît au personnage d'Holmes une sorte d'objectivité, il s'agit du même genre d'objectivité que celle qui est attribuable aux billets de banque ou aux contrats de location. De ce point de vue, Holmes est un objet social : son mode d'objectivité est intentionnel et dépend d'un nombre important de règles complexes et de conventions culturelles. Une recherche ontologique au sujet des personnages de fiction consiste alors à enquêter sur leur mode particulier d'existence. Une telle recherche implique de considérer la manière dont ces objets sociaux s'intègrent à une ontologie plus globale, plus précisément, comment des objets purement intentionnels comme Sherlock Holmes se rattachent aux autres choses qui existent, comme les détectives, les deerstalker, les chiens, les meurtres et les villes. Il y a principalement deux manières de voir les choses pour caractériser les relations entre les personnages de fiction et les réalités physiques. Soit on considère que l'ontologie des objets physiques est insuffisante pour expliquer l'objectivité qui touche les personnages de fiction et qu'il faut la compléter par un réservoir d'objets plus exotiques, de type platonicien ou meinongien. Soit on prétend au contraire que la réalité des objets physiques est suffisante pour rendre aussi compte de l'objectivité des personnages. Cela revient alors à admettre une objectivité épistémique, et à rejeter une objectivité métaphysique.

La fiction devrait sans doute être discernée du réel par l'argument ontologique, en ce qu'elle décrit des choses qui n'existent pas. Nous n'avons pas d'expérience directe des choses qui sont décrites dans les fictions, alors que nous pouvons faire des expériences directes du réel [5]. On pourra toutefois objecter qu'une connaissance directe de choses réelles n'est pas toujours possible. Napoléon a existé, mais n'existe plus, et le prochain lauréat du prix Nobel de physique existera

[4]Pour Margaret MacDonald, la démarcation est linguistique. Bien qu'ils réfèrent à des entités réelles, des noms comme « la Russie » ou « Napoléon » deviennent néanmoins purement fictionnels aussitôt qu'ils apparaissent dans un texte de fiction. La Russie de *La Guerre et la Paix* n'est donc pas celle qui a été véritablement envahie par les troupes napoléoniennes. En outre, les publications littéraires fournissent généraient des indications métatextuelles assurant qu'il s'agit bien d'y lire une histoire de fiction et non un récit factuel. L'annonce liminaire traditionnelle qui avertit que « les personnages et les situations de ce récit étant purement fictifs, toute ressemblance avec des personnes ou des situations existantes ou ayant existé ne saurait être que fortuite » nous garantit du caractère imaginaire du récit. Mais je ne m'attarderai pas sur les critères linguistiques de discrimination qui, il me semble, ont été réfutés de manière convaincante, par exemple dans (RENAULD, 2014).

[5]Selon Russell, la connaissance directe est assez ordinaire, il s'agit de situations où l'expérience permet une saisie immédiate par les sens. Je laisse de côté la question de la fréquence de telles expériences, qui peuvent aussi être considérées comme des phénomènes rares liés à une réceptivité passive par des sortes de dés-inhibition des filtres épistémiques constitués par nos croyances.

certainement. L'étendue du réel ne doit certainement pas se limiter à ce qui existe ici et maintenant, mais doit aussi prendre en compte tout le passé (ce qui a été il y a une seconde et il y a mille ans) et tout le futur (ce qui sera dans une seconde et ce qui sera dans mille ans). Mais si l'on doit ainsi reconnaître une certaine contingence de la connaissance directe du réel, en revanche, la connaissance directe des « objets » de fiction est quant à elle nécessairement impossible. Les personnages de fiction n'existent pas, n'ont pas existé et n'existeront jamais.

Rosset a-t-il alors tort de ne pas souscrire à l'incompatibilité proclamée par Sartre entre imaginaire et réalité ? Je ne le crois pas, parce que la fiction n'est pas nécessairement un double de remplacement. Mais il semble alors que l'argument de Rosset doive être amendé. Car la fiction relève bien de l'irréel et il ne faut pas nécessairement la doter d'un coefficient ontologique pour lui éviter un dénigrement cognitif ou pour lui faire tenir un rôle d'usurpation à l'égard du réel. Nous devrons donc avancer en gardant en vue deux pôles, l'un consistant à admettre que les fictions sont irréelles, l'autre consistant à admettre qu'elles ont bien un intérêt cognitif. Il me semble alors que la question peut être posée comme le fait Manuel Rebuschi (REBUSCHI, 2000) : comment désolidariser l'objectivité de l'ontologie ? Comment admettre des propositions vraies à propos de choses qui ne sont pas, et cela, sous aucune forme ? En fiction, il peut y avoir du vrai, quoique, à proprement parler, il n'y ait rien de réel. N'est-ce pas d'ailleurs le cas des nombres *imaginaires*, et plus généralement de tout l'édifice des mathématiques ?

3.1.2 Théorie mimétique et représentations

Mimésis et diégésis

Platon a focalisé de manière durable la philosophie des fictions sur l'activité mimétique décrite comme la production de doubles de substitution par une imitation imparfaite de la réalité. C'est par le terme de *mimésis* que Platon qualifie la représentation des choses par le moyen du langage. Le terme est essentiellement péjoratif, se rapportant à l'idée d'usurpation. Deux livres de *La République* (PLATON, 1955, p. 935) sont consacrés à ostraciser la fiction hors des murs de la cité idéale. La première critique est située dans le livre 3. Platon y distingue trois modes d'énonciations (*lexis*) : d'une part le mode de l'exposition pure et simple ou *diègèsis* (diègèsis) et d'autre part le mode de l'imitation ou *mimésis* (mimésis), et enfin un mode mixte.

Le premier mode consiste à raconter une histoire à la troisième personne comme le fait Homère dans le passage suivant du début de l'*Iliade* :

> Pour racheter sa fille au prix de grands trésors, Chrysès était venu vers les sveltes vaisseaux de la flotte achéenne, et, sur un sceptre d'or, de l'archer Apollon portant les bandelettes, il priait les ariens, mais surtout les deux chefs de guerre, fils d'Atrée. (HOMÈRE, 1955b, p. 93)

Comme on le voit, dans ce texte, le narrateur n'intervient pas dans le récit, il se contente d'en relater les circonstances comme s'il avait été le témoin passif *mais absent* des circonstances. Il est un témoin omniscient. Sa relation narrative est *extradiégétique*.

Le second mode d'énonciation consiste à faire comme si les faits décrits étaient racontés par quelqu'un d'autre que celui qui les récite. Par exemple, dans la suite du texte de l'Iliade, Homère décrit les circonstances comme s'il était Chrysès lui-même :

> Atrides, et vous tous, Argiens aux belles guêtres, puissent les Immortels, habitants de l'Olympe, vous donner de piller la ville de Priam, puis de rentrer dans vos demeures sains et saufs ! Mais rendez-moi ma fille, acceptez sa rançon, par égard pour l'Archer Apollon, fils de Zeus. (HOMÈRE, 1955b, p. 93)

Platon commente ainsi ce procédé :

> Le poète dit des paroles prononcées comme si lui-même était quelqu'un d'autre. Mais celui qui s'assortit lui-même à un autre, soit pour l'intonation de la voix, soit pour l'apparence extérieure, nous disons qu'il « imite » cet autre auquel il se sera assorti (PLATON, 1955, p. 945)

Sous le mode mimétique, le narrateur emprunte l'identité d'un autre. L'activité d'imitation consiste donc à feindre le déroulement d'une aventure qui n'est pas *réellement* vécue. Celui qui raconte décentre sa position en se plaçant du point de vue de celui qui est censé vivre l'intrigue. Dans le passage qui précède, Homère se distingue d'un narrateur omniscient exposant simplement des faits parce qu'il commente des événements comme s'il était autre que lui-même. L'imitation procède en une forme de refoulement de la personnalité qui cède sa place à l'incarnation d'une autre. C'est ainsi qu'au fil de son récit, Homère s'attribue diverses identités d'emprunt. Bien sûr, il n'y a pas une véritable usurpation d'identité en ce que cette manière de simulacre n'a pas pour volonté de tromper les récepteurs de la fiction. Il ne s'agit que d'un procédé narratif destiné à rendre plus expressif un passage par un procédé ressortissant à l'empathie.

Platon utilise la dichotomie entre les deux modes d'énonciation diégétique et mimétique pour établir une division théorique des arts poétiques. Les dithyrambes sont la forme paradigmatique du mode diégétique. La comédie et la tragédie relèvent quant à elles de l'exposition mimétique. Et comme nous l'avons vu avec l'Iliade, les épopées ressortissent d'un mode mixte en ce qu'elles contiennent à la fois des passages diégétiques et des passages mimétiques.

Cette distinction formelle est assortie chez Platon d'une distinction axiologique. Plus un contenu est lié à l'imitation, plus il est considéré comme pernicieux. Ce que Platon reproche en effet au procédé narratif mimétique, c'est qu'il incite à la duplication des rôles des citoyens dans la Cité. Celui qui prétend imiter les autres hommes, celui qui prend provisoirement la place d'un autre, celui-là indique qu'il pourrait occuper toute fonction, se substituer à tout un chacun. Mais pour Platon, la cité idéale est celle où « il n'existe pas d'homme qui soit un double, ni qui soit un multiple de lui-même, puisque chacun ne fait qu'une seule chose » (PLATON, 1950b, p. 951). Dans la cité idéale, chacun doit être, dit Platon en substance, cantonné à une seule fonction. On doit ainsi y trouver « le cordonnier, cordonnier, et non pas capitaine de navire » (PLATON, 1950b, p. 951). Quant à celui qui sait imiter les rôles, il n'aura qu'*à aller voir ailleurs s'il y est* :

> [...] un homme ayant le pouvoir, conditionné par un talent, de se diversifier et d'imiter toutes choses, un tel homme, s'il parvenait à entrer dans notre Cité avec l'intention d'y présenter au public et sa personne et ses poèmes, nous lui ferions profonde révérence comme à un personnage sacré, hors pair, délicieux, et d'autre part, nous lui dirions qu'il n'y a pas chez nous d'homme comme lui dans la Cité, et qu'il n'est point permis qu'il en vienne à s'y produire ; nous l'éloignerions en direction d'une autre Cité, après avoir sur son chef répandu du parfum et l'avoir couronné de laine ! (PLATON, 1950b, p. 951)

Homère, qui a le tort de vouloir faire passer les vessies fictionnelles pour des lanternes factuelles est donc banni de la cité platonicienne.

La deuxième critique platonicienne de la fiction intervient dans le dixième livre de *La République*. Il ne s'agit plus d'examiner le cas de la narration, mais celui de la représentation picturale. La critique est bien connue : la représentation picturale rejoue à un niveau ontologique plus bas la différence entre les idées intelligibles et les choses sensibles. Les formes intelligibles sont caractérisées par le fait d'être simples, non mixtes, d'être exclusivement ce qu'elles sont. Le beau, ce n'est pas ce qui est aussi un cheval, une cuillère ou une jeune fille, c'est ce qui *seulement* beau.

Platon considère le cas de la représentation picturale d'un lit selon une hiérarchie ontologique distinguant trois modes d'instanciation : le lit comme idée générale du lit, le lit comme forme sensible particulière produite par le menuisier et enfin le lit comme représentation par un peintre [6]. Du premier type (idée intelligible) au dernier (représentation mimétique), on passe d'une singularité intelligible à une forme sensible qui en est une duplique parmi d'autres (le lit en fer, le lit en bois, etc.), puis à l'imitation d'une de ces copies. Du monde intelligible au monde des représentations mimétiques, on a donc un dénivelé ontologique de deux niveaux. Dans la hiérarchie platonicienne, la représentation du lit s'avère être une forme doublement dégénérée, puisqu'elle est une copie d'une copie. Or plus une chose a de doubles, moins elle a de poids ontologique. Puisque le peintre n'imite pas les choses intelligibles elles-mêmes, mais seulement les choses sensibles, sa représentation éloigne l'esprit de la connaissance réelle, qui doit toujours être dirigée par la seule quête des formes intelligibles. Le peintre produit donc des simulacres, des multiplicités de multiplicités. La peinture procède en des sélections impropres de qualités qui nous éloignent des pures formes intelligibles encore plus que les sélections produites par les artefacts. Puisque l'objectif de la philosophie platonicienne est essentiellement sélectif en regard de l'incitation à l'imitation des formes intelligibles, la représentation picturale est rejetée comme une sorte galvaudée d'imitation.

Le reproche platonicien à l'égard de la représentation mimétique ne provient pas du caractère imparfait de sa reproduction, qui n'est jamais une reproduction exacte, comme nous l'avons déjà évoqué dans la première partie en examinant le cas d'un éventuel double parfait de Cratyle. Quand bien même un artiste produirait une œuvre dont l'apparence serait quasi-indiscernable de la réalité, comme ces fameux raisins peints par Zeuxis qui trompaient l'appétit des oiseaux, le caractère *spéculaire* de la représentation mimétique ne devrait pas valoir de caractère *spectaculaire*. La surface d'un miroir, rappelle d'ailleurs Platon, reproduit à l'exactitude les apparences des choses qui viennent s'y refléter mieux qu'un peintre ne pourrait jamais y parvenir. La technicité mimétique du peintre ne mérite donc guère de louanges particulières.

De fait, l'imitation narrative et picturale sont rapportées à la même erreur fondamentale : elles ne sélectionnent pas les bons sujets. Telles les ombres projetées sur la paroi de la caverne dans la fameuse allégorie, la représentation mimétique ne donne à voir que l'apparence des choses, mais non les choses elles-mêmes. Il y a donc, chez Platon, une évidente démarcation entre ce qui relève de la fiction et du réel, la première étant une reproduction imitative de la seconde, mais une reproduction condamnée à une grossière inexactitude en ce qu'elle ne choisit pas le bon modèle. La conception mimétique de la fiction, et plus généralement de la représentation est donc ancrée à

[6] Pour illustrer ceci, on pourra penser à l'œuvre créée par Joseph Kosuth en 1965, intitulée *One and Three Chairs*, qui servit de manifeste à l'art conceptuel. L'artiste a exposé, contre un mur, une étiquette mentionnant la définition du mot « *chair* » assortie d'une véritable chaise en bois et d'une photographie de cette même chaise.

partir de Platon sur une relation de ressemblance incomplète. La fiction imite des choses, mais ne saurait guère se substituer à la réalité qui lui échappe. Cette conception a pour corollaire que les œuvres ne sont pas elles-mêmes des réalités, mais des simulacres de la réalité. Elles ne peuvent avoir un rapport au vrai, parce qu'elles ne se rapportent pas aux choses (celles ayant un véritable poids ontologiques), mais à une illusion. Chez Platon, l'imaginaire est donc complètement rivé à l'illusoire.

La fiction éloignée de la vérité

Pour Platon, *mimesis* et *logos*, autrement dit, fiction et science, sont donc deux types irréconciliables de discours. Après Platon, Frege a enfoncé le clou en ancrant à son tour la notion de vérité à la notion de dénotation, liant complètement l'ontologie à la sémantique, réduisant la notion d'objet à celle de référence d'un terme linguistique. Si Frege admet une sémantique double, évaluée selon les modes du sens (*Sinn*) et de la dénotation (*Bedeutung*), le premier mode est conçu comme tout aussi objectif que la dénotation, relevant ainsi d'un troisième monde par delà le monde des réalités physiques et par delà le monde des représentations subjectives :

> Les pensées ne sont ni des choses du monde extérieur ni des représentations. Il faut admettre un troisième domaine. Ce qu'il enferme s'accorde avec les représentations en ce qu'il ne peut pas être perçu par les sens, mais aussi avec les choses en ce qu'il n'a pas besoin d'un porteur dont il serait le contenu de conscience. Telle est par exemple la pensée que nous exprimons dans le théorème de Pythagore, qui est vraie de manière intemporelle et indépendamment du fait que quelqu'un la tienne pour vraie ou non. (FREGE, 1971b, p. 12)

En sortant la signification de l'esprit, Frege la rabat conséquemment sur l'ontologie. Ainsi conçoit-il qu'énoncer des propositions vraies, c'est toujours parler de choses réelles.

Les termes qui dénotent des particuliers peuvent s'avérer simples, comme par exemple « Napoléon ». Ils peuvent aussi être complexes, comme par exemple « le vainqueur de la bataille d'Austerlitz ». Mais les uns et les autres sont considérés comme relevant de la catégorie des noms propres au sens logique du terme. Certains sujets grammaticaux, en dépit du fait qu'ils prennent bien l'apparence de noms propres, ne dénotent pourtant rien. Dès lors, Frege est amené à concevoir qu'un terme comme « Sherlock Holmes », qui semble de prime abord être un nom propre, ne dénote cependant aucun particulier. C'est donc que le terme n'est pas un authentique nom propre, au sens logique du terme. « Sherlock Holmes » est un nom vide.

Rappelons que selon le principe de compositionnalité, la signification d'une proposition, autrement dit, sa valeur sémantique, est tributaire de la signification de ses composantes. Un énoncé contenant nom vide est donc non-dénotant, autrement dit, il est sans valeur de vérité. L'inclusion d'un nom vide implique inévitablement l'absence de valeur de vérité de la proposition qui le contient.

> La proposition « Ulysse fut déposé sur le sol d'Ithaque dans un profond sommeil » a évidemment un sens, mais il est douteux que le nom d'Ulysse qui y figure ait une dénotation ; à partir de quoi il est douteux que la proposition entière en ait une. Il est certain toutefois que si l'on prend sérieusement cette proposition pour une proposition vraie ou fausse, on attribue, ce faisant, une dénotation au nom d'Ulysse, en plus du sens. (FREGE, 1971b, p. 109)

La fiction relève ainsi, uniformément, de représentations qui sont en deçà de la vérifaction, qui ne sont ni vraies ni fausses, parce qu'elles ne dénotent rien. Selon Frege, tout ce qui signifie quelque chose existe d'une manière ou d'une autre. Le monde est reflété dans le langage. Il s'agit, comme chez Platon, d'une théorie mimétique du langage.

En refusant de conférer une valeur de vérité aux énoncés contenant des noms vides, la sémantique de Frege échoue à rendre compte d'expressions triviales comme « Hercule Poirot est Hercule Poirot » et de disjonctions du type « Sherlock Holmes a un grand-père daltonien ou n'a pas de grand-père daltonien », de propositions existentielles négatives vraies comme « Hercule Poirot n'existe pas », ou encore de vérités fictionnelles internes comme « Hercule Poirot est un détective ». En particulier, la proposition exprimée par « Hercule Poirot est Hercule Poirot » nécessite un traitement différent de la proposition exprimée par « Emmanuel Macron est Emmanuel Macron ».

La sémantique de Frege aboutit ainsi à l'articulation de deux thèses, l'une ontologique, l'autre épistémique :

Le domaine de référence des termes fictionnels est vide.

Le contenu épistémique des propositions fictionnelles est vide.

Il n'en demeure pas moins que Frege conçoit qu'un nom vide puisse manifester une forme de signification, à savoir, celle qui relève du sens. En outre, la double valeur sémantique du sens et de la dénotation est étendue aux propositions elles-mêmes. Parallèlement au principe de compositionnalité des dénotations, il existe ainsi chez Frege un principe de contexte qui stipule que la phrase est l'unité sémantique de référence de la signification. Cela explique pourquoi deux expressions comme « Napoléon est le vainqueur de la bataille d'Austerlitz » et « Napoléon a perdu la bataille de Waterloo » co-dénotent la vérité, mais diffèrent dans leur manière d'y accéder. Des expressions, même si elles demeurent sans valeur de vérité, contribuent nécessairement à la signification de la proposition qui les contient. Ainsi, un nom vide complexe comme « la planète située entre le soleil et Mercure » possède, sinon une dénotation, du moins un sens conféré par chacune des dénotations de ses constituants. Le principe de contexte explique donc aussi pourquoi « Sherlock Holmes est un détective » n'a pas la même signification que « Sherlock Holmes est Sherlock Holmes » bien que ces deux expressions soient sans dénotation.

En autorisant ce clivage, Frege semble ainsi ouvrir la voie à une conception intensionnelle des fictions. Car, par la médiation du sens, Frege construit une sémantique qui n'est pas exclusivement référentielle. Mais soulignons que le sens d'une expression reste encore extrêmement tributaire de la référence, puisqu'il est défini comme mode d'accès à cette dernière. Un tel attachement, trop étroit, entre la signification et la réalité semble compromis au sens où, comme le fait remarquer Gareth Evans, il devient difficile de considérer que les noms vides puissent vraiment avoir un sens : « It is really not clear how there can be a mode of presentation associated with some term when there is no object to be presented » EVANS, 1982, p. 22. Comment pourrions-nous avoir une forme d'accès à la dénotation en l'absence de celle-ci ?

Du point de vue de Frege, la notion de vérité demeure complètement corrélée à la référentialité. Sous cette perspective, les fictions ne jouent aucun rôle épistémique, elles n'entretiennent aucun

rapport à la connaissance, elles se rapportent seulement à l'esthétique, à la sensibilité. Elles sont, au mieux, une affaire de style :

> Si l'on écoute une épopée, outre les belles sonorités de la langue, seuls le sens des propositions et les représentations des sentiments que ce sens éveille tiennent l'attention captive. À vouloir en chercher la vérité, on délaisserait le plaisir artistique pour l'examen scientifique. De là vient qu'il importe peu de savoir si le nom d' « Ulysse », par exemple, a une dénotation, aussi longtemps que nous recevons le poème comme une œuvre d'art. (FREGE, 1971b, p. 109)

Aussi, si l'avantage de la théorie de Frege est de reconnaître la vacuité ontologique du domaine des termes fictionnels, elle le fait cependant au détriment d'un quelconque apport épistémique. La fiction n'occasionne donc pas une duplication du réel. Ainsi, en refusant de concéder une valeur de vérité aux affirmations contenant des noms vides, Frege tend à assimiler l'imaginaire à l'illusoire.

Le problème des assertions existentielles négatives

La théorie de Frege se heurte également au problème des assertions existentielles négatives vraies comme « Sherlock Holmes n'existe pas ». De fait, il arrive fréquemment, dans nos discussions ordinaires, que nous rejetions l'hypothèse d'une entité n'ayant pas de substance concrète, comme par exemple le croque-mitaine du placard à balais, l'âme, les statues, les anges ou Dieu. Bien souvent, notre interlocuteur, s'il est convaincu du contraire (précisons que c'est plus fréquent pour Dieu et pour les anges que pour le croquemitaine), nous demande alors de prouver l'inexistence de la chose qu'il attribue, quant à lui, au domaine de ce qui est. Cet état de fait pourrait paraître assez étonnant, voire irrationnel, puisqu'il sollicite la preuve de l'inexistence d'une chose, au demeurant invisible et immatérielle. En outre, on attendrait une justification plutôt du côté de celui qui prêterait réalité à de telles putatives entités, ces dernières n'ayant aucune manifestation tangible pour témoigner de leur présence.

Abstraction faite de l'éventuelle mauvaise foi de l'interlocuteur, cette pratique discursive semble relever d'une forme d'argumentaire dans lequel celui qui rejette les entités non matérielles souffre au départ d'un désavantage logique. En effet, les affirmations existentielles négatives vraies comme « les basilics n'existent pas » paraissent impliquer d'emblée une contradiction. Car en rejetant l'existence d'une chose, il semble que nous admettions paradoxalement que nous parlons bien de *quelque chose*, bien que ce quelque chose soit privé d'existence. Le désavantage de celui qui se refuse à admettre une certaine entité est souligné par Quine dans un passage célèbre de « On What There Is » où il lui donne le nom de problème « de la barbe de Platon ». Un dialogue oppose deux protagonistes. Le proposant admet l'hypothèse d'une entité que l'opposant rejette. Le proposant exprime sa différence d'opinion de la manière suivante : « il y a quelque chose que vous refusez de reconnaître ». L'opposant pourrait rectifier cette formulation, qui semble d'emblée donner raison au proposant, de la manière suivante : « il n'y aucune entité du genre que vous proposez que je puisse reconnaître ».

En ancrant la vérité à l'ontologie, toute théorie sémantique qui affronte la question de la signification des noms vides doit faire face aux deux contraintes contradictoires suivantes :

1. Toute représentation vraie est représentation d'un objet.

2. Il y a des représentations vraies n'ayant pas d'objet.

Si Sherlock Holmes n'existe pas, nous devrions alors assumer la vérité de la proposition « Sherlock Holmes n'existe pas ». Or, du point de vue de Frege, affirmer que « Sherlock Holmes » ne dénote rien de réel, c'est affirmer que la proposition exprimée par « Sherlock Holmes n'existe pas » est sans valeur de vérité. On voit que l'ancrage du vrai au réel bute directement sur l'entreprise de rejeter les fictions hors de l'ontologie.

Serait-ce là l'indice d'un faux problème ? Une tentative de réponse consisterait à simplement refuser de donner un sens à de telles questions. Notre méfiance à l'égard de la métaphysique, et notre positivisme pourrait en effet nous en détourner. De telles questions ne sont-elles pas, au fond, des absurdités ? À quoi bon se demander en quoi consiste l'(in)existence de « Sherlock Holmes » ? À en croire Rudolf Carnap, posées ainsi, les questions d'existence n'ont strictement aucun sens et, de ce fait, n'attendent pas vraiment de réponse. Un tel énoncé d'existence n'est ni faux comme l'est « 4 est un nombre premier », ni contradictoire comme l'est « mon frère est fils unique ». Il n'est pas non plus intrinsèquement absurde comme « l'atmosphère augure des nombres verts », parce qu'il semble bien porter une information. Mais c'est une promesse toujours ajournée. L'énoncé n'est ni vrai, ni faux, ni contradictoire, ni absurde. Il est pris dans un « idiome matériel » : il est mal formulé. Il faut cependant le traduire dans l'idiome logique pour en faire saillir l'exigence première : on voit alors qu'il n'exprime pas une thèse mais une préférence pour un certain cadre linguistique. Ainsi, l'expression « ce charançon est une chose », retraduite dans l'idiome formel, consiste à affirmer que « le mot "charançon" est un nom de chose », autrement dit, à privilégier le langage des objets physiques par rapport au langage des expériences sensorielles.

Comparons la question de l'existence des personnages de fiction à celle des nombres. Y a-t-il des nombres ? Nous pouvons fournir la réponse suivante : il est incontestable qu'il existe des nombres premiers. Donc les nombres existent. Pourtant, en passant de la prémisse à la conclusion, il semble que nous ne tenons plus tout à fait le même discours. Nous glissons brutalement d'une affirmation qui n'a pas de conséquence ontologique à une affirmation métaphysique qui, elle, revendique l'existence d'une entité exotique. Or Carnap fait justement remarquer que, pour éviter ce type indu de glissement, il est tout à fait essentiel de bien démarquer deux ordres de questions, qui doivent être différenciées par un certain crible. À cet effet, Carnap distingue les questions internes des questions externes. Les premières sont caractérisées par le fait qu'elles s'inscrivent dans un cadre linguistique précis, cadre qui dote les termes d'une signification. Ainsi, parvenir à comprendre la question « y a-t-il des personnages de fiction » nécessite-t-il :

> [...] de reconnaître une distinction fondamentale entre deux genres de questions concernant l'existence ou la réalité d'entités. Si quelqu'un veut parler dans son langage d'un nouveau genre d'entités, il doit introduire un système de nouvelles façons de parler, assujetties à des règles nouvelles ; nous parlerons alors de construction d'un *cadre* (*framework*) linguistique pour les nouvelles entités en question. Et maintenant, nous devons distinguer deux genres de questions d'existence : d'abord, les questions portant sur l'existence de certaines entités de ce nouveau genre *à l'intérieur du cadre* ; nous les appelons les *questions internes* ; en second lieu, les questions portant sur l'existence ou la réalité *du système d'entités pris comme un tout*, que nous appelons *questions externes*. (CARNAP, 1997, p. 315)

Les questions internes sont donc les questions qui sont propres à un système linguistique donné. Il s'agit de questions comme « y a-t-il un morceau de papier blanc sur mon bureau ? », ou comme « y a-t-il un nombre premier entre 10 et 12 ? ». La signification de telles questions est entièrement gouvernée par le système linguistique qui prévoit ou non l'existence des entités concernées. La réponse peut être donnée en recourant à l'analyse logique du cadre linguistique et éventuellement en faisant appel à l'observation empirique. Les questions internes peuvent ainsi être réglées de manière triviale : elles sont décidables de manière empirique (il y a bien une feuille sur mon bureau) ou de manière logico-mathématique (il y a bien un nombre premier entre 10 et 12). Demander s'il existe un personnage du nom de Sherlock Holmes, c'est répondre par l'affirmative en découvrant que le corpus d'un certain auteur, Conan Doyle, contient bien un personnage de ce nom.

En revanche, demander si Sherlock Holmes existe en dehors du cadre de l'histoire de la littérature est proprement insignifiant, et n'appelle, de ce fait, aucune réponse. Carnap souligne que les questions comme « y a-t-il des personnages de fiction » sont le type même de questions que posent les métaphysiciens.

> Nobody who meant the question 'Are there numbers' in the internal sense would either assert or even seriously consider a negative answer. This makes it plausible to assume that those philosophers who treat the question of the existence of numbers as a serious philosophical problem and offer lengthy arguments on either side, do not have in mind the internal question. (CARNAP, 1950, p. 209)

De telles questions ne sont ni celles de l'homme de la rue, ni celles du scientifique (CARNAP, 1950, p. 207). Elles sont des questions sorties de leur contexte théorique, elles sont « externes ». Elles n'ont alors, dit Carnap, aucune signification, parce qu'elles échappent par définition à un cadre linguistique qui pourrait leur conférer la décidabilité attendue. Il s'agit donc moins de véritables questions que de pseudo-questions.

Le déroulement successif d'arguments contradictoires dans l'histoire de la philosophie devrait d'ailleurs, prévient Carnap, nous avertir du caractère non fondé des questions externes, questions auxquelles « les réalistes répondent par l'affirmative, les idéalistes subjectifs par la négative, et la controverse se poursuit depuis des siècles sans être jamais résolue » (CARNAP, 1997, p. 316). Les questions externes restent sans réponse définitive, pour la simple raison qu'elles ne peuvent en avoir. L'échec est assuré par le fait que ces questions sont tout simplement mal conçues. Exister, c'est posséder une signification dans un système linguistique. Mais poser des questions sur la réalité des systèmes eux-mêmes, c'est demander plus que ce à quoi l'on peut répondre :

> Être réel au sens scientifique veut dire être un élément du système ; ce concept ne peut donc être appliqué de manière sensée au système lui-même. Celui qui pose la question de la réalité du monde des choses lui-même a peut-être en tête non pas une question théorique, comme ses formulations semblent le suggérer, mais une question pratique, une décision d'ordre pratique concernant la structure de notre langage. Nous avons à choisir si, oui ou non, nous acceptons et utilisons les formes d'expression du cadre en question. (CARNAP, 1997, pp. 315-316.)

Demander *simpliciter* s'il y a des personnages de fiction est donc une question qui relève d'un pseudo-énoncé, une question « dépourvue de contenu cognitif » (CARNAP, 1950, p. 31). On

admettra donc que l'importance d'une telle question ne relève que d'un sens pratique, en attendant par exemple une réponse qui relève seulement de l'histoire littéraire, et non de l'ontologie. Bien sûr, la question de savoir s'il faut ou non admettre un certain cadre théorique reste significative. La véritable question que soulèvent les questions métaphysiques est celle de la préférence pour un cadre linguistique. Autrement dit, il s'agit d'évaluer l'opportunité d'adopter un cadre plutôt qu'un autre. En soi, ce genre d'activité ne pose pas de problème métaphysique particulier. Cela implique de voir le travail du philosophe comme essentiellement destiné à construire le cadre théorique afférent à une question ou à en discuter la pertinence. Mais en tout cas, ce travail ne consiste pas à légitimer l'existence d'entités hors de tout contexte.

L'antidéflationnisme de Quine

La stratégie déflationniste de Carnap est essentiellement basée sur l'idée qu'il est toujours possible, dans un système de croyances, de démarquer les contributions qui relèvent de l'esprit humain de celles qui relèvent exclusivement du monde lui-même. Quine a récusé cette thèse dans deux articles. Dans *Two Dogmas of Empiricism* (W. V. O. QUINE, 1951b), Quine rejette la distinction entre jugements analytiques et jugements synthétiques sur laquelle repose la dichotomie des questions internes et externes. Et dans « On Carnap's view on ontology » (W. V. O. QUINE, 1951a), Quine récuse directement la possibilité de démarquer ces deux types de questions.

Le problème est que si l'on accepte la proposition synthétique « il y a un personnage de fiction du nom de Sherlock Holmes » [7], cela implique alors qu'il existe un personnage de fiction. Chez Carnap, la distinction entre assertions externes et internes n'a aucun rapport avec le fait d'admettre des croyances. Or Quine a montré que favoriser un système linguistique revient justement à privilégier certaines croyances. Et réciproquement, admettre des croyances, c'est admettre tel système de langage plutôt que tel autre. Les engagements ontologiques ne portent pas sur les prédicats eux-mêmes. Par exemple dans « il y a des poissons », ils ne portent pas sur « poisson ». Ils portent sur des variables de quantification affectées à des prédicats (il y a des choses qui poissonnisent).

> Thus I consider that the essential commitment to entities of any sort comes through the variables of quantification and not through the use of alleged names. The entities to which a discourse commits us are the entities overwhich our variables of quantification have to range in order that the statements affirmed in that discourse be true. (W. V. O. Quine, 1951a, p. 67)

De fait, c'est devenu un lieu commun que d'identifier la renaissance de la métaphysique à la parution, en 1948, de l'article de Quine intitulé « De ce qui est ». Dans cet article, Quine souligne à la fois la simplicité apparente de la question ontologique et son caractère discursif :

> Ce qu'il y a de curieux avec le problème ontologique, c'est sa simplicité. On peut l'énoncer en trois syllabes : « qu'y a-t-il ? ». Et l'on peut, qui plus est, lui apporter une réponse en un mot : « tout » - et chacun acceptera cette réponse comme vraie. Cependant, cela revient à dire simplement qu'il y a ce qu'il y a. Cela laisse la possibilité de désaccords au cas par cas ; et ainsi la question continue de se poser depuis des siècles. (W. V. O. QUINE, 2003b, p. 25)

[7] En recourant par exemple à l'histoire de la littérature, ce que Carnap n'aurait peut-être pas admis, le modèle même des propositions synthétiques relevant chez lui seulement des propositions des sciences dites « dures ».

L'engagement ontologique consiste en l'inclusion d'entités dans le domaine de quantifications des variables. Selon Quine, exister, pour une entité, c'est répondre au critère suivant : une entité est postulée par une théorie si elle doit être incluse parmi les valeurs des variables liées afin que les affirmations stipulées par la-dite théorie soient vraies. Des énoncés manifestant des formes grammaticales ressemblantes pourront revêtir des traductions logiques n'ayant pas le même nombre de variables liées. Par exemple, « Charles vise une cible » pourra par exemple se traduire par $\exists x \mathrm{V} cx$. Le critère d'adhésion ontologique nous demande de reconnaître l'existence de deux choses, à savoir, un individu appelé Charles et une cible posée à une certaine distance et distincte de Charles. Considérons maintenant « Charles vise un objectif ». Dans ce nouveau cas, nous pourrions discuter plus longuement du genre de traduction pertinente. D'un côté, on pourrait prétendre que la théorie semble stipuler l'existence d'une seule entité, à savoir Charles, auquel est appliqué le prédicat « viser un objectif ». La traduction adéquate serait alors du type $\exists x Px$. Mais un platonicien ayant trop lu Baudelaire pourrait préférer considérer au contraire qu'il y a bien là deux entités qui font l'enjeu de l'affirmation, à savoir, Charles et l'Idéal, avec un grand « I », que vise Charles. Comme on le constate ainsi, le critère ontologique de Quine ne nous relève donc pas ce qui existe indépendamment de toute autre chose. Il nous relève ce qu'une théorie *affirme* qu'il y a.

Comme l'écrit Quine,

> Names, in fact, can be dispensed with altogether in favor of unnaming general terms, plus quantification and other logical devices ; the trick of accomplishing this elimination is provided, in its main lines, by Russell's theory of descriptions. Thence forward the variable of quantification becomes the sole channel of reference. For ontological commitment it is the variable that counts. (W. V. O. QUINE, 1951a, p. 67)

Seule la variable est ontologiquement chargée, et l'article défini « le » est dissous dans le langage du calcul logique des propositions.

L'engagement est donc associé à la théorie sémantique de Russell, dont Quine garde l'esprit en vidant toute trace de noms propres logiques. Rappelons brièvement cette théorie. Elle se donne, entre autres objectifs, pour but de réviser la sémantique de Frege. Comme nous l'avons dit, selon Frege, une expression comme « l'actuel roi de France est chauve » est une expression complexe, mais un nom vide. Russell purge la classe des noms propres logiques en la ramenant aux seuls termes afférents à des connaissances par accointance (*acquaintance*). Ces connaissances sont définies comme des expériences directes *via* les *sense-data*, autrement dit, *via* les données immédiates des sens :

> Nous dirons que nous avons l'expérience par accointances d'une chose quand elle est directement devant nous, que nous en avons conscience, sans l'intermédiaire d'aucun processus d'inférence ou de quelque connaissance de vérité que ce soit. (RUSSELL, 1965, pp. 69-70)

La classe des authentiques noms propres ne contient donc que les seuls déictiques comme « ceci », « lui », « je », « ici », etc. Les noms qui ne sont pas directement associés à des expériences de connaissance par accointance sont, à l'inverse, tous traités comme des descriptions définies. Ainsi, l'expression « l'actuel Roi de France », relève d'une description. Et il en va de même d'expression comme « Emmanuel Macron ».

La stratégie quantificationnelle de Russell possède deux intérêts notables. Le premier, c'est que tous les énoncés existentiels négatifs contenant des inexistants deviennent trivialement vrais, puisqu'ils n'affirment rien d'autre que l'inexistence de ce qui n'existe pas. Car, sous l'apparence d'une pensée singulière portant sur un nom vide, « l'actuel roi de France » cache en fait une quantification portant sur une description. En notant Rx pour « x est roi de France » et Cx pour « x est chauve », on peut alors formaliser l'expression sous la forme d'une quantification :

$$\exists x(\mathrm{R}x \wedge \forall y(\mathrm{R}y \rightarrow y = x) \wedge \mathrm{C}x)$$

Étant donné qu'on n'affirme ici rien d'autre que l'inexistence de quelque chose qui correspondrait à la description « actuel roi de France », voilà donc résolu le problème des propositions existentielles négatives.

Le second avantage de la théorie des descriptions définies est qu'elle permet à Russell de rétablir l'universalité du tiers-exclu, dont la théorie de Frege ne parvenait pas à rendre compte, puisque des énoncés contenant des noms vides n'étaient ni vrais, ni faux. Chez Russell, la proposition exprimée par « l'actuel roi de France est chauve ou l'actuel roi de France n'est pas chauve » peut être interprétée selon deux possibilités alternatives dotées chacune d'une valeur de vérité. En effet, on peut l'interpréter soit par la proposition

$$\exists x\ (\mathrm{R}x \wedge \forall y(\mathrm{R}y \rightarrow y = x) \wedge (\mathrm{C}x \vee \neg\mathrm{C}x))$$

qui est fausse en raison de ce qu'elle quantifie sur un inexistant, soit encore par la proposition :

$$\exists x\ (\mathrm{R}x \wedge \forall y(\mathrm{R}y \rightarrow y = x) \wedge \mathrm{C}x) \vee \neg\exists x\ (\mathrm{R}x \wedge \forall y(\mathrm{R}y \rightarrow y = x) \wedge \mathrm{C}x)$$

qui, elle, est trivialement vraie en ce qu'elle ressortit de la disjonction « P ou non-P ».

Le mérite de la théorie des descriptions définies est que l'emploi de termes ne contraint pas à stipuler l'existence de choses qu'ils dénoteraient. Quand nous faisons usage du mot « Pégase » et du mot « rouge » nous ne sommes pas dans la nécessité de croire qu'il y a des choses réelles, dans le monde, auxquelles réfèrent effectivement ces noms. Nous n'avons ni besoin de reconnaître un inexistant cheval ailé ni un universel de rougeur. L'idée que tout terme singulier présuppose une entité désignée par ce terme est une illusion. Comme le souligne Quine, « il y a un gouffre entre *signifier* et *nommer* » (W. V. O. QUINE, 2003b, p. 35). Un terme singulier n'a pas besoin de nommer pour être signifiant. En particulier, les assertions existentielles négatives n'ont pas besoin de dénoter pour gagner leur signification.

> Nous nous engageons à une ontologie contenant des nombres quand nous affirmons qu'il y a des nombres premiers plus grands qu'un million, nous nous engageons à une ontologie contenant des centaures quand nous disons qu'il y a des centaures ; et nous nous engageons à une ontologie contenant Pégase quand nous disons que Pégase est. Mais ce n'est pas le cas quand nous disons que Pégase, ou l'auteur de Waverley, ou la coupole ronde carrée qui surmonte Berkeley College n'est pas. Nous ne sommes plus contraints à être victimes de l'illusion qui consiste à croire que la signification d'un énoncé contenant un terme singulier présuppose une entité nommée par le terme (W. V. O. QUINE, 2003b, p. 34-35)

En conséquence, nos engagements ontologiques n'ont pas à porter sur les noms propres comme « Pégase », puisque ceux-ci peuvent toujours être réduits à des description définies par le moyen de périphrases. L'engagement porte donc ailleurs, précisément, sur l'existence des entités qui constituent le domaine de quantification des variables liées. C'est parce que nous estimons qu'il est vrai que $\exists x Fx$ que nous pouvons affirmer l'existence des Fs. Autrement dit, pour Quine, les questions ontologiques ne sont rien d'autre que des questions de quantification. Il résume ceci en un slogan qui sert de critère d'engagement ontologique : « être, c'est être la valeur d'une variable quantifiée. » Pour savoir s'il y a des Fs (mettons, s'il y a des nombres), il suffit de parcourir le domaine de quantification de ce qu'une science dit qu'il y a et ensuite de déterminer si une valeur satisfait à la quantification « $\exists x\ Fx$ ».

Cette façon de considérer la notion d'existence en renforçant la stratégie sémantique russellienne par une réduction drastique des noms propres aux descriptions définies a un évident effet sur le dénombrement des entités. Avec Quine, les notions d'existence, d'unité et d'objet sont entièrement corrélées. Exister, c'est être un objet, et être un objet, c'est être « un » : pas d'entité sans identité, pas plus que d'identité sans entité.

En un sens, l'ontologie de Quine est relativisée à un système de représentation scientifique qu'elle prolonge. La métaphysique n'a rien à voir avec un discours pré-scientifique qui élaborerait de manière autonome les bases *a priori* d'un discours spéculatif. Déterminer ce qu'il y a, c'est préalablement s'inscrire dans un système capable d'organiser *a posteriori* les données éparpillées de nos expériences. En ce sens, une ontologie n'a de raison d'être que si elle constitue un schème « arrangeant les fragments désordonnés de l'expérience brute » (W. V. O. QUINE, 2003b, p. 44). C'est cette capacité à organiser de manière cohérente ces données qui permet de sélectionner entre deux ontologies rivales : la gagnante devrait être celle qui offre l'explication la plus simple au plus grand nombre d'observations. Étant donné que nos engagements ontologiques sont étroitement dépendants de nos meilleures théories scientifiques, il faut en conclure qu'ils sont par nature évolutifs. Quine prescrit donc une tolérance basée sur l'expérimentation, seule garante de l'évolution positive d'une discipline scientifique. En rejetant l'idée d'une métaphysique absolue, Quine amorce ainsi un pluralisme ontologique.

La théorie russellienne des descriptions définies, qui traduit les propositions dans une langue formelle permettant d'appliquer le calcul des propositions, incitait Carnap à considérer qu'une analyse purement grammaticale permet de démarquer d'une part les propositions bien formées, relevant des énoncés de la science et, d'autre part, les pseudo-propositions non signifiantes. Cela promettait une vaste entreprise de nettoyage visant à récurer, par l'effet détergent de la logique, le plancher philosophique de toutes traces métaphysiques [8]. Chez Quine, l'ontologie matérielle est épistémisée, relativisée. Le réservoir de ce qu'il y a est le domaine de quantification de nos théorie. Un F existe signifie qu'il y a une variable liée qui « F-ise » . L'ontologie trouve sa source dans l'état des connaissances scientifiques que nous en avons, dans nos croyances communes, dans ce que nous acceptons comme vrai. Ce réservoir peut, selon les croyances qu'il occasionne, s'accroître ou régresser. L'histoire des sciences retrace ainsi les aléas du catalogue ontologique, voyant certaines entités ajoutées (les nombres imaginaires, les fermions, les gènes, les groupes sociaux) ou au contraire éliminées (les cercles de même aire qu'un carré donné constructibles à la

[8] Il faut toutefois souligner qu'à partir de 1942, sous l'influence des travaux de Tarski, Carnap admet l'importance des concepts sémantiques de référence et de vérité pour déterminer le statut cognitif des énoncés.

règle et au compas, la planète Vulcain, les humeurs sanguines, les races humaines).

Bien sûr, ce ne sont pas les mêmes variables qui existent selon qu'on a affaire aux mathématiques, à la physique, à la biologie, voire à l'histoire de la littérature. Mais qu'est-ce qui permet de défendre une ontologie plutôt qu'une autre ? Peut-on défendre une ontologie qui donnerait crédit aux personnages de fiction ? Comment départager entre ontologies rivales ? Quine répond que l'engagement ontologique relève d'un argument d'indispensabilité : nous devons nous engager sur l'existence de toutes les entités qui s'avèrent indispensables à nos meilleures théories. Dans l'esprit de Quine, ces théories relèvent exclusivement des sciences de la nature. Bien sûr, l'indispensabilité relève d'une croyance : certains scientifiques crurent que l'éther, le phlogiston ou la planète Vulcain étaient indispensables à l'explication des phénomènes qu'ils prétendaient respectivement expliquer. Cela indique que les entités dites « indispensables » peuvent l'être de manière temporaire, avant de tomber dans la catégorie des fictions scientifiques. La sémantique de Quine vaut essentiellement pour expliquer le fonctionnement des langages scientifiques, c'est-à-dire, de langages idéalisés, mais, en donnant sens aux discours scientifiques, elle prend trop aux discours non formels.

Le mérite de la stratégie de Quine est de rappeler que l'enquête ontologique ne doit pas consister en un procès du monde en comparution devant un tribunal jugeant de ce qu'il y a. Elle découle de nos croyances les plus indispensables. Malheureusement, cette stratégie est aussi assortie d'un rejet global de la vérité en fiction : tout discours comportant des termes fictionnels est inévitablement faux. En particulier, toutes les propositions existentielles affirmatives contenant des noms fictionnels, comme « Sherlock Holmes est un détective », ou « Ulysse est un personnage de l'Iliade » ou encore comme « Ulysse est identique à Ulysse » s'avèrent inévitablement fausses. En outre, comment justifier l'usage scientifique de fictions comme les milieux sans résistance ou les gaz parfaits ?

Avec Quine, la fiction est certes discernée ontologiquement du réel, mais elle l'est aussi de manière sémantique. S'il convient bien de la démarquer du réel, et ainsi de considérer qu'elle n'est pas un double métaphysique de la réalité, le tribut à payer est lourd : les discours fictionnels sont systématiquement considérés comme des affirmations fausses. La dette que doit payer la fiction est cependant trop grande pour en rester à l'idée que l'ontologie cartographie intégralement le discours vrai. La question qui se pose alors est de savoir comment aménager une place au vrai lorsqu'on parle de choses qui, de manière unanime, n'existent pas.

Boris Vian écrit, dans sa préface à *L'écume des jours* que son histoire est entièrement vraie puisqu'il l'a inventée de toute pièce (VIAN, 1963, p. 9). La fiction relève certes de l'irréel, mais cela n'implique pas nécessairement qu'elle soit fausse ou futile. Et même, comme l'invitent à penser Manuel Rebuschi et Marion Renauld (REBUSCHI & RENAULD, 2010), la fiction peut être considérée comme une connaissance indispensable à nos connaissances, autant que le sont les sciences dures. Il s'agit alors d'élargir le critère ontologique de Quine aux œuvres littéraires.

3.1.3 La valeur didactique des représentations mimétiques

Contrairement à Platon qui dissocie imitation et diégèse, Aristote fait de la *diégésis* l'un des types particulier de la *mimésis*. Le récit est l'une des deux formes de l'imitation, l'autre étant la représentation directe des événements joués par des acteurs devant un public. La *poièsis* est ainsi

complètement incluse dans la *mimésis*. Par contre, les récits racontés à la première personne ne rentrent pas dans le domaine de la fiction [9]. Seules les productions qui font appel à l'imitation, par représentation scénique ou par le récit, sont véritablement appelées des œuvres poétiques.

La poétique est partie intégrante de la *poïétique*, c'est-à-dire de la sphère des productions dont la finalité est externe, autrement dit, qui vise une fin en dehors de son propre processus. En cela, elle s'oppose à la *praxis* dont la finalité réside au contraire dans son propre mouvement. Chez Aristote, il ressort que l'articulation entre *mimésis* et *poiésis* est paradoxale, puisque l'œuvre fictionnelle est à la fois une re-production parce qu'elle imite un réel qui vient à manquer, mais aussi parce qu'elle est une production, autrement dit, surgissement d'une réalité n'ayant pu être produite par la nature. L'artiste producteur fabrique un objet qui, dès lors, appartient à la dimension du réel, s'y ajoutant en complétant la réalité de possibilités non réalisées. Cette tension entre imitation et innovation fait tout l'enjeu et la spécificité de la *poiésis*.

L'imitation peut se dire de trois manières : elle consiste à peindre les choses soit telles qu'elles furent ou telles qu'elles sont réellement, soit telles qu'on les croit être ou telles qu'elles semblent être, soit telles qu'elles devraient être. La *mimésis* est donc divisée en trois catégories : épistémique, doxastique et déontique. Lorsque l'artiste échoue à imiter la réalité, cet échec advient par l'impuissance du poète à mimer correctement ce qu'il s'était proposé de peindre. La représentation fautive d'un cheval au galop balançant simultanément ses deux jambes en avant ne provient que de l'ignorance de l'artiste. C'est donc le manque d'éducation scientifique de l'artiste qu'il faut critiquer mais non pas condamner les actes d'imitation eux-mêmes comme le fait Platon.

Loin d'être une illusion menant à la déroute des connaissances, la fiction a donc une véritable valeur épistémique. L'image des choses ne doit pas être rejetée dans les marges de l'illusoire puisqu'elle autorise leur connaissance. C'est ce que montre, d'après Aristote le plaisir que nous prenons à nous représenter des choses qui nous rebuteraient dans d'autres circonstances.

> Nous trouvons plaisir à regarder les images les plus soignées des choses dont la vue nous est pénible dans la réalité, par exemple les formes d'animaux tout à fait ignobles ou de cadavres ; la raison en est qu'apprendre est un plaisir non seulement pour les philosophes, mais également pour les autres hommes [...] ; en effet, si l'on aime à voir des images, c'est qu'en les regardant, on apprend à connaître et on conclût ce qu'est chaque chose comme lorsqu'on dit : celui-là, c'est lui. (ARISTOTE, 1996, p. 82)

Outre ces qualités, les fictions ont un rôle éminemment social puisqu'elles ont le mérite de produire des effets thérapeutiques, autant bénéfiques au niveau individuel que collectif. Car elles encouragent et stimulent la purgation du pathos (*catharsis*) : « La tragédie [...] suscitant pitié et crainte, opère la purgation propre à pareilles émotions (ARISTOTE, 1996, p. 88) ». La représentation tragique de destins funestes convie les spectateurs à l'apaisement social en représentant les désastreuses conséquences de la démesure née d'*hybris*.

Bien sûr, il arrive que les fictions mentionnent des événements, des choses impossibles. Aristote considère qu'elles sont alors fautives. Néanmoins de telles infractions ne sont pas rédhibitoires à condition qu'elles aient pour vocation d'atteindre plus profondément la finalité propre que

[9] Gérard Genette souligne que le champ de la littérature, autant chez Platon que chez Aristote, est donc complètement ravalé aux catégories de la *mimésis*. Y échappent donc, par exemple, les poésies lyriques de Sapho et de Pindare ainsi que le poème philosophique d'Empédocle (GENETTE, 2004).

se donne l'œuvre en tant que telle. Aristote illustre ici son propos en mentionnant l'exemple de la poursuite d'Hector dans l'Iliade. Pour rappel, alors qu'Achille a rattrapé Hector, il ordonne à ses compagnons, d'un signe de tête, de lui abandonner le sort du troyen. Or les troupes grecques ne s'étant pas lancées à la poursuite d'Hector, elles sont manifestement hors d'atteinte des prescriptions d'Achille. Selon Aristote, Homère s'acquitte de cette aberration parce qu'elle apporte encore plus d'expressivité à la représentation de la volonté sans faille d'Achille. Mais le philosophe rejette néanmoins l'utilisation d'éléments irrationnels qui ne seraient pas justifiés par des motifs stylistiques. Les histoires ne doivent pas déroger gratuitement à la logique. Le possible élaboré par l'œuvre de fiction ne doit pas être absurde. Le récit de fiction doit obéir à des règles coercitives de cohérence dictées par la logique de l'action. La *Poétique* esquisse une approche modale de la fiction qui la restreint au probable et qui rejette les fictions impossibles :

> Il est clair d'après ce que nous venons de dire, que ce n'est pas raconter les choses réellement arrivées qui est l'œuvre propre du poète, mais bien de raconter ce qui pourrait arriver. Les événements sont possibles suivant la vraisemblance ou la nécessité (ARISTOTE, 1996, p. 88)

Ainsi, l'imitation chez Aristote a pour enjeu la façon dont le monde aurait pu être, le possible.

Puisque le mécanisme de la fiction repose sur le vraisemblable, la poétique doit être rapprochée de la rhétorique, autrement dit l'art de persuader, puisque les deux disciplines travaillent sur la dimension du probable. La vérité fictionnelle s'articule entre d'une part les raisonnements analytiques, déductions produites à partir de prémisses vraies, et d'autre part la dialectique, dont les prémisses sont des opinions simplement probables. La logique sert alors à maintenir une cohérence dans l'enchaînement des arguments qu'elle articule. La tragédie, par exemple, répond de manière exemplaire à ces enchaînements inflexibles, en déployant les conséquences logiques d'une situation initiale seulement probable.

Contrairement à la théorie mimétique platonicienne, la poétique aristotélicienne ouvre la perspective du vrai en fiction.

3.2 Théories ontologiquement non démarquantes

Nous voudrions ainsi laisser une place à la fiction et, pour le dire comme Clément Rosset, rendre grâce aux doubles de « proximité » que produit l'imaginaire. Comment rendre alors la vérité qui revient à certains énoncés fictionnels ? Nous devrions nous mettre en route vers une conception intensionnaliste. De fait, le leg de Frege, de Quine et de Russell concernant la sémantique des fictions peut alors être résumé en ces termes : nous n'avons jamais que deux options concernant notre compréhension d'une assertion existentielle négative comme « Sherlock Holmes n'existe pas » :

- soit nous la considérons vraie, mais nous devons alors renoncer à faire porter sa vérité directement sur la référence du sujet grammatical (anti-réalisme des *ficta* [10]).

[10] Dans ce qui suit, j'appellerai « objet fictionnel », ou, pour abréger, *fictum*, ce que les réalistes désignent comme les entités associées aux personnages de fiction, quelle que soit la nature qui leur est attribuée.

- soit nous la considérons fausse parce qu'elle traite d'une véritable entité qui est la référence du sujet grammatical (réalisme des *ficta*) ;

Le critère ontologique de Quine repose sur l'idée d'une association entre la vérité et l'ontologie. Ce qui est vrai est réel, et ce qui est réel est vrai. *Exit*, donc, toutes les assertions fictionnelles vraies. En effet, elles ne font pas partie des connaissances vraies. Elles n'ont donc aucun rapport au réel, et en conséquence, ne font pas partie de nos savoirs indispensables.

L'échec de rendre compte de vérités aussi évidentes que celle portée par la phrase « Sherlock Holmes est Sherlock Holmes » a conduit certains philosophes à rejeter la thèse que « Sherlock Homes » soit un nom vide. Une autre tentative d'explication consiste selon eux à admettre, contre Quine, Russell et Frege, que nos assertions fictionnelles sont bien dénotantes. Cette perspective implique de considérer que les entités fictionnelles sont ontologiquement indiscernables des entités réelles, certes non pas d'un point de vue absolu, mais au moins d'un point de vue relatif.

3.2.1 Les théories intentionnelles de la représentation

La représentation comme processus de duplication

Les philosophes ont entretenu de nombreuses croyances, plus ou moins vraisemblables, au sujet des représentations vues comme des sortes de duplication du réel. La représentation et son objet y entretiennent un rapport mimétique parfois réduit à une simple relation d'indiscernabilité chez ceux qui considèrent les représentations mentales comme des sortes de répliques des choses hors de l'esprit.

Avec la théorie des simulacres, Lucrèce explique nos représentations mentales comme des sortes de saisies d'émanations provenant directement des choses qui nous font face :

> De tous les objets, il existe ce que nous appelons les simulacres : sortes de membranes légères détachées de la surface des corps, et qui voltigent en tous sens parmi les airs. (LUCRÈCE, 1978, p. IV, 30-35)

Lucrèce explique d'ailleurs le phénomène de le représentation en l'absence et du rêve par le même processus de réception de simulacres, qui, dans ce cas, ne se fait plus en présence, mais de manière différée, les simulacres ayant été retenus ou propagés dans les airs :

> Et dans la veille comme dans le rêve, ce sont ces mêmes images dont l'apparition vient jeter la terreur dans nos esprits, chaque fois que nous apercevons des figures étranges ou les ombres des mortels ravis à la lumière ; souvent elles nous arrachent du sommeil, tout frissonnants et glacés d'effroi. N'allons donc pas croire que des âmes puissent s'échapper de l'Achéron, ou des spectres voltiger parmi les vivants ; ne croyons pas davantage que rien de nous puisse subsister après la mort [...]. (LUCRÈCE, 1978, p. IV, 35-45)

Si Lucrèce peut ainsi expliquer l'image mentale de disparus, sa théorie tente également d'éclairer le phénomène de l'imagination fictionnelle, bien que, dans ce cas, il n'y ait nul objet d'où émaneraient des simulacres. Mais Lucrèce considère que les simulacres sont de différents types. A côté

des simulacres-copies qui sont générés par un objet original, certains simulacres sont engendrés de manière « spontanée », c'est-à-dire sans modèle préalable d'où ils s'échapperaient :

> Mais ne va pas croire que seuls parcourent l'espace les simulacres qui émanent des corps. Il est d'autres images, engendrées spontanément, qui se constituent d'elles-mêmes dans cette région du ciel qu'on nomme l'atmosphère. Formées de mille manières, elles s'élèvent dans les hauteurs, et ne cessent dans leur course, de se fondre et de se transformer, et de prendre les aspects les plus divers : tels ces nuages que nous voyons parfois se rassembler dans les hauteurs et qui, caressant l'air de leur vol, violent la sérénité du ciel. Souvent alors il nous semble voir voler des faces de géants qui répandent leur ombre, ou encore s'avancer de hautes montagnes, entraînant des roches arrachées à leurs flancs et dont la marche masque le soleil : puis c'est quelque autre figure bizarre qui attire à elle d'autres nuages pour s'en affubler. (LUCRÈCE, 1978, p. IV, 130-1405)

En outre, les simulacres ont les mêmes propriétés de combinaison que les atomes, ce qui explique ainsi les images mentales composites comme les chimères et les centaures :

> [...] de toutes parts errent en foule les simulacres de toute espèce, subtils, et qui dans les airs n'ont pas de peine, en se rencontrant, à se souder les uns aux autres, comme des toiles d'araignées ou des feuilles d'or. [...] C'est ainsi que nous voyons des Centaures, des formes de Scyllas, des museaux de Cerbères, et les fantômes des morts dont la terre recouvre les os. En effet, des simulacres de toute sorte sont emportés dans l'espace, les uns engendrés spontanément dans l'air même, d'autres échappés des différents objets, d'autres enfin, formés de la réunion des diverses images. Ce n'est certes pas d'un modèle vivant que provient l'image du Centaure, puisqu'un tel animal n'a jamais existé : seulement si le hasard rapproche l'image du cheval de celle d'un homme, elles se soudent sans peine aussitôt l'une à l'autre, comme nous l'avons dit plus haut, grâce à leur nature subtile et à leur tissu délié. Toute autre idée semblable a la même origine. (LUCRÈCE, 1978, p. IV, 720-745)

Pour une autre catégorie de philosophes, les représentations mentales sont dues à la saisie directe des essences. Chez Aristote, l'acte de représentation est vu comme un acte spéculaire, une sorte de réflexion en miroir du réel, par le truchement de l'information de l'âme imitant directement la forme de l'objet à travers l'acte de représentation. Et chez Leibniz, la théorie théologique de l'harmonie structurelle préétablie permet de rendre compte de la compatibilité entre la réalité et sa représentation. De telles conceptions permettaient en particulier d'expliquer le fait que des individus numériquement distincts puissent former la même représentation indépendamment des conditions proprement subjectives. Mais bien sûr, l'une et l'autre nécessitaient une souscription coûteuse, respectivement à la théorie de l'hylémorphisme et à l'existence d'un démiurge bienveillant.

Avant les travaux de Bolzano sur l'intentionnalité, les opinions concernant ce qui peut constituer une représentation étaient elles aussi fort diverses : parmi elles comptaient les sensations et parfois même les jugements. Mais Bolzano a réduit la notion de représentation en la caractérisant depuis un point de vue strictement logique : une représentation est la partie d'une proposition (au sens frégéen du terme) qui n'est pas elle-même une proposition. L'analyse des propositions

passe ainsi par la décomposition atomique de ses constituants. Le point de vue, proche de l'analyse propositionnelle de Frege, est donc appuyé sur la compositionalité des représentations. Là où Frege considère que la signification de la proposition dépend de la signification de ses composantes, Bolzano affirme qu'une pensée représente quelque chose si les composantes de cette pensée représentent elles-mêmes des choses. La vérité naît de l'association pertinente entre un objet de représentation et une propriété qui lui est attribuée dans une proposition alors même qu'il exemplifie cette propriété. Certaines représentations servent à désigner des objets, contrairement à d'autres, comme celles rattachées aux conjonctions « et », « ou ». Ces dernières, appelées syncatégorématiques, remplissent une autre fonction mais sont néanmoins véritablement considérées par Bolzano comme des représentations [11].

Il reste alors la question de savoir comment une représentation atomique est associée à son extension. Sans répondre à la question, qu'il semble juger de toute façon absurde, Bolzano prévient toutefois qu'il ne faut surtout pas faire appel ici à la relation de ressemblance :

> Quiconque assimile cet être représenté à une sorte de figuration ; quiconque veut parler comme si le rapport entre un objet et sa représentation était plus ou moins semblable à celui entre un objet sensible et son image ; quiconque suppose seulement qu'il y a une certaine similitude entre la représentation et l'objet, un accord entre les propriétés de l'un et l'autre ; et surtout, quiconque (comme nos philosophes de l'identité) parle de l'égalité complète, même de l'identité des deux — celui-là, je vous dis, s'avance déjà sur un chemin des plus dangereux, un chemin sur lequel, s'il continue de le suivre, il gardera pour toujours son dos tourné à la pensée claire et distincte. Car, quelle sorte de similitude existe-t-il ou pourrait-il exister entre la représentation de quelque chose d'un côté et ses objets de l'autre côté — c'est à dire, tout objet qu'il y a ? Qu'on me montre deux choses qui sont *moins semblables* que les représentations *propriété, objet spatial, machine, concept, proposition*, etc., et les objets qui sont représentés par ces représentations (et ici je ne spécifie pas si on doit penser ces représentations au sens objectif ou au sens subjectif — c'est-à-dire comme des phénomènes dans l'esprit d'un être pensant). (BOLZANO, 1843, p. 74-75)

En rejetant la ressemblance dans l'explication du lien entre l'objet et sa représentation, Bolzano rompt ainsi avec la tradition philosophique qui conçoit les actes mentaux comme des duplications directes du réel. Il rejette les théories concevant la représentation comme une sorte de duplication mimétique, au profit d'une théorie de l'expression : la représentation n'imite pas, elle exprime la réalité. Les genres d'expressions s'avèrent variés :

> Le modèle exprime la machine, le dessin perspectif exprime le volume sur un plan, le discours exprime les pensées et les vérités, les chiffres expriment les nombres, l'équation algébrique exprime le cercle ou toute autre figure ; et ce qui est commun à ces expressions est que, à partir du seul examen des rapports de l'exprimant nous pouvons parvenir à la connaissance des propriétés correspondantes de la chose à exprimer. (BOLZANO, 1843, p. 73)

Selon Bolzano, si la similitude échoue à rendre compte de l'adéquation entre un objet et sa représentation, c'est notamment parce que deux choses quelconques sont toujours ressemblantes

[11] Précisons que ce point de vue singulier n'a pas manqué de susciter des objections de la part des successeurs de Bolzano, notamment Brentano et Twardowski.

sous quelque aspect. En outre, il y a un second problème. D'un côté, nos pensées se rapportent parfois à des choses matérielles : je peux penser aux charançons qui dévorent l'hortensia de mon balcon. En associant des représentations, je peux alors prononcer des propositions vraies comme « les charançons sont friands de feuilles d'hortensia ». Mais d'un autre côté, il arrive aussi parfois que nous pensions aussi à des choses qui ne se trouvent pas dans le monde de manière concrète. Je peux chercher un charançon dans mon hortensia, alors qu'en fait, c'est une sauterelle qui cause les dommages. Et si mon esprit était davantage perturbé, je pourrais éventuellement croire qu'il s'agit aussi du basilic [12] végétarien qui se cache sous les tuiles du grenier. Dans ces deux cas, nous aurions bien sûr encore affaire à une représentation au sens que Bolzano confère à cette notion, mais une représentation qui est sans objet réel. Dans la *Wissenschaftslehre* Bolzano qualifie ce type de représentations d'imagination sans objet (*gegenstandlose Vorstellungen*). Bolzano y inclut autant les objets impossibles (le frère de ce fils unique) que les objets fictifs probables (le charançon recherché qui est en réalité une sauterelle) et improbables (le basilic). Il est évident ici que la relation de ressemblance échoue à rendre compte de ces représentations puisqu'il n'y a rien dans le monde qui puisse offrir comparaison avec elles.

Lorsque nous nous représentons des charançons, nous pouvons loger l'objectivité de notre représentation dans leur réalité physique. Évidemment, lorsque nous avons des représentations sans objet, une telle stratégie ne semble plus possible. Les représentations sans objet sont par définition indépendantes de l'existence d'entités physiques qui les susciteraient. Mais selon Bolzano, ces représentations ne devraient cependant pas être décrites comme de purs contenus subjectifs. En effet, Bolzano constate qu'il serait contradictoire de croire que les représentations soient à la fois de nature purement psychologique et que deux esprits puissent avoir la même représentation. Penser au même objet, remarque Bolzano, nécessite de saisir plusieurs fois la même représentation objective, quoique l'acte de saisie demeure, quant à lui, subjectif.

On a donc affaire à deux ordres de représentations. L'un est certes subjectif, mais l'autre est idéal, objectif. Ce deuxième type de représentations est tel que les représentations sans objet existeraient même en dehors de tout acte mental. Ce second niveau de représentations vient ainsi doubler le monde des choses matérielles par un monde idéal de représentations ancrées sur la signification. Les représentations sans objet seraient donc objectives au sens où elles existeraient indépendamment de toute instanciation subjective.

Après Bolzano, Franz Brentano accentue encore le lien entre représentation et objet en soulignant qu'un phénomène psychique est toujours « à propos » d'un objet, et que tout acte de conscience est toujours conscience « de quelque chose » (BRENTANO, 1944, p. 109), autrement dit, toujours dirigé vers un objet. Cet objet peut parfois s'avérer être une chose concrète et réelle, comme un caillou, un chat, une table, une personne. Mais il peut aussi viser quelque chose n'ayant aucune forme d'existence spatio-temporelle, comme Sherlock Holmes, le nombre $\sqrt{2}$, Ahura-Mazda, la planète Vulcain ou le carré-rond. De telles choses sont sujettes à des visées intentionnelles indépendamment de leur signification.

> Every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction toward an object (which is not to be

[12] La créature mythique, pas la plante !

> understood here as meaning a thing), or immanent objectivity. Every mental phenomenon includes something as object within itself, although they do not all do so in the same way. In presentation something is presented, in judgement something is affirmed or denied, in love loved, in hate hated, in desire desired and so on. This intentional inexistence is characteristic exclusively of mental phenomena. No physical phenomenon exhibits anything like it. We can, therefore, define mental phenomena by saying that they are those phenomena which contain an object intentionally within themselves. (BRENTANO, 1944, p. 68)

Une telle théorie semble, comme on le voit, profilée pour expliquer nos discours fictionnels. Mais quelle charge ontologique demande-t-elle de charier ?

Avant 1905, Brentano soutenait l'existence de deux types d'entités : d'une part les *entia realia* et d'autre part les *entia irrealia*. Dans la première catégorie figuraient les cailloux, les bateaux, les chats, les cercles mais aussi les basilics en tant qu'individus particuliers ou en tant que groupes d'individus. Dans la deuxième catégorie figuraient les possibilités, les concepts, les propriétés, les états de choses et les propositions. Comme on peut le constater, les exemples donnés dans la liste des *entia realia* suggèrent qu'il ne s'agit pas seulement d'entités matérielles. Mais il s'agit toujours de « particuliers » au sens où ces choses sont toujours pensées comme individuées, même si certaines n'existent pas. D'après Brentano, quand on pense à un basilic, on se le représente comme s'il était *un* animal particulier, et non comme l'idée générale de basilic.

Après 1905, Brentano revient sur la distinction entre deux types d'entités et soutient finalement que les *entia realia* sont les seules entités auxquelles nous sommes capables de penser, et que les pensées dirigées vers des *entia irrealia* sont, en dernier recours, elles-mêmes analysables en termes d'individu. Ce point de vue consiste à rejeter la distinction entre individu et concept. D'après Brentano, une proposition de type sujet-prédicat peut en effet toujours être retraduite sous la forme d'une combinaison ne mêlant que des noms d'individus. Par exemple, l'expression « *ein Rose ist rot* » peut en effet s'interpréter sans prédicat, par exemple comme « *Ein Rose ist Rotes* ». Le principal mérite de cette réduction extensionnelle des prédicats est alors d'éviter des interprétations philosophiques trompeuses et des hypostases illégitimes. Par rapport à la conception de Bolzano, cela représente un manifeste bénéfice ontologique.

Le refus le plus saillant, chez Brentano, concerne l'hypothèse de parvenir à penser le non-individué. Contrairement à ce que pensait Bolzano, les termes syncatégorématiques comme « ou », « et » ne relèvent donc pas du tout de représentations. Comme l'écrit Roderick Chisholm, pour Brentano, nous ne pouvons penser à une chose que dans la mesure où nous pensons à un particulier qui l'exemplifie.

> we shouldn't say that unicorns have horns; we should say that to think of a unicorn is to think of something that has a horn. Brentano's more general point may be put this way: We can think only of *entia realia*; and to think of an *ens reale* is to think of something which, if it existed, would be an individual. (CHISHOLM, 1986, p. 10)

Lorsque nous pensons à quelque chose qui n'existe pas, nous y pensons toujours en admettant que, si elle existait, elle serait singulière. Pouvons-nous nous représenter un charançon sans penser à un charançon en particulier ? Brentano répond que non, la représentation est toujours au sujet d'un particulier. Nous ne pouvons donc penser qu'à des choses qui sont individuées. Pouvons-nous nous

représenter des bosons, comme des particules non individuées ? À quoi penserions-nous au juste ? Sans doute pas à des particules au sens classique du terme. Mais à quoi d'autre alors ? Comme on le voit, la représentation n'est pas généralisée à la concevabilité mais reste ancrée dans la capacité à imaginer, au sens visuel du terme, quelque chose.

En outre, contrairement à Bolzano qui considère que les objets de pensée sont objectifs, Brentano souligne que la définition du basilic en tant qu'objet demeure cependant toute psychologique : une chose est un objet intentionnel si elle peut être « présentée » par et dans un acte de pensée consciente. La signification d'une phrase comme « le basilic est un objet de pensée » nécessite donc sa ré-interprétation par une paraphrase comme « telle personne pense à un basilic », qui ne concerne plus directement un objet inexistant qui serait visé en tant que tel, mais à l'acte d'une personne pensant à une chose qui n'existe pas.

Lorsqu'elles visent des choses qui n'existent pas, les *entia realia* demeurent donc des entités purement psychiques, ce qui pourrait les bannir de l'objectivité. Lorsque Brentano tente de préciser ce qu'il entend par (in)existence, il emprunte la notion aristotélicienne de forme. L'objet purement intentionnel est immanent à la pensée par sa forme. Mais ainsi constitué comme contenu d'une représentation mentale, l'objet de pensée ne dénotant pas un objet matériel semble devoir être aussi purement subjectif. Dès lors, si l'on considère que les personnages de fiction sont des objets purement intentionnels, se pose alors la question de déterminer comment deux personnes pourraient avoir une pensée à propos du même personnage et même, comment une seule personne pourrait avoir deux pensées successives à propos du même personnage. C'est à partir de ces deux objections qu'ont été déployées les théories de l'intentionnalité chez les élèves de Brentano.

Dans *On the content and object of representations*, Twardowski a introduit la triple distinction entre l'acte mental de représentation, le contenu psychologique des représentations (*Ausdruck*) et l'objet des représentations (*Bedeutung*). L'intérêt de cette distinction réside dans le fait de révéler de manière systématique les confusions entre les représentations et leurs objets. Elle permet également de mieux cerner la notion d'existence lorsqu'elle est appliquée aux phénomènes psychiques. Ainsi, dire qu'un contenu de représentation existe, ce n'est pas dire qu'il y a réellement un objet que ce contenu représenterait :

> (...) If something 'exists' as represented in the sense of an object of representation, then its existence is no existence in the proper sense of the term. By adding the formulation 'as an object of representation' we modify the meaning of the word 'existence'; something wich exists as the object of a representation does not in fact exist at all, but is merely represented. The phenomenal, intentional existence of an object is opposed to the real existence of that object, an existence wich forms the content of an acknowledge judgement. (TWARDOWSKI, 1979, p. 11)

Comme Brentano, Twardowski considère qu'un jugement est une affirmation ou une négation à propos d'un objet représenté. Juger qu'il n'existe pas de basilics consiste donc à nier l'objet basilic : c'est affirmer ~~basilic~~. Twardowski souligne que ce qui est nié, ce n'est pas le contenu de la représentation, mais le basilic. Cela revient à dire qu'il y a donc un objet, même lorsque la représentation traite de quelque chose qui n'existe pas. De ce point de vue, l'expression « représentation sans objet » semble contradictoire :

> L'expression « représentation sans objet » est de telle sorte qu'elle contient une contradiction interne. Car il n'y a pas de représentation qui ne représenterait pas quelque chose en tant qu'objet ; il ne peut pas y avoir de pareille représentation. Il y a, au contraire, de très nombreuses représentations dont l'objet n'existe pas. (TWARDOWSKI, 1993, p. 114)

Twardowski admet ainsi une équivalence entre le fait d'être un terme catégorématique, et le fait d'être un objet de représentation. Si tous les objets de représentation relèvent de termes catégorématiques, Twardowski admet aussi la converse, à savoir que tous les termes catégorématiques relèvent tous d'objets de représentation, même ceux, comme « Sherlock Holmes » ou « le carré rond » qui ne désignent rien dans la réalité.

Penser à « Sherlock Holmes », c'est d'abord se représenter un objet, et se poser seulement après coup la question de l'existence de cet objet. La question de l'existence relève exclusivement du jugement, et non de la représentation qui peut stipuler de tels objets fictionnels nonobstant leur évidente absence dans le monde. Autrement dit, pour qu'un objet soit jugé inexistant, il faut d'abord qu'il apparaisse comme un contenu de pensée. Ainsi, Twardowski introduit le terme d'existence phénoménale de l'objet, mais pour aussitôt nier que la signification d'une telle existence relève du sens habituel donné au terme :

> À l'existence effective d'un objet, telle qu'elle forme le contenu d'un jugement de reconnaissance, s'oppose l'existence phénoménale, intentionnelle de cet objet (TWARDOWSKI, 1993, p. 110)

Comme l'écrit Jocelyn Benoist dans *Représentations sans Objet* :

> (l)'objet en tant que visé et purement en tant que visé n'a pas l'existence, pas plus que l'inexistence : il est indifférent ontologiquement à de telles propriétés contrairement au « contenu », qui, quant à lui, renvoie à un morceau de réalité : la réalité mentale. (BENOIST, 2001, p. 92)

L'une des stratégies visant à éclaircir les caractéristiques de l'intentionnalité consiste ainsi à désarticuler l'association existence-concrétude, en désolidarisant les relations intuitivement instituées entre les références de ces termes. Le puzzle de l'existence peut être résumé par la question de la légitimité de cette manière de procéder, ou plus précisément, par les deux questions suivantes :

- L'existence est-elle ou non une propriété des individus ?
- Si c'est le cas, y a-t-il des individus auxquels manqueraient une telle propriété ? Autrement dit, l'existence est-elle séparable de l'individuation ?

La question, à rebours de Kant et Frege, de savoir si l'existence est une propriété consiste à se demander si le livre posé présentement devant moi, l'hortensia occupant ma terrasse et Emmanuel Macron instancient la propriété d'exister, en plus de celles d'être respectivement constitué de 220 pages, d'être en floraison et d'être l'actuel président de la République française.

En quel sens disons-nous alors qu'« il y a des personnages de fiction » ? Dans cette formulation, que signifie au juste « y avoir » ? En quoi démarquer cette formulation de celle d'existence ?

Pour certains philosophes, dont Aristote, l'existence coïncide complètement avec l'essence. Rien ne sépare l'existence d'un individu de ce qui le définit essentiellement. La tradition médiévale a globalement suivi l'identification aristotélicienne de l'existence à l'essence avant que Thomas d'Aquin n'en conteste le dogme. L'objection consiste à dire qu'il est possible de penser à une chimère nonobstant qu'une telle chose n'existe pas. Thomas d'Aquin en conclut que l'essence de la chimère est indépendante de son (in)existence. Et puisque l'existence ne coïncide pas avec l'essence, on doit alors en conclure qu'il s'agit d'une propriété qui s'ajoute à l'essence. L'intérêt de cette conception réside principalement dans l'explication qu'elle confère à notre possibilité de penser à des choses qui n'existent pas, comme les êtres impossibles, les créatures mythiques, les hypothèses fausses ou les personnages de fiction.

De fait, nous attribuons des caractéristiques à des choses qui existent. Nous disons par exemple, « un certain S est F ». Cette proposition est vraie si s_1 est F, ou si s_2 est F, etc. où s_1, s_2, etc. sont des énumérations des Ss qui existent. Mais nous pouvons aussi attribuer des caractéristiques à des choses qui ont existé, mais qui n'existent plus et nous pouvons aussi parler de choses qui n'existent pas encore. Les nominalistes du XIVe siècle expliquent la manière dont un nom peut ainsi étendre la portée temporelle de sa signification par le « principe d'ampliation ». Ce principe consiste à désolidariser l'existence de l'essence. Ainsi, on dira par exemple que la proposition « un certain S sera F » est vraie seulement si s_1 *sera* F, ou s_2 *sera* F, etc. où $s_1, s_2, etc.$ sont des énumérations de tous les Ss qui existent *ou qui existeront*. En outre, Marsile d'Inghen a souligné que le principe d'ampliation pouvait aussi s'appliquer de manière modale et intentionnelle. La proposition exprimée par « possiblement un certain S est F » est vraie seulement si possiblement s_1 est F ou possiblement s_2 est F, etc. où $s_1, s_2, etc.$ sont des énumérations de tous les Ss qui existent ou pourraient exister.

Avant 1905, Russell distingue encore deux modes d'existence par l'utilisation des verbes « exister » et « être ». De ce point de vue, l'existence est une propriété des choses, et certaines choses, qui sont, ne disposent pas de cette propriété [13]. Ainsi, « exister » et « être » sont vus comme deux prédicats de premier ordre référant à des propriétés différentes. À côté de l'existence physique, spatio-temporelle, concrète et actuelle, il y aurait donc encore une autre propriété de premier ordre, exprimée par le verbe « être » ou la locution « il y a », propriété qui appartiendrait aux choses qui seraient à la fois existantes et actuelles mais pas concrètes. La distinction entre le livre qui est actuellement sur cette table et la montagne d'or serait ainsi que le premier instancierait les deux propriétés existentielles (exister et être) là où le deuxième n'en instancierait qu'une seule (être) :

> Numbers, the Homeric gods, relations, chimeras, and four-dimensional spaces all have being, for if they were not entities of a kind, we could make no propositions about them. Thus being is a general attribute of everything, and to mention anything is to show that it is. Existence, on

[13] Avant de proposer sa théorie des descriptions définies, Russell adhéra en effet à une conception réaliste des objets fictionnels qu'il rejeta ensuite violemment en l'attribuant au passage à la théorie de l'objet d'Alexius Meinong. Comme le souligne Francis Jacques, les critiques que Russell adresse à partir de 1905 à la théorie de Meinong relèvent plus vraisemblablement de sa propre interprétation de la théorie de l'objet : « En fait, la doctrine de Meinong n'est pas tout à fait celle que Russell lui prête et qu'il s'est plu à réfuter en 1905. Par une ironie bien connue des historiens de la philosophie, la thèse attribuée à Meinong ressemble étrangement à la position prise par Russell lui-même deux ans plus tôt, qui consistait - au mépris du fameux sens de la réalité qu'il allait bientôt revendiquer contre Meinong - à accorder l'être à tout objet de pensée, « aux nombres, aux dieux homériques, aux relations, aux chimères, aux espaces à quatre dimensions » (JACQUES, 1973, pp. 266-267).

> the contrary, is the prerogative of some only amongst beings. (RUSSELL, 2010, p. 67)

Cette distinction entre exister et être est contestée par Brentano comme quelque chose d'assez absurde :

> I do not wish to bring this discussion of mental reference to a close without having given a word of consideration to the view that there is a distinction between "being" and "existing". According to this view both are to be taken in a very peculiar sense. Namely, a person might be led to say that if someone is mentally referring to an object, the object really always has being just as much as he does, even if it does not always exist as he does. (...) I confess that I am unable to make any sense of this distinction between being and existence. (BRENTANO, 1944, p. 213-214)

Plus généralement, Brentano récuse la distinction entre être et existence au niveau des universaux :

> As far as universals are concerned, the assumption that they have being is just as absurd as the assumption that they exist, for it leads to contradictions. And the law of contradiction not only denies that the same thing can both exist and not exist at the same time but also that the same thing can be and not be at the same time. (...) it is impossible to make use of that distinction – of which I can make no rational sense – between being and existence (...) (BRENTANO, 1944, p. 213-214)

Résumé des stratégies

Les différentes théories sémantiques des fictions peuvent ainsi être classées selon leur manière de désolidariser ou non le quadruplet existence, actualité, concrétude et dépendance. La manière dont on conçoit la capacité objectuelle des contenus intentionnels est en effet tributaire de la façon dont on considère les significations des trois couples d'antonymes existant/inexistant, actuel/inactuel, concret/abstrait, dépendant/indépendant, quitte à désolidariser leurs putatives articulations réciproques. On obtient ainsi $2^4 = 16$ possibilités.

Commençons par constater que certains cas échouent à correspondre à une possibilité théorique. Par exemple, on aurait du mal à imaginer ce que pourrait signifier, pour quelque chose, d'exister sans être concret ou actuel. De même, associer inexistence et concrétude semble absurde et à ma connaissance, aucune théorie métaphysique ne paraît admettre ce genre d'hypothèse. En réduisant ainsi la liste aux cas théoriquement pertinents, on obtient alors les six cas indiqués dans le tableau ci-dessous (figure 3.1).

Le premier cas relève des théories anti-réalistes des fictions parce que toute théorie affirmant que les noms fictionnels dénotent un objet qui ne serait pas seulement mental devrait admettre ou l'existence, ou la concrétude, ou l'actualité de cet objet. Le second cas peut être assimilé à la théorie de l'objet de Meinong qui admet des objets inexistants, immatériels mais actuels, comme l'affirment par exemple les théories respectives de Terence Parsons et d'Edward Zalta. Le troisième cas, admettant des objets existants et concrets mais non actuels, relèverait d'une forme de

Cas	Existant	Concret	Actuel	Dépendant	Théorie(s)
Cas 1	Non	Non	Non	Non	anti-réalisme
Cas 2	Non	Non	Oui	Non	Théories meinongiennes
Cas 3	Oui	Oui	Non	Non	Possibilisme
Cas 4	Oui	Non	Oui	Non	Théories platoniciennes
Cas 5	Oui	Non	Oui	Oui	Théories des objets abstraits
Cas 6	Oui	Oui	Oui	Non	Réalisme modal

FIGURE 3.1 : Stratégies référentialistes

possibilisme qu'on pourrait trouver dans la théorie nonéiste de Graham Priest ou dans le modalisme meinongien de Francesco Berto. Le quatrième cas correspond ou à une théorie platonicienne logeant les *ficta* dans un arrière monde, et le cinquième cas à une forme de théorie ontologique stratifiée en niveaux de dépendance, comme celle de Roman Ingarden reprise par Amie Thomasson. Enfin, le sixième cas correspond au réalisme modal de David Lewis, puisque dans cette théorie, Sherlock Holmes existe, en tant qu'individu concret d'un monde possible, et qu'il est actuel du point de vue de ce monde possible. Dans ce qui suit, nous allons nous attarder sur la question de savoir quel rapport à la duplication entretiennent les différentes théories qui admettent des objets inexistants, abstraits ou inactuels.

3.2.2 La théorie de l'objet

Représentations et assomptions

Dans *Ueber Annahmen* (MEINONG, 1910), Meinong reprend une idée esquissée par Frege dans « Fonction et Concept » (FREGE, 1971b, p. 94), consistant à introduire dans la typologie des actes psychiques un intermédiaire entre la représentation et le jugement. Contrairement à la représentation, le jugement n'est jamais neutre à l'égard de son objet, mais toujours ancré dans un rapport doxastique polarisé selon l'acceptation ou le rejet. Car « juger que A », c'est soit « croire que A », soit « croire que non-A ». La question posée par Meinong est de savoir s'il est possible de dissocier la croyance de sa polarisation. Cela l'amène à concevoir des pensées qui seraient certes polarisées, mais sans faire l'objet d'une croyance quelconque. Mais que seraient de telles pensées ? Ni de simples représentations, puisque polarisées, mais pas non plus des jugements, puisque neutres à l'égard de la croyance. De telles pensées occuperaient donc une place intermédiaire entre la représentation et le jugement. Meinong les appelle « assomptions » (*Annahme*)

Les assomptions jouent un rôle d'importance dans la connaissance. Elles « ne font pas qu'éclairer des vieux problèmes à propos de la psychologie et de la théorie de la connaissance, mais révèlent des chemins vers des domaines de faits restant jusque là virtuellement inconnus » (MEINONG, 1910, p. 94). Pour Meinong, l'assomption est bien un état mental authentique, comme le sont aussi les représentations, les jugements, les sentiments et les désirs. Les assomptions ressor-

tissent à l'état mental de quelqu'un qui aurait une pensée à propos d'une chose sans admettre la réalité de cette chose. On peut rapprocher les assomptions des pensées contrefactuelles. Lorsque j'observe les feuilles méticuleusement perforées de mon hortensia, je peux admettre que des charançons s'acharnent à le dévorer. J'émets ainsi un *jugement*, parce que je présuppose d'emblée l'existence de ces insectes sur ma plante. Mais si je réfléchissais plus largement aux diverses causes possibles expliquant que des feuilles soient tailladées, je pourrais évoquer les charançons sans croire particulièrement qu'ils existent sur ma plante en particulier. Dans ce cas, ma pensée relèverait non d'un jugement en ce que ce dernier serait toujours déterminé sur l'existence de la chose visée : il s'agirait d'une assomption.

À titre d'illustration, Meinong fournit l'exemple suivant. Admettons qu'à l'occasion d'un article, un journaliste relate sa stupéfaction à l'égard de l'issue sereine d'un débat parlementaire dont la teneur aurait pourtant dû soulever une vive polémique. Il y a bien ici à la fois un objet de représentation (la polémique) et un jugement (elle aurait dû se produire). Cependant, puisque la polémique attendue n'a pas eu lieu, cet objet manque à l'existence, il demeure absent du monde. Or cette polémique aurait bien pu être décrite sans que l'événement soit particulièrement attendu. C'est dans ce cas de figure que nous aurions affaire alors à une assomption.

On voit bien toute l'importance de la notion d'assomption pour la fiction et plus largement pour l'imaginaire. En fait, Meinong confère plusieurs rôles épistémiques aux assomptions. Premièrement, elles permettent de viser des objets complexes avant même de prononcer un jugement concernant leur réalité. Deuxièmement, elles permettent de penser des objets complexes dotés du statut d'entités idéales. Troisièmement elles rendent possibles des expériences de pensée. Enfin, quatrièmement, elles relèvent du type d'acte impliqué dans la pensée de certains modèles scientifiques. Elles occupent donc une place tierce, dans la phénoménologie de Brentano, entre les rôles joués par les représentations et les jugements. À l'instar des jugements, elles présupposent des objets, elles sont ancrées sur ces représentations et elles peuvent caractériser l'objet de ces représentations par l'attribution de propriétés. Mais elles diffèrent des jugements en ce qu'elles neutralisent la portée existentielle des représentations.

Meinong va ensuite focaliser son attention sur le rôle des objets qui se rapportent à la théorie des assomptions. Dans son ouvrage *Über Gegenstände höherer Ordnung und deren Verhältnis zur inneren Wahrnehmung* (MEINONG, 1899), il reprend la tripartition émise par Twardowski entre acte mental, contenu psychologique (*psychologischer Inhalt*) et objet de représentation (*Gegenstand einer Vorstellung*). Mais Meinong souligne que connaître quelque chose, c'est nécessairement connaître un objet, indépendamment du fait que celui-ci existe ou non. Un contenu de pensée doit donc toujours se rapporter à un objet, que celui-ci existe ou non. Car limiter la connaissance à ce qui existe, c'est accorder trop de faveur à l'existence. Or « ce qui est destiné à être un objet de connaissance n'a nullement pour nécessité d'exister. » (MEINONG, 2000, p. 71). La totalité des choses qui ont existé, qui existent et qui existeront est donc considérablement moins élevée que la totalité des objets de connaissance.

Aussi, Meinong admet-il des objets inexistants et même des objets impossibles, dont l'admission motive l'émergence d'une nouvelle science, appelée « théorie de l'objet », science concurrentielle de la phénoménologie et de la psychologie, mais dont la vocation est plus générale puisqu'elle prétend se consacrer à l'étude la plus complète des objets, sans discriminer sur leur (in)existence.

Ce qui motive principalement la théorie de l'objet relève de nos représentations à l'égard de ce qui n'a aucune sorte d'existence. Le fait est que nous produisons des attitudes intentionnelles portant sur des choses absentes du monde. Dire que Ponce de Léon cherchait la fontaine de jouvence, c'est dire que Ponce de Léon cherchait quelque chose, bien que cette chose soit demeurée absente du monde. Dire que les Grecs vénéraient Athéna, c'est dire qu'ils vénéraient quelque chose. Or Athéna n'existe pas. Donc les Grecs vénéraient quelque chose qui n'existe pas. Quand Tolkien attribue trois anneaux aux rois elfes, il semble qu'il faille admettre qu'il y ait trois anneaux appartenant aux rois elfes. La question est précisément celle de savoir s'il est correct de lier la quantification logique à l'ontologie. Selon son auteur, et contrairement aux reproches de Russell, la théorie de l'objet n'a pas de prétention métaphysique, mais seulement épistémique. Il se s'agit pas de dire qu'il y a de l'être au delà de l'existant. Meinong ne dit pas que lorsque nous prédiquons des propriétés à un nom vide, nous disons qu'il *y a* une chose que désigne ce nom. Il dit plutôt que lorsque nous pensons à une chose qui n'existe pas, celle-ci peut quand même avoir, comme objet de pensée, les attributions que nous lui prêtons, en dépit de son inexistence.

Les traitements respectifs qu'offrent Russell et Meinong en réponse à la question des énoncés existentiels négatifs permettent de mesurer tout leur écart théorique initial. Rappelons que la théorie de Russell consiste à expliquer la vérité de « l'actuel roi de France n'existe pas » en parcourant l'ensemble de toutes les choses, et en n'en trouvant aucune ne pouvant être décrites comme « actuel roi de France ». La théorie de Meinong consiste, inversement, à penser directement à cet objet en lui attribuant d'emblée l'inexistence. Une chose est de penser à l'actuel roi de France et à nier ensuite son existence, une autre est de penser aux choses réelles et à se demander si l'une d'entre elles correspond ou non à un actuel roi de France. Meinong renonce donc à admettre que l'équivalence logique entre « l'actuel roi de France n'existe pas » et « aucune chose réelle n'est actuellement roi de France » soit également une équivalence sémantique.

> Si je dois pouvoir, à propos d'un objet, juger qu'il n'est pas, il semble que je sois dans la nécessité d'appréhender une première fois l'objet pour pouvoir en prédiquer le non-être, plus exactement, lui imputer (*zuurteilen*) ou l'en démettre (*arbuteilein* (MEINONG, 2000, p. 74)

Le point de vue de Meinong est donc que toute assertion d'existence négative relève d'une proposition intensionnelle et qu'elle ne peut d'aucune manière être réduite, comme le prétend Russell, à une proposition extensionnelle.

Meinong élargit l'extension de la catégorie frégéenne d'objet qui était définie au seul niveau propositionnel comme sujet saturé d'un prédicat dans une expression. Selon Meinong, ce sont au contraire les caractéristiques intentionnelles qui doivent expliquer les caractéristiques sémantiques du langage. L'adéquation des mots provient de l'*à propos* des pensées. En dépit de son inexistence, la montagne d'or pensée par Pierre doit bel et bien être entendue comme un objet, plus précisément comme un objet qui est intrinsèquement intentionnel. L'un des aspects de cette perspective est qu'elle offre la possibilité de rendre compte d'une référence, qui, dépassant le cadre nominal, est élevée au niveau de la proposition. La théorie de l'objet souscrit donc à une forme de réalisme propositionnel. Tout ce qui peut être éprouvé, d'une quelconque manière, comme étant la visée d'un acte mental, Meinong l'appelle objet (*Gegenstand*). Par définition il s'agit d'une chose individuée.

Meinong souligne que différents genres d'actes de pensée correspondent à différents genres

d'objets. Il distingue ainsi d'une part les objets (*Objekt*), qui correspondent aux choses données comme telles à la représentation, et d'autre part les objectifs (*Objektiv*) qui correspondent aux jugements et aux assomptions, aux phrases énoncées à l'aide de la conjonction de subordination « que ». Alors que Frege reléguait les objectifs au domaine du sens, Meinong les déplace au contraire dans le domaine de la référence. Précisons en outre qu'une représentation mentale n'apparaît jamais de manière isolée, elle est toujours mêlée d'attitudes propositionnelles. D'un côté, tout acte de représentation mentale présuppose un ensemble d'attitudes intentionnelles qui l'accompagne. Nous ne nous représentons pas seulement « le rouge », mais qu'une chose quelconque est rouge, ou que le rouge est quelque chose ((MEINONG, 1910, §§38, 45, 63) ; (MEINONG, 1915, §33 ; 1917, §§1, 8)). D'un autre côté, et en retour, tout acte de pensée dépend aussi d'un contenu de représentation. Ainsi, un objectif dépend toujours d'un objet.

La stratégie de Meinong consiste ainsi à réduire la différence entre l'imaginaire et le réel. Il ne s'agit pas d'une théorie mimétique, au sens où les objets mentaux ne sont pas des copies d'objets matériels, mais des objets d'une nature différente.

Pour nombre de néo-meinongiens [14], contrairement à ce qu'a pu affirmer Russell à son propos, la théorie de l'objet ne prescrirait pas de considérer différents niveaux d'existence/d'être [15]. Meinong rejetterait complètement la manière dont Russell décrit sa théorie parce qu'elle échouerait à rendre compte de certains types d'objets. Car de fait, la théorie de l'objet stipule non pas deux, mais trois sortes d'objets. Les premiers possèdent la propriété de localisation spatio-temporelle. Ce sont les objets matériels et les événements mentaux perçus en première personne. On y trouve donc les hortensias, les charançons, les camemberts, les pensées conscientes, les désirs et volitions, les jugements, les assomptions, les sensations, les sentiments. Le deuxième type d'objet sont les objets *subsistants* qui sont essentiellement des objets *fondés*. Ils sont caractérisés par le fait que leur saisie nécessite la médiation d'une activité mentale. Celle dernière ne génère cependant pas l'objet fondé, mais le sélectionne. On peut inclure les nombres dans cette catégorie. Elle correspond grosso-modo à l'ensemble des objets abstraits logés dans un arrière-monde platonicien. Mais comme nous l'avons dit, ces deux premières catégories ne suffisent pas à rendre compte de la totalité des objets. Aussi, Meinong en adjoint-il une troisième, celle des objets ne pouvant ni exister, ni subsister. Parmi ces objets figurent en particulier les objets incomplets et les objets impossibles.

Il arrive en effet que certains objets de pensée soient sans objectifs. Ce sont alors des objets ontologiquement neutres qui n'ont ni existence ni même subsistance. Meinong dit d'eux qu'ils sont *ausserseiende*, c'est-à-dire, par delà l'être, où, pour le dire d'une manière plus littéraire, ils sont comme des « Horla ». Par exemple, la fontaine de Jouvence que cherchait Ponce de Léon n'est ni un objet concret, ni même un objet idéal. Cet objet est au-delà de l'être. Il est donc sans logis ontologique, ou comme l'écrit Meinong, « *heimatlos* ». Ainsi, pour Meinong, de tels objets, à la fois inexistants et actuels, ne sont pas logés dans un arrière-monde platonicien. Ils sont plutôt des choses *ontologiquement affranchies*. C'est aussi ce que résume Roderick Chisholm :

> We may think of physical things and of persons as *concreta* and of attributes, classes and numbers as being *abstracta*. And we might say that a homeless object is an object that is nor

[14] Par exemple (PRIEST, 1995).

[15] Nous pouvons toutefois émettre de sérieux doutes sur le fait que la théorie de l'objet n'imposerait pas différents niveaux d'existence, il semble que cela soit une conséquence inévitable de la dissociation de l'existence et de la quantification.

> a *concretum* nor an *abstractum*. Such an object would be homeless, not only because it is not covered by the usual branches of knowledge, but also because there would seem to be no place for it, either in Plato's heaven or on earth. (CHISHOLM, 1973a, p. 207)

Les personnages de fiction rentrent ainsi dans cette catégorie.

Constatons cependant que la théorie de l'objet implique que les objets dits inexistants sont inévitablement des objets *actuels*. N'en déplaise à Meinong, on a donc bien affaire à une question ontologique. Car il paraît difficile de distinguer « être actuel », d'« être réel ». En outre, la théorie de l'objet repose non seulement sur la thèse que tout objet actuel peut dans l'absolu constituer un objet de pensée, ou thèse de concevabilité de l'actuel (CA), mais aussi, et réciproquement, sur la thèse d'actualité du concevable (AC) stipulant que tout ce qui est concevable est actuel. La thèse AC constitue une théorie spéculaire métaphysique de la représentation, car si tout ce qu'on peut concevoir est actuel, cela implique une sorte de duplication ontologique du réel dans le langage. La théorie des objets engendre une béance entre l'irréel et le réel : le non-réel n'est pas l'irréel (ce qui n'est rien sous aucun rapport). Car il y a à la fois des choses qui sont non-réelles, mais qui sont cependant quelque chose. Il y a des objets qui ne sont pas. En ce sens, la théorie de l'objet contredit le principe Parménidien affirmant que l'être est, et que le non-être n'est pas.

De ce point de vue, la théorie pan-référentialiste défendue par Meinong n'est pas vraiment une théorie mimétique reflétant la réalité de manière spéculaire dans le langage. La référence à des objets purement fictifs est bien autorisée, mais on peut questionner en quoi cette référence est encore une dénotation, en ce sens que de tels objets sont, comme le dit Meinong, « sans logis ».

Caractère ascriptif des termes fictionnels

Selon Meinong, la pensée obéit essentiellement à deux fonctions. La première fonction est thétique, elle renvoie à la capacité à appréhender un « être » (*sein*). Je peux par exemple penser à mon hortensia. La deuxième fonction, dite synthétique, renvoie à la capacité d'appréhender un « être-tel » (*sosein*). Il reste à expliquer comment attribuer des propriétés à des inexistants. En 1903, Ernst Mally (MALLY, 1912) a complété la théorie de l'objet en proposant le principe d'indépendance qui stipule que l'*être-ainsi* (*sosein*) d'un objet n'est pas affecté par son inexistence. Pour un objet, le fait d'avoir ou non telle propriété ne dépend pas du fait d'exister. Ce principe combine en fait deux affirmations différentes : la première est le principe de caractérisation qui dit que tout objet possède les propriétés qui le caractérisent comme étant cet objet-là de pensée, la deuxième est l'innocence ontologique ou assomption ontologique qui admet qu'il y ait des propositions vraies à propos des inexistants. Lors même que des choses seraient inexistantes, elles pourraient toutefois posséder certaines propriétés. Admettons que Pierre pense à la montagne en or. Nous pouvons raisonnablement dire de cette montagne qu'elle est en or, qu'elle est une montagne, qu'elle est inexistante, et qu'elle est pensée par Pierre. Il s'agit justement, pour Meinong, d'expliquer, par une théorie adéquate, comment des propriétés peuvent être attribuées, de manière intentionnelle, à des inexistants.

C'est un fait que nous associons aux noms fictionnels certaines descriptions. Un lecteur de Conan Doyle associera ainsi le nom de Sherlock Holmes à la qualité d'être un brillant détective, de résider à Londres, etc. Les anciens Grecs associaient le nom de Zeus au roi des dieux vivant

au sommet du Mont Olympe, et l'astronome Urbain Le Verrier associait le nom de Vulcain à une planète dont l'orbite serait située entre le Soleil et Mercure. Associer un nom à une description est une chose, mais la théorie de l'objet ne se contente pas de cela, elle considère en outre que ces noms sont eux-mêmes associés à des objets qui, en retour, instancient réellement les propriétés contenues dans les descriptions qui en sont faites. La manière dont un objet possède les propriétés utilisées pour le décrire a été désignée par Richard Routley (ROUTLEY, 1980) sous le nom de « postulat de caractérisation ». Il est aussi connu sous le nom de « principe de compréhension ». Ce principe a été diversement décliné par les successeurs de Meinong, nous y reviendrons.

Il reste à connaître plus précisément encore la manière dont un objet inexistant est attaché à des propriétés. Comment sont-elles transférées aux objets *heimatlos* ? D'où proviennent de telles capacités ascriptives et génératives ? Du langage ? Des pensées ? Il semble qu'admettre ceci, c'est doter l'un ou l'autre d'une extraordinaire puissance ontologique. Mais Meinong ne peut pas l'attribuer au langage, puisque comme nous l'avons déjà vu, il prétend expliquer la pertinence du langage par la pertinence des pensées. Or ce pouvoir ontologique ne peut pas non plus provenir de la seule pensée. Comme le souligne Alberto Voltolini (VOLTOLINI, 2006, p. 10), la théorie de l'objet n'est pas une théorie phénoménologique des objets inexistants. De fait, même si Meinong concède à Bolzano qu'une intentionnalité vise toujours un objet, il soutient en revanche, et avec Twardowski, qu'un tel objet demeure indépendant des pensées. Ce ne sont donc pas les pensées qui confèrent leurs propriétés aux objets *ausserseiende*. Il faut ainsi y voir un mystère que les meinongiens ne parviennent pas à expliquer, et qui est d'ailleurs la source principale des désaccords occasionnés entre leurs différentes interprétations de la théorie de l'objet.

En outre, la théorie de l'objet semble se heurter systématiquement à l'explosion logique parce que les objets inexistants semblent nécessairement être également des objets contradictoires, autrement dit, des objets impossibles. Détaillons cette difficulté. Comme nous l'avons expliqué plus haut, l'idée générale de la théorie de l'objet consiste à traiter l'existence comme une propriété des particuliers. Il s'agit bien sûr d'une conception qui s'oppose à toute la tradition, de Hume à Quine, qui considère que l'existence n'est pas un prédicat. Meinong admet cependant qu'on pourrait y voir un paradoxe : « Qui aime les paradoxes pourra fort bien dire : il y a des objets à propos desquels on peut affirmer qu'il n'y en a pas. » (MEINONG, 2000, p. 73) Mais puisque « être existant » est une propriété, comme le sont également « être en or » et « être une montagne », le prédicat peut ou non être appliqué selon l'objet qui est visé. On pourra ainsi appliquer les prédicats « être en or », « être une montagne », « être inexistante » et « être l'actuel objet de ma pensée » à la Montagne d'or, mais en revanche, il ne convient pas de lui appliquer « être existant ». Affirmer qu'« il y a des choses qui n'existent pas » n'est donc pas, en soi, contradictoire, en ce que la phrase n'est pas équivalente à « certaines choses existantes n'existent pas ».

Après la publication de « On Denoting », c'est-à-dire après avoir pris ses distances avec la théorie de l'objet, Russell l'accuse de mener à des contradictions explosives. En effet, le principe d'indépendance est une entorse aux principes logiques de non-contradiction :

> (...) the chief objection is that such objects, admittedly, are apt to infringe the law of contradiction. It is contended, for example, that the existent present King of France exists, and also does not exists ; that the round square is round, and also not round; etc. But this is intolerable; and if any theory can be found to avoid the result, it is surely to be preferred. (RUSSELL,

1905)

Russell souligne que postuler l'être d'un carré rond mène à des raisonnements absurdes. Si toute association de propriétés constitue un objet, c'est alors encore le cas lorsqu'on agglomère deux propriétés contradictoires comme le fait d'être rond et de ne pas être rond. On ne peut alors éviter l'explosion logique. Admettre des objets inexistants consiste non seulement à admettre des objets contradictoires, comme la coupole ronde-carrée, mais encore à tolérer l'existence de toute chose, puisque, du point de vue de la théorie de l'objet, l'existence peut être attribuée de manière arbitraire à n'importe quel contenu de pensée. En effet, la théorie de l'objet stipule qu'un objet de pensée possède littéralement les propriétés qui servent à le décrire. Or il ne semble pas que ses propriétés soient limitées à celles-ci. Car il en possède inévitablement d'autres, comme d'être mon actuel contenu de pensée ou d'être inexistant. Or, afin de constituer un nouvel objet, il est toujours d'adjoindre aux propriétés servant à décrire un objet inexistant la propriété supplémentaire d'existence. Pour revenir à l'exemple de la montagne d'or, nous pourrions concevoir à partir d'elle un nouvel objet constitué des propriétés « d'être une montagne », « d'être en or » et « d'être existante ». Dès lors, cet objet serait à la fois existant (par ascription) et à la fois inexistant (par définition d'un objet inexistant). Le caractère ascriptif des inexistants semble donc impliquer de charrier dans l'ontologie un nombre infini d'objets contradictoires.

On peut d'ailleurs généraliser le problème. Graham Priest (PRIEST, 2005, p. 83) souligne que pour n'importe quelle propriété A, on peut considérer la propriété B définie par $x = x \wedge \mathrm{A}$. En appliquant le principe de caractérisation à la propriété B, on a alors un objet c_{B} tel que $c_{\mathrm{B}} = c_{\mathrm{B}} \wedge \mathrm{A}$, d'où il découle que A est prouvée.

Différentes réponses ont été élaborées par les néo-meinongiens pour obvier l'explosion logique et répondre aux objections de Russell soulignant le caractère foncièrement contradictoire des objets inexistants.

Dans son essai « Nonexistent objects », Parsons introduit une comparaison entre la croyance en la réalité de propriétés non instanciées, et la croyance d'objets inexistants (PARSONS, 1980, p. 1).

Comme le déplorait déjà Meinong, Parsons constate que le paradigme scientifique contemporain est préjudiciable aux inexistants. Ce paradigme proviendrait des critiques de Russell et de Quine à leur encontre, jugées habituellement définitives, mais que Parsons s'emploie à rejeter. L'une des thèses fortes du paradigme est le refus de distinguer la forme logique de la proposition exprimée par une phrase comme « il y a des tables » de la forme exprimée par la phrase « les tables existent » (PARSONS, 1980, p. 7). Parsons considère ainsi que l'expression « il y a » (ou *some*) n'est pas nécessairement chargée de manière existentielle :

> By using "some F", or "an F", or "at least one F" we usually communicate "some existing F", "an existing F" or "at least one existing F." and one natural explanation of this is that the word "exists" is logically redundant, so that these pairs of phrases are all synonyms. However, that is not the only possible explanation, and it is a wrong one. (PARSONS, 1982, p. 365)

Parsons souligne ainsi l'importance de distinguer l'acte de « référer à rien » de celui de « ne rien référer ». Il ne faut pas confondre le fait de référer à quelque chose qui n'existe pas et le fait

de ne pas référer. Le nom « Sherlock Holmes » et la description définie « le détective d'*A Study in Scarlet* » dénotent bien un objet, quoique celui-ci soit inexistant. En ce sens, le manque de référence à un objet existant n'est que relative. En revanche, l'évocation du « fameux détective d'Alice au pays des merveilles » n'est pas un acte de référence relatif à un inexistant, mais un échec absolu de référence.

Parsons argumente l'admission ontologique des objets inexistants en légitimant les entités par leur effet explicatif. Si les objets matériels sont admis dans l'ontologie en raison de ce qu'ils simplifient nos explications du monde, le même argument devrait être encore valable pour les objets inexistants : puisqu'ils simplifient les rapports à nos expériences, nous devrions les admettre. En ce sens, si leur statut peut faire l'objet de soupçons, ce n'est pas très différent des objets concrets :

> Quine thinks that our belief in physical objects is a convenient myth. Perhaps nonexistent objects are no better off. But they are no worse off. (PARSONS, 1982, p. 370)

Parsons introduit une distinction entre le type de propriétés prédicables afin de constituer un ensemble corrélé à un objet. Les propriétés nucléaires sont les propriétés que l'on peut attribuer librement aux objets existants ou inexistants, comme par exemple d'« être en or » et d'« être une montagne ». En revanche, les propriétés extra-nucléaires ne sont pas librement attribuables. Il s'agit des propriétés existentielles (existe, est mythologique, est fictionnel, etc.), intentionnelles (est pensé, est cru, est adoré, etc.), modales (est possible, est nécessaire, etc.), techniques (est complet, etc.). La stratégie de Parsons consiste donc à bloquer toutes les prédications qui pourraient rendre contradictoire l'objet corrélé à l'ensemble des propriétés qui le décrit. Les propriétés extra-nucléaires sont donc exactement celles qui engendrent des problèmes logiques. Cela consiste finalement à affirmer que la réalité obéit à la logique, ou pour le dire autrement, à constituer une solution *ad hoc* taillée sur mesure pour éluder certains problèmes des objets meinongiens.

Toutefois, il peut paraître assez étrange d'affirmer que certaines propriétés soient attribuables aux objets fictionnels, et que d'autres non. Edward Zalta à proposé une théorie alternative qui évite ce recours en développant une idée d'Ernest Mally :

> [Mally] distinguished two relationships which relate objects to their properties. On Mally's view, properties can DETERMINE objects which do not in turn SATISFY the properties. (Edward ZALTA, 1983, p. 11)

Pour Zalta ce ne sont pas les types de propriétés qui divergent, mais leur manière d'être instanciées. Contrairement à Parsons, Zalta considère que ce n'est pas la nature des propriétés attribuables aux objets fictionnels qui est restreinte, mais la manière dont elles sont attribuées. Autrement dit, alors que pour Parsons il y a deux types de propriétés et une seule manière de les attribuer, pour Zalta, il n'y a qu'un type de propriétés mais deux manières de les attribuer. Plutôt que de relativiser le principe de caractérisation à différents types de propriétés, nucléaires ou extra-nucléaires, Zalta considère que c'est l'auxiliaire « être » qui doit être désambiguïsé dans les expressions du type « m *est* une montagne en or ». Les objets fictionnels possèdent leurs propriétés d'une manière différente de celles des objets existants. Des propriétés extra-nucléaires peuvent également être attribuées aux objets inexistants, et on peut dire de Sherlock Holmes qu'il existe. Mais contrairement aux objets existants, les objets abstraits n'exemplifient pas leurs propriétés, ils les *encodent* :

> By relying on the model of Fregean senses as Meinongian objects that I've developed elsewhere, we can say that the sense of the name "Holmes" with respect to Doyle at this point in time is also an abstract, Meinongian object which encodes certain properties. This object encodes the properties attributed to Holmes in the "core story". (Edwards ZALTA, 2003, p. 252)

Ainsi, Sherlock Holmes encode l'existence, mais il ne l'exemplifie pas. Par contre, il exemplifie l'inexistence, mais ne l'encode pas. La différence entre l'encodage et l'exemplification semble donc marquer une façon de produire la propriété : stérile dans un cas, productive dans l'autre.

Zalta confère deux interprétations au langage obtenu en distinguant l'existence de la concrétude. La première interprétation consiste à considérer que $E!$ exprime l'existence et que le quantificateur $\exists$ est ontologiquement neutre, exprimant « quelque » (*some*). Et donc, en ce cas, les objets abstraits sont les objets à qui manque nécessairement l'existence. Le principe de caractérisation affirme alors que pour toute formule Φ, qu'il y a un objet abstrait (autrement dit nécessairement non-existant) qui encode les propriétés F satisfaisant Φ et uniquement celles-là. La deuxième interprétation consiste à considérer que $E!$ exprime la concrétude et que le quantificateur $\exists$ exprime l'existence. Sous cette interprétation, les objets abstraits sont nécessairement non-concrets et le principe de caractérisation affirme alors que pour toute formule Φ, qu'il y a un objet abstrait qui encode les propriétés F satisfaisant Φ et nécessairement celles-là.

De fait, la conception néo-meinongienne développée par Parsons et Zalta consiste en une forme de logique libre qui utilise deux quantificateurs ontologiquement déchargés Π (pour tout) et Σ (pour quelque) adjoint à un prédicat d'existence $\mathrm{E}!$. Ainsi, ce qu'on écrit traditionnellement $\forall x \Phi x$ est retraduit : $\Pi x(\mathrm{E}!x \rightarrow \Phi x)$ et $\exists x$ est retraduit par $\Sigma x(\mathrm{E}!x \wedge \Phi x)$. C'est donc un point de vue qui est en opposition totale avec le slogan de Quine « exister, c'est être la valeur d'une variable liée ». Cette conception pose d'inévitables problèmes d'individuation dont nous nous occuperons dans le chapitre suivant. Mais pour l'instant, soulignons que de tels objets rejoignent les objets abstraits platoniciens, et qu'en quelque sorte, les théories de Parsons et de Zalta s'éloignent ainsi de la théorie de l'objet de Meinong. Car en admettant les *ficta* comme des objets idéaux, ces théories renouvellent l'idée d'une duplication du réel dans le langage.

Dans ce qui précède, nous avons vu que le développement de la théorie de l'objet par les successeurs de Meinong l'a progressivement ramenée à une théorie mimétique de type platonicienne. Néanmoins, dans la théorie de Zalta, il y a un clivage entre les propriétés selon le type de rapport externe ou interne qui les gouverne. Cette distinction est l'enjeu principal d'une autre grande théorie pan-référentialiste, la théorie des objets dépendants, que nous allons examiner maintenant.

3.2.3 La théorie des objets abstraits

Remarquons que, lorsqu'on parle d'existence à propos d'un personnage de fiction, comme Sherlock Holmes, on évoque toujours deux choses en même temps. D'une part, nous pensons à quelque chose qui est interne au jeu de fiction, évoquant un personnage censé être fait de chair et d'os, né en 1850 de l'union de deux êtres humains. D'autre part, nous pensons aussi à la création d'un auteur, création qui apparaît pour la première fois en 1887 dans *A Study in Scarlet*. De ce point de vue, on peut dire que le statut ontologique des personnages de fiction est à cheval entre la réalité et

l'irréel. En un premier sens, Sherlock Holmes n'existe pas, au sens où existe Emmanuel Macron. Nous ne pourrions rencontrer Sherlock Holmes, aussi propices fussent les conditions temporelles de cette rencontre. Néanmoins, en un second sens, Sherlock Holmes existe comme fait un culturel : il a ses admirateurs, un musée lui est dédié et différents films et séries télévisées le représentent. Ainsi, il semble que nous puissions à la fois parler de l'inexistence de Sherlock Holmes, et dire aussi qu'il a été créé par Conan Doyle. Il y a donc, en quelque sorte, une double caractéristique, contradictoire, mêlant l'existence à l'inexistence. Pour sortir de l'impasse, il s'agit de distinguer deux choses plutôt qu'une : d'un côté, le personnage de chair et d'os qui manque au monde, mais qui existe dans un cadre fictionnel, et d'autre part, l'artefact culturel qui, lui, fait bien partie de notre monde mais qui manque au monde de la fiction. Nous pourrions admettre que la question de l'existence de Sherlock Holmes a un sens relativement au cadre où elle est posée. Nous dirions à la fois que, en dehors de son cadre linguistique de référence, Sherlock Holmes n'existe pas. Mais qu'en revanche, à l'intérieur du cadre fictionnel prescrit par Conan Doyle, il existe bien.

Les *ficta* sont-ils des objets éternels ?

Caractériser les personnages de fiction par le principe d'ascription les assimile inévitablement à des réalités platoniciennes, autrement dit, à des objets éternellement disponibles, à des idéaux, des objets abstraits. Car de tes objets n'attendent pas d'être pensés par un auteur pour être des objets. Ils ne sont pas liés à des contraintes historiques de création et de persistance. Que Conan Doyle existe ou non ne changerait rien concernant le *fictum* associé au nom « Sherlock Holmes ». La désolidarisation des *ficta* meinongiennes à toute intentionnalité et à toute temporalité est pourtant contre-intuitive. Le caractère intentionnellement créé des personnages semble être une composante *fondamentale* de leur mode d'existence. Les personnages de fiction sont créés à un certain moment, selon certaines intentions, par un certain auteur, dans un certain texte. Ils peuvent également disparaître à un autre moment si leurs conditions de persistance ne sont pas satisfaites, par exemple s'ils disparaissent de la mémoire collective.

Si l'on accepte le fait qu'un contenu de pensée renvoie systématiquement à un objet, une autre perspective, développée par Roman Ingarden, consiste à développer, à côté d'une ontologie formelle qui s'intéresse à la forme de ce qu'il y a, et à côté d'une ontologie matérielle qui s'intéresse à l'existence des particuliers concrets, une ontologie existentielle basée sur le principe qu'il y a plusieurs modes d'être, c'est-à-dire d'adhérer à un pluralisme ontologique. Il ne s'agit plus d'opposer être et exister dans un dualisme ontologique comme le faisait encore Russell avant 1905, ni d'admettre que l'existence est une propriété attribuable aux choses, mais de convenir que les choses ont plusieurs façons d'exister. L'idée est de situer d'emblée les objets de pensée (question ontologique) et de déterminer ensuite si de tels objets existent ou non, où plutôt, de quelle manière ils existent (question métaphysique). Ainsi, la question de savoir s'il existe des objets d'un certain type, comme les répliques, les doubles purs ou les coïncidents purs, n'est plus vue comme une question ontologique mais comme une question métaphysique.

L'intérêt d'une telle manière de voir les choses est qu'elle permet une analyse *a priori* de la teneur de nos représentations, indépendamment de la question de savoir si elles constituent des objets existants. Mais cette forme de libéralisme des représentations peut induire, en contrepartie, une ontologie coûteuse, et tombe sous la menace d'une sorte de pan-réalisme.

La théorie des objets abstraits repose ainsi sur l'idée que les choses peuvent exister de différentes manières. Soit de manière concrète comme les hortensias, les charançons, les camemberts et les personnes, soit de manière purement abstraite et autonome, comme les entités mathématiques des platoniciens, soit encore comme des entités abstraites et dépendantes. L'existence est vue comme un principe primitif, il s'agit d'un mode d'être qui peut prendre différentes *intensités*. Certains objets existent en soi, indépendamment de toute autre chose, et d'autres n'existent qu'à travers l'existence d'objets plus fondamentaux. Cette idée est donc ancrée sur une conception stratifiée de l'existence. La dépendance existentielle se rapproche ainsi d'une conception du double par proximité : les entités ontologiquement dépendantes le sont de la même façon que l'ombre est dépendante de l'objet matériel dont elle trace la silhouette, ou que le reflet dépend de l'objet qui en produit l'image spéculaire. Le rapport aux objets fictionnels n'est cependant pas mimétique, mais hétéronomique, comme nous allons le préciser maintenant.

Des modes d'existence chez Roman Ingarden

Roman Ingarden a établi les bases d'une théorie intentionnaliste des œuvres littéraires par la thèse de dépendance ontologique que les *ficta* entretiennent avec les actes mentaux. Le point de départ d'Ingarden consiste à répondre à la problématique de la disparition. Comment pouvons-nous comprendre la disparition d'une chose particulière ? L'auteur nous invite à considérer une lampe particulière, puis à imaginer qu'elle cesse brusquement de persister, soit qu'elle soit détruite, soit qu'elle soit transformée en autre chose. Remarquons dans ce cas que la lampe ne serait pas « partie » comme si on la déplaçait d'une pièce à une autre. Déplacée, la lampe verrait certaines de ses propriétés modifiées, mais elle resterait encore une lampe. Imaginons plutôt que la lampe soit réduite en cendre. Brûlant, la lampe est encore là, jusqu'au moment où elle n'est plus. De même, la chenille est encore là jusqu'au moment où elle devient un papillon. Plutôt que d'y voir un vague lié à nos usages maladroits des concepts pour décrire des processus continus de transformations physico-chimiques complexes, Ingarden considère qu'il y a là un véritable mystère [16]. Il parle d'une « rupture radicale », l'occasion d'une transition de l'être au non-être. L'auteur souligne que nous avons, de ce point de vue, peu de moyens de comprendre ce que signifie, pour une chose, de disparaître.

> The lamp, along with all of the properties which a moment ago it still "possessed," has vanished. Somewhere within this process of annihilating transformation a radical rupture of being occurs, a radical leap by the lamp into a nothing, from being into non-being, irrespective of how difficult it might be for us to grasp the "locus" of the rupture, of the leap, and how difficult it might be to comprehend the sense of that "transition" from being into non-being. (INGARDEN, 2013, p. 103)

Que signifie exister ? L'existence est-elle une propriété qu'une chose pourrait perdre ? Dans *The Controversy over the Existence of the World*, Roman Ingarden propose une expérience de pensée d'indiscernables. Considérons deux lampes indiscernables, sinon que l'une des lampes existerait, mais pas l'autre. Qu'est-ce qui différencierait de tels objets de pensée ? Ingarden, qui soutient la thèse parménidienne, prévient que le non-être est la négation de tout mode d'être :

[16] Le mystère de la disparition prend bien entendu une intensité encore plus grande quand il s'agit d'une personne.

« Non-being [...] is no mode of being but rather the outright *privation* of all being ». L'être n'est pas séparable de l'existence, il ne s'étend pas par delà l'existence dans une forme de subsistance immatérielle. Il ne s'agit donc pas, contrairement au Russell d'avant 1905, de considérer qu'une « subsisterait » en deçà de l'existence. Il ne s'agit donc pas d'ajouter ou de retrancher des propriétés d'un objet à l'autre.

> [...] what would have to somehow be attached to that non-existing "lamp" in order to get it to exist. What distinguishes it from the existing lamp? There is only one possible answer to this question, which was given by Kant (and actually, already by Hume): no new property and no new formal moment attaches to that non-existent "lamp". And yet something ultimately primitive, though not quite absolutely simple, makes itself obtrusive (*bemerkbar*) in the most remarkable fashion in everything that constitutes the existing lamp, in everything that characterizes it and belongs to its form: that it does indeed exist – and it is precisely this, and only this, that distinguishes it from the non-existing "lamp". Now it is the totality of what makes the existing lamp differs in this re- markable manner from the non-existent "lamp" (while holding firmly fixed in our mind all of the latter's material and formal moments) that we term its mode of being. (INGARDEN, 2013, p. 102)

Si aucune propriété ne permet de différencier une lampe existante d'une lampe qui n'existerait pas, on doit en déduire que l'existence n'est pas une propriété des objets. L'existence n'est pas analysable, c'est une donnée primitive, irréductible, que Ingarden désigne aussi comme un *mode d'être*. La conception de l'existence d'Ingarden est donc classique et s'oppose à celle de Meinong. La notion d'existence est primitive et non analysable.

Ingarden définit quatre modes d'existence : absolu, réel, idéal et possible (INGARDEN, 2013, p. 99) : « The "being-real [Real-Sein], being-ideal [Ideal-Sein], being-possible [Möglich-Sein] of something", and the like, are modes of being of this something » (INGARDEN, 2013, p. 99). Il s'agit pour l'auteur de défendre l'idée que les choses ne sont pas seulement diversifiées par leur nature, mais aussi par leur manière d'exister. Ingarden avance ainsi l'idée que deux entités pourraient être indiscriminables par leur propriétés, mais avoir chacune un mode d'existence différent. Le mode d'être absolu est celui que les théologiens confère au démiurge. Le mode d'être idéal est le mode d'être que les platoniciens confèrent aux entités qui sont indépendantes de l'esprit et qui n'existent pas de manière physique, comme les nombres. Le mode d'être réel est celui des entités spatio-temporelles qui sont indépendantes de l'esprit, les concrets particuliers. Le mode d'être possible, intentionnel, est celui des entités qui sont dépendantes de l'esprit comme les objets sociaux et les objets conventionnels.

Mais la thèse d'Ingarden ne s'arrête pas à cette distinction des modes d'existence. Car ces modes sont eux-mêmes structurés selon des « moments existentiels », moments qui sont regroupés en quatre sortes d'oppositions autour de la notion de dépendance ontologique à travers les couples antithétiques suivants : autonomie/hétéronomie, originalité/dérivation, séparabilité/inséparabilité et indépendance/dépendance [17].

L'autonomie désigne le moment d'un mode d'existence tel que l'entité existe directement par

[17] Je reprends ici le vocabulaire tel qu'il est traduit par Sébastien Richard dans un exposé au collège de France (https://www.college-de-france.fr/site/claudine-tiercelin/symposium-2017-06-30-17h10.htm)

elle-même, n'ayant pas besoin d'un autre fondement qu'elle-même afin de subsister en tant qu'entité. Un objet autonome est tel que son fondement est immanent. Telle est le moment existentiel d'entités concrètes comme Emmanuel Macron ou les charançons. On parlera au contraire d'hétéronomie pour qualifier le mode d'existence d'une entité dont la persistance dépend d'autre chose qu'elle-même, comme par exemple Sherlock Holmes (INGARDEN, 2013, p. 109-110).

Ensuite, l'originarité désigne le mode d'existence d'une entité qui n'est pas *produite* par une autre entité. Une entité originelle dotée d'une *quiddité* existe dans la mesure où elle ne peut pas ne pas exister par le fait même de posséder cette essence. On dira au contraire qu'une entité est dérivée si elle résulte d'une création par une autre entité, si elle est « incapable d'exister en vertu de sa seule essence » (INGARDEN, 2013, p. 118). Par exemple, puisque Emmanuel Macron et que Sherlock Holmes sont tous les deux produits par autre chose qu'eux-mêmes, ils tombent tous deux sous le moment existentiel de dérivabilité.

Le moment de séparabilité (ou autosuffisance) qualifie le moment du mode existence des entités dont l'existence ne nécessite pas qu'elles soient parties propres d'une entité plus grande dans une relation entre une partie et un tout. Dans le cas contraire, si l'existence d'une entité nécessite son étroite collaboration comme partie d'un tout, on parlera d'inséparabilité. Ainsi, ce que certains philosophes pourraient qualifier de « trope de la démarche de Macron » est une entité qui ne peut être séparée de Macron, elle-même étant séparable. Contrairement à l'autonomie, la relation de séparabilité est donc méréologique. Sherlock Holmes est donc séparable, il ne dépend pas de manière continue de l'existence ontologique de Conan Doyle.

Enfin, le dernier couple de moments existentiels exprime l'indépendance existentielle à proprement parler, qui corrèle les substances entre elles sans nécessiter la coexistence dans un tout.

> Namely, it is possible for an entity to be selfsufficient and still require, in virtue of its essence, the existence of some other selfsufficient entity for its own "continued subsistence". We then refer to the first entity as existentially dependent. (INGARDEN, 2013, p. 153)

Macron est, relativement à ses parents, existentiellement dépendant en tant que fils, et Sherlock Holmes est existentiellement dépendant de l'existence d'Arthur Conan Doyle en tant que créature.

Ces moments existentiels sont eux-mêmes encore déclinables selon huit moments en rapport à la temporalité, précisément, par trois moments temporels que sont l'actualité (présent), la post-actualité (passé) et la possibilité empirique (futur), mais aussi par le moment de non-actualité (les objets qui n'existent pas dans le temps). Il y a encore deux couples de moments : la permanence/fragilité (existe ou non de manière permanente) dans le temps et enfin par le couple de moments fissuration/non-fissuration qui se rapporte à la manière dont une entité présente ne peut pas sortir du présent, ce qui permet le clivage entre les conceptions endurantiste ou perdurantiste de la persistance.

Les modes d'être sont ainsi combinables selon $2^4 \times 2^4 = 256$ moments. Bien sûr, il y a certaines associations impossibles, parce que certains moments existentiels sont incompatibles ou parce que certains en impliquent d'autres. Par exemple, les objets originaux sont nécessairement autonomes. Mais certaines combinaisons demeurent possibles. Par exemple, les personnages de fiction peuvent être définis comme des objets dont le mode d'existence est idéal (ils n'existent pas dans l'espace) et dont ce mode est structuré par les moments existentiels suivants : hétéronomie (ils

tirent leur existence de quelque chose qui leur est extérieur), non-originalité (ils sont créés par un auteur), autosuffisance (ils ne nécessitent pas d'appartenir à une entité plus grande) et dépendance (ils sont dépendants des esprits qui y pensent). Aussi, Ingarden ne retient-il finalement que quinze moments existentiels.

Les fictions, et les objets culturels plus généralement, ont la particularité d'avoir un double mode d'existence, puisqu'ils dépendent d'une part, des actes de créations de leur auteur et d'autre part des supports matériels qui convoient leurs descriptions en tant qu'objet de fiction. Ils dépendent à la fois des êtres intentionnels qui les pensent et des livres qui les véhiculent.

Les objets fictionnels sont donc des objets hétéronomes dérivés en ce qu'ils relèvent de productions culturelles qui dépendent d'auteurs. En outre, ils sont séparés, puisqu'ils ne sont pas des parties propres de leur auteur. Ainsi, pour le mode d'être purement intentionnel, on a les deux catégories suivantes :

Mode d'être intentionnel	
Hétéronomie	Hétéronomie
Dérivabilité	Dérivabilité
Non-actualité	Non-actualité
Séparabilité	Non-séparabilité
Dépendance	
Exemples	
Sherlock Holmes	La sagacité de Sherlock Holmes

En quoi la théorie d'Ingarden est-elle réaliste ? Cette théorie considère que les *ficta* ne sont pas des objets concrets. Il s'agit donc d'objets abstraits, au sens où ils n'ont pas d'extension spatiale. Il pourrait alors s'agir d'une théorie conceptualiste. Mais Ingarden refuse aussi que les *ficta* soient de purs contenus mentaux. Ils sont réifiés en objets, autrement dit, ils ont une pondération ontologique. La stratégie d'Ingarden reste donc réaliste en ce qu'elle affirme la nécessité de stipuler des objets intentionnels pour expliquer les contenus de conscience. Les relations de dépendance semblent néanmoins fructueuses, nous pourrons donc les utiliser sans néanmoins souscrire à une théorie des objets intentionnels. En dépit de ce qui apparaît encore comme un réalisme soucieux de légiférer le statut ontologique des fictions, Ingarden reconnaît leur caractère essentiellement dépendant. L'œuvre est moins un programme qu'un dessin technique, un dessin de définition destiné à être rempli par le récepteur de l'œuvre. L'œuvre n'est ainsi construite que par un effet de collaboration entre l'auteur et le récepteur et n'existe qu'à travers ce processus.

La théorie des objets abstraits

La théorie d'Ingarden a connu des développements à travers les travaux de Saul Kripke. La théorie sémantique de Kripke n'a pas directement pour objet de répondre au problème des fictions, mais constitue une critique du descriptivisme au profit d'une conception millienne de la référence des noms propres. Selon Kripke, non seulement la théorie sémantique de Russell ne permet pas de

garantir que la référence visée par un nom est bien atteinte par le moyen de la description définie employée, mais plus encore, la description n'est même pas nécessaire à la dénotation. Car l'on peut employer le nom d'une personne pour y référer sans en connaître la moindre description. Le nom dénote directement l'individu sans passer par la médiation d'une description. La fixation d'un nom est généralement établie par un acte initial de désignation directe d'un individu. La communication de l'emploi du nom est lié à une chaîne de transmission qui remonte à cet acte inchoatif de baptême. La fixation de la référence étant directe, cela explique qu'on puisse viser le bon objet même lorsqu'on fait usage de descriptions erronées à propos de l'objet que dénote ce nom.

Mais qu'en est-il des noms fictionnels comme « Sherlock Holmes » ? En quoi pourrait consister un acte de baptême d'un immatériel ? Quelle chaîne historique pourrait véhiculer la signification du nom sachant que le personnage n'a aucune existence causale ?

Plusieurs travaux de Kripke, ultérieurs à *Naming and Necessity*, concernent l'application de la théorie des désignateurs rigides aux noms fictionnels. L'article intitulé « Vacuous Names and Fictional Entities » et les lectures données à Oxford en 1973 dans le cadre des « John Locke Lectures » indiquent une évolution de sa théorie des fictions vers une sorte de créationnisme promulguant les fictions au rang d'entités abstraites créées par leurs auteurs.

Dans un premier temps, Kripke a soutenu une version descriptiviste des noms fictionnels. Puisque ceux-ci ne réfèrent à rien de réel, les noms propres fictionnels ne se comportent pas vraiment comme d'authentiques noms propres. De fait, ce que nous savons au sujet d'un personnage de fiction ne réside que dans les attributions conférées par des mots par son auteur. Il ne peut y avoir ici une transmission par une chaîne causale. Mais cette conception conduit alors à distinguer des noms fonctionnant de deux manières différentes. D'une part les authentiques noms propres qui dénotent directement leur objet et qui se comportent comme des désignateurs rigides, et d'autre part, ceux, à l'instar des noms fictionnels, qui fonctionnent comme des descriptions définies. La signification d'un nom propre ne devrait pourtant pas varier selon le contexte réalité/fiction. Et comment comprendre alors le nom de personnages historiques utilisé dans des fictions, comme « Napoléon » qui apparaît dans les romans de Balzac et de Stendhal ? En outre, la coexistence d'une théorie de la référence directe pour les noms d'existants avec une théorie descriptive des quasi-noms produirait une bien étrange migration de la référence lorsqu'un objet qu'on croyait fictif s'avère finalement existant. Et inversement, des descriptions définies d'individus présumés inexistants peuvent avoir des occurrences différentes dans la réalité. Le Verrier avait fait l'hypothèse de la planète Vulcain, désignée alors comme la planète la plus proche du soleil. Bien que Vulcain n'existe pas, cette description correspond bien à quelque chose de réel, à savoir, la planète Mercure. Il apparaît que tenter de juxtaposer deux théories antagonistes de la référence, l'une pour un contexte factuel et l'autre pour un contexte fictionnel, conduit à des paradoxes inextricables. Une telle stratégie s'avère donc peu convaincante et Kripke va finalement l'abandonner.

Dans un deuxième temps, le traitement que consacre Kripke aux noms propres fictionnels part du constat que nous devrions accepter deux types de réponses aux questions existentielles à propos des personnages de fiction.

D'une part, nous devrions admettre qu'il n'y a rien dans le monde réel qui corresponde à « Sherlock Holmes ». Nous devrions donc, extérieurement à la fiction, souscrire à l'idée que Sher-

lock Holmes n'existe pas. En outre, rien n'aurait pu même être Sherlock Holmes. Sherlock Holmes ne peut exister dans *aucun* monde possible. Si un individu réel avait rempli toutes les descriptions qui échoient à Holmes, cet individu n'aurait pas été Holmes lui-même, mais au mieux une réplique d'Holmes. Cette conclusion est dérivable de deux hypothèses. La première, c'est que la réalité ne doit pas être partagée entre l'existence et l'inexistence. Il n'y a que de l'existence. La réalité aurait certes pu être plus fournie en individus, mais il n'y a aucun sens à dire d'un personnage de fiction qu'il aurait pu exister. La seconde hypothèse, c'est que dans l'activité de fiction, Holmes est représenté comme un particulier, quelque chose d'individué, et non comme une sorte de cahier des charges que diverses personnes pourraient parvenir à remplir. Les noms propres indiquant de manière rigide leurs référents, si quelqu'un venait à remplir les descriptions auxquelles répond Holmes, cela ne ferait pourtant pas de lui Holmes :

> Under what circumstances, to mention the case that I noted before, would Sherlock Holmes have existed, given that the name in fact has no referent? Well, not simply if someone or other would have done these things in the story, because the name 'Sherlock Holmes' is supposed to refer to a particular person rigidly. One cannot say, 'Well, it does not designate a real person, but just a (merely) possible person', whether one likes such an ontology or not, because many possible people might have done the things in the story. In fact, some actual people might have done the things in the story, if the circumstances had been different, in another possible world. Charles Darwin, if had he decided to go into another line of work, might have made an excellent detective around London at the time and fought with some analogue of Moriarty. This is not to say of him, or of anyone else, that he would have been Sherlock Holmes or might have been Sherlock Holmes. He could have played the role of Sherlock Holmes, he could have fulfilled the stories that are told about Sherlock Holmes. But if the pretense about Sherlock Holmes is that 'Sherlock Holmes' designates someone rigidly, one cannot say which person would have been designated. There is no criterion to pick out one as opposed to another; one should just say that this name does not designate. (KRIPKE, 2011, p. 59)

Mais d'autre part, du point de vue interne à la fiction, il serait faux de prétendre que Sherlock Holmes n'existe pas. Dans « Une étude en rouge », Sherlock Holmes n'est pas une fiction, mais bien une personne réelle. Ici, Kripke utilise la théorie des *speech acts* de John Searle, en particulier, la théorie des assertions feintes (*pretense*) afin de caractériser de manière pragmatique l'usage que nous faisons des noms fictionnels. Les noms fictionnels doivent fonctionner de la même manière que les autres noms propres. Or, c'est bien le cas, constate Kripke, parce que les assertions fictionnelles ne doivent pas être prises au sérieux, mais considérées sous le filtre d'un jeu de simulation consistant à faire comme si les noms qu'elles contiennent étaient les noms d'authentiques individus :

> To write a work of fiction is to imagine - spin a certain romance, say - that there really is a Sherlock Holmes, that the name 'Sherlock Holmes' as used in this story really refers to some man, Sherlock Holmes, and so on. It is therefore presumably part of the pretense of the story that the name 'Sherlock Holmes' is really a name and really has the ordinary semantic function of names. (KRIPKE, 2011, p. 58)

Si l'on s'arrête ici, d'un point de vue de la référence directe, il vient alors que les noms fictionnels feignent de faire référence :

> If this is so, the name, of course, doesn't really have any referent, it is pretended to have a referent; and if some view like Mill's is right, and the semantic function of naming is reference, then it follows that here we are only pretending to refer to a certain person and say something about him. (KRIPKE, 2011, p. 59)

Et dans ce cas, les propositions contenues dans les fictions ne sont pas de véritables propositions à propos de certaines personnes, mais seulement des propositions qui feignent d'en parler.

Kripke va alors défendre une conception ontologique réaliste des personnages en admettant que « Sherlock Holmes » réfère bien à une entité, sans pour autant céder aux intuitions meinongiennes ou au possibilisme : « fictional characters are not ghostly entities or merely possible entities—they are entities of a certain sort that exist in the real world. » (KRIPKE, 2011, p. 62) Mais comment assurer alors une référence à ce nom sans souscrire à un vaste domaine d'inexistants ?

À cette fin, Kripke va utiliser, en la modifiant, une idée de Keith Donnellan. Admettant l'universalité de la théorie de la référence rigide, Donnellan considère que, puisque l'on ne rencontrera jamais un individu concret en remontant la chaîne de communication des usages d'un nom fictionnel, on doit considérer que celle-ci est arrêtée par ce qu'il appelle un « bloc ». Qu'est-ce au juste ? Donnellan en donne des exemples : un acte d'imagination, une méprise d'identification, une œuvre de fiction (DONNELLAN, 1974, p. 23-24). Il précise qu'un bloc « ends with events that preclude any referent from being identified » (DONNELLAN, 1974, p. 23) ou encore que « the history of our use of a name [...] does not end in the right way » (DONNELLAN, 1974, p. 24). Ce qu'est exactement un bloc n'est donc pas complètement défini, mais stigmatisé par une anormalité : ne pas parvenir à aboutir à un référent lorsqu'on traque un nom sur la piste de la chaîne de retransmission du nom. Pour un individu qui recherche la référence d'un nom fictionnel, le « bloc » caractérise ainsi le fait de devoir s'arrêter sans parvenir à désigner une chose réelle.

Les noms fictionnels fonctionnent bien comme des noms propres, mais leur chaîne de communication sont interrompues. Il convient cependant d'expliquer ce qui caractérise alors la chaîne d'un nom comme « Sherlock Holmes » de celle d'un nom comme « Hercule Poirot ». Kripke reprend l'idée de Donnellan en substance, mais modifie la nature de l'échec de remontée de la chaîne causale. Car il y a bien un acte de baptême, même s'il ne relève pas d'une entité concrète, ni d'une entité meinongienne :

> So in this sense, instead of saying that the name "Hamlet" designates nothing, we say that it really does designate something, something that really exists in the real world, not in a Meinongian, shadowy land. In this sense a term in fiction refers to a fictional character. This is a usage parasitic on the usage where it is merely a fictional name of a person. We have, so to speak, a transformation which creates a new kind of object from a fictional name of a person, by transforming it into a name of a fictional person. This allows us to have all names in fiction really refer. (KRIPKE, 2015, p. 78)

Ce à quoi réfère « Sherlock Holmes », c'est à un acte inchoatif de création du personnage par son auteur. Ceci permet de distinguer le fonctionnement de différents noms propres. Le nom fictionnel a un référent qui n'est pas un bloc, mais un acte de création par un auteur. Kripke précise d'ailleurs la nature culturelle des entités fictionnelles :

> The question of their existence [of ficta] is a question about the actual world. It depends on whether certain works have actually been written, certain stories in fiction have actually been told. The fictional character can be regarded as an abstract entity which exists in virtue of the activities of human beings, in the same way that nations are abstract entities which exist in virtue of the activities of human beings and their interrelations. A nation exists if certain conditions are true about human beings and their relations; it may not be reducible to them because we cannot spell them out exactly (or, perhaps, without circularity). Similarly, a fictional character exists if human beings have done certain things, namely, created certain works of fiction and the characters in them. [...] [Fictional characters] are not ghostly possible entities; they are abstract entities of a certain sort that exist in virtue of the activities of people. (KRIPKE, 2011, p.62-64)

En outre, cela permet aussi d'expliquer les confusions entre personnage réel et fictionnel. Il est en effet toujours possible qu'un lecteur se trompe sur la nature du nom, et qu'il considère qu'un nom fictionnel désigne une chose réelle. La théorie des chaînes débouchant sur des actes de création permet d'expliquer la possibilité de telles erreurs :

> If one mistakenly believed the name to be non-empty rather than empty, it would be part of the mistake that this is a name having the ordinary semantic function of names. (KRIPKE, 2011, p. 58)

La question de départager entre noms propres et noms fictionnels relève donc d'enquêtes empiriques :

> A name of fictional character has a referent. One might then suppose that the name definitely has a referent (the fictional characters). It will be a matter of empirical investigation, concerning a given work, whether it is about a fictional character or a real person. (KRIPKE, 1963, p. 65)

Cette conception est reprise par Amie Thomasson, développée à partir de la théorie ingardienne des moments existentiels. Les *ficta* sont des *artefacts abstraits*. « Artefacts » parce qu'il sont des constructions, des créations d'auteurs et « abstraits » parce qu'ils n'ont pas d'étendue spatiale mais qu'ils ne sont pas pour autant réductibles à des contenus mentaux subjectifs.

> En tant que réaliste en ce qui concerne les personnages de fiction, je maintiens qu'il y a des personnages de fiction tels que Meursault, Hamlet et Precious Ramotwe. Cela ne consiste bien sûr pas à dire qu'il y ait (ou qu'il y ait jamais eu) de tels gens, pas plus ici que dans un 'monde fictionnel' modalement ou spatialement lointain. Les personnages de fiction, à mon avis, sont un type particulier d'artefacts abstraits, créés dans le processus de narration des œuvres de fiction, et pas très différents d'autres créations culturelles abstraites telles que les lois, les contrats, et même les histoires. (THOMASSON, 2010)

L'idée est de constater que les fictions peuvent être entendues de deux manières différentes : d'une part en tant de médias qui véhiculent l'information, c'est-à-dire en tant qu'artefacts, histoires, pièces faisant partie prenante du monde matériel et d'autre part comme médiums, c'est-à-dire

comme information. L'idée de Thomasson est donc de réifier les fictions, de considérer la partie informative de la fiction de la même manière que la partie médiatique.

L'existence des artefacts abstraits est dépendante de l'existence d'autres entités et d'états de choses, comme les intentions des auteurs et les textes matériels qui les diffusent. En ce sens, le mode d'existence des *ficta* les rapproche du mode d'existence des lois et des contrats. Leur forme d'existence abstraite ne relève donc pas d'une ontologie exotique puisqu'elle ne nécessite finalement que la seule condition de satisfaire à certains protocoles validant leur apparition :

> Given such a work of literature (one containing purely fictional names), nothing more is required, no extra ingredients are needed, to 'get' a fictional character, just as nothing is more required for there to be a pair of gloves than for there to be a matching left glove and right glove, and nothing is more required for there to be marriages that certain legal principles be accepted and their criteria fulfilled. In each case, the existence of the former entities (according to our ordinary understanding of terms like 'pair', 'mariage' and 'fictional character') is guaranted by the existence of the latter entities. (THOMASSON, 2005, p.221)

La référence du nom fictionnel est donc fondée sur l'association que ce nom entretient avec le texte où il est mentionné, et dont il dépend de manière ontologique, en tant qu'objet intentionnel. La dépendance que les *ficta* entretiennent aux textes leur procure une temporalité propre. Les *ficta* comme artefacts abstraits naissent à un moment donné et peuvent disparaître si tous leurs supports disparaissent également. En un mot, les artefacts abstraits semblent répondre à nos intuitions concernant le caractère temporel des personnages de fiction.

L'analyse de Thomasson n'est pas seulement conceptuelle, mais métaphysique. Thomasson reprend ainsi la perspective du pluralisme ontologique de Roman Ingarden en la simplifiant considérablement. Thomasson subdivise en effet le domaine de ce qui est en deux catégories relativement bien démarquées : d'un côté les objets qui sont purement matériels, logés de manière spatio-temporelle, et d'un autre côté, les objets généralement qualifiés d'abstraits comme les associations et les états qui n'ont pas d'étendue spatiale mais qui ont néanmoins une étendue temporelle.

> (La) capacité intrinsèque à certains états mentaux, d'aller au-delà d'eux-mêmes et de représenter quelque chose (...) permet de stratifier de nouvelles propriétés et de nouveaux objets à partir d'un monde naturel indépendant. (THOMASSON, 2011, p. 169)

Comme Ingarden, Thomasson souligne que la méthodologie de l'enquête ontologique doit suivre le chemin suivant : d'abord établir les relations de dépendances qui se tissent entre les putatives entités matérielles et mentales (tâche de l'ontologie), pour ensuite fixer les catégories ontologiques, quitte à en élimer ou à réduire certaines à d'autres (tâche de la métaphysique). Aussi, Thomasson soutient que la prétention du moniste, qu'il soit matérialiste ou idéaliste et de l'éliminativiste à respectivement récuser ou réduire certaines entités ne peut s'avérer pertinente qu'à la condition d'avoir préalablement établi les relations de dépendances qui relient états matériels et états mentaux :

> Le fait de clarifier les dépendances apparentes à l'égard des entités mentales ou spatio-temporelles, permet de clarifier aussi les coûts apparents liés à l'élimination des entités qui les soutiennent,

> et donc, d'indiquer à l'éliminativiste comme à l'idéaliste, quelles sont les autres entités auxquelles ils vont devoir renoncer ou qu'ils vont devoir expliquer autrement. (THOMASSON, 2011, p. 171)

Une catégorisation ontologique basée sur la relation de dépendance pourrait *a priori* paraître assez absurde en ce que traditionnellement, les essences ne relèvent pas des propriétés relationnelles. Mais la dépendance n'est pas une propriété extrinsèque comme « être à 2 m d'une fenêtre », il s'agit d'une propriété délivrant les conditions d'existence d'un objet qui n'est pas strictement matériel.

Si l'on cherche, comme le préconise Thomasson, à déterminer le genre de dépendance qu'occasionnent les *ficta*, nous pouvons relever d'une part une dépendance historique à un acte intentionnel de création par un auteur et d'autre part une dépendance ontologique à l'entité qu'est cet auteur. Une entité x dépend ontologiquement d'une entité y si x ne pourrait exister sans qu'existe aussi y. Autrement dit, l'inexistence de y impliquerait inévitablement l'inexistence de x. Mais cette notion de dépendance, très faible en ce qu'elle ne précise pas les conditions de synchronicité des entités x et y, peut encore être nuancée en plusieurs versions plus fortes. La dépendance ontologique constante est la plus forte de toutes : une chose x dépend constamment d'une chose y si x ne peut exister que de manière contemporaine à y, autrement dit, s'il est nécessaire qu'à tout moment de l'existence de x, y existe aussi. On dit qu'une entité x dépend historiquement d'une entité y si x nécessite l'existence de y pour apparaître à l'existence, mais peut ensuite subsister indépendamment de y. On dit d'une dépendance historique qu'elle est rigide si elle dépend d'un individu en particulier. On dira qu'elle est générique si elle ne dépend pas d'un individu en particulier.

Les *ficta* sont donc des entités ayant une double contrainte ontologique : d'une part, ils sont historiquement dépendants d'actes de création. En effet, les *ficta* viennent à l'existence par l'acte de création d'un auteur particulier à un certain moment. En outre, ils peuvent persister par delà cet acte de création. En tant qu'artefacts abstraits, ils doivent donc être considérés comme historiquement et rigidement dépendants des actes mentaux de leur auteur.

Mais la caractérisation de Thomasson ne peut s'arrêter là parce que si cette dépendance historique explique pourquoi les personnages naissent, en revanche, elle n'explique pas l'éventualité d'une disparition. Or, comme l'écrit Sartre dans l'Imaginaire, les objets purement intentionnels semblent souffrir d'une carrière extrêmement fluette et discontinue :

> La faible vie que nous leur insufflons vient de nous, de notre spontanéité. Si nous nous détournons d'eux, ils s'anéantissent. (SARTRE, 1940, p. 240)

Il semble donc que la dépendance ontologique qui les oblige ne soit pas seulement historique, sans quoi ils ne seraient rien, mais bien constante. Une fois que la pensée se détourne de l'objet fictionnel, celui-là disparaît. La vie sporadique de tels objets serait tout simplement inexplicable sans le recours d'une autre dépendance ontologique. Selon Thomasson, cela s'explique parce que les *ficta* sont, d'autre part, constamment et génériquement dépendants de leurs supports ou des pensées des récepteurs de la fiction. Cette dépendance ontologique est en effet générique puisqu'elle ne dépend pas de l'occurrence qui sert de support à l'œuvre (une réplique du livre pourra aussi bien

faire l'affaire qu'une autre) et elle est constante en ce qu'il doit au moins y avoir une occurrence servant de support au personnage pour assurer la persistance du *fictum* qui lui correspond.

En quoi la théorie des artefacts abstraits est-elle une théorie réaliste ? Un *fictum* est généré par un acte *purement* intentionnel de création, il vient à l'existence par cet acte. Or, comme nous l'avons déjà dit, d'après Brentano, un objet purement intentionnel est une entité immanente à l'acte de pensée qui le cible. Mais la théorie des artefacts abstraits, qui se veut réaliste et non conceptuelle, ne peut se contenter de cette existence fragile, immanente, des *ficta* dans les pensées. Les objets fictionnels sont considérés comme des choses qui persistent, dont la carrière est prolongée au delà des actes mentaux qui s'y réfèrent. Un artefact abstrait subsiste à son acte de création inchoatif, dont il est rigidement et historiquement dépendant, et est, pour ainsi dire, maintenu en vie par les supports médiatiques qui en assure la persistance continue dans le flux du temps par une dépendance ontologique générique continue.

Tout comme l'explication meinongienne des *ficta*, celle de Thomasson assigne aux pensées un étonnant pouvoir génératif. Elles sont capables de produire des entités subsistant par-delà l'acte de pensée, tant qu'une occurrence médiatique quelconque les supporte dans l'existence. Le fait que ces entités soient à la fois produites par des actes mentaux et prolongées dans l'existence par des supports pose d'évidents problèmes d'individuation, nous en reparlerons dans le chapitre suivant.

La thèse des artefact abstraits pose également d'autres problèmes. Si les *ficta* dépendent d'actes de création, par exemple l'écriture de textes où figurent leurs noms, ces textes doivent être considérés comme des *hapax*. Mais cette caractérisation devient problématique pour les noms comme « Ulysse » et « Richard III » qui dénomment respectivement un personnage mythique et un personnage historique préexistants aux personnages fictionnels d'Homère et de Shakespeare.

En outre, même dans le cas où le personnage n'aurait pas d'occurrence hors de l'œuvre, déterminer l'origine de l'acte de création d'un personnage demeure encore problématique. Car quelle version du manuscrit ou quelle édition publiée faudrait-il choisir ? Faut-il prendre en compte la première occurrence du nom quelque part dans l'avant-texte de l'écrivain ? Ou bien ne faut-il prendre en compte que la version publiée du texte ? Il semble que nous n'ayons que des moyens arbitraires de départager entre de telles options.

Mais cela n'est peut-être pas une objection très grave à l'égard de la théorie de Thomasson : peut-être qu'un peu de vague peut être admis sur le moment du baptême du personnage. Ce qui paraît plus néfaste à l'égard de la théorie des artefacts abstraits, c'est par ailleurs qu'un personnage de fiction n'apparaît pas comme un lapin hors d'un chapeau de magicien. Il est construit au fur et à mesure de l'œuvre de l'écrivain, et élaboré dans l'avant-texte. La théorie de Thomasson ne permet pas de rendre compte de telles difficultés parce qu'elle est trop ancrée sur un aspect réifiant et statique qui échoue à rendre compte des aspects dynamiques de l'identification des personnages qui évoluent dans le temps [18]. Nous y reviendrons quand nous évoquerons le problème de l'identité des *ficta* dans le chapitre qui suit.

Terminons enfin avec une dernière objection : Sainsbury a caricaturé la théorie artefactuelle par la croyance au syllogisme suivant : les histoires appartiennent au monde actuel et les personnages de fiction appartiennent aux histoires, donc les personnages de fiction appartiennent au monde actuel (R. SAINSBURY, 2010). En effet, la théorie artefactuelle ne présuppose-t-elle pas ce qu'elle

[18] Voir par exemple (ERDRICH, 2018).

tente justement de démontrer ? Que signifie pour un personnage de fiction d'être contenu dans une histoire ? N'est-ce pas une utilisation métaphorique de la relation d'inclusion ? Et qu'est-ce que cela signifie qu'une histoire « existe » ?

Dans les théories artefactuelles, les entités produites par les auteurs persistent en dehors des actes d'imagination. Ils ont une existence qui n'est pas seulement intentionnelle, liée à un acte de pensée à un moment t, mais qui se prolonge au delà de cet acte. C'est en cela qu'on peut dire que les théories artefactuelles sont des théories réalistes des entités fictives.

3.2.4 Le pan-fictionnalisme

Nous pourrions être tentés de dire que notre incapacité à parvenir à une théorie expliquant suffisamment nos représentations fictionnelles vient du fait que jusqu'à présent, nous avons admis le principe de mimésis, c'est-à-dire le principe que les énoncés doivent nécessairement trouver leur signification dans la réalité. Parmi les auteurs s'intéressant à la question des rapports de démarcation entre la fiction et le réel, certains ont, en conséquence, renoncé au principe mimétique. À coté du pan-réalisme meinongien, qui étend le principe de référence à tout ensemble de descriptions, nous devons aussi mentionner une solution contraire, que je qualifierai d'attitude pan-fictionaliste, et dont le principe consiste à supprimer tout simplement la possibilité d'un rapport discursif au réel. C'est une manière radicale de répondre à la question de la connexion entre réalité et représentations vraies, en supprimant complètement l'accès au réel. Le discours factuel est ainsi complètement rabattu sur le discours de fiction, les énoncés scientifiques étant alors, eux-mêmes, rapprochés des discours fictionnels. Le principe de mimésis est ainsi révoqué, puisqu'il n'y a plus de modèle de référence sur lequel appuyer les représentations.

De fait, l'accès à la vérité peut être refusé de différentes manières. Soit qu'on supprime l'idée de réel, comme semble l'entreprendre par exemple Gorgias de Léontinium. Soit qu'on conserve l'idée de réel, mais qu'on nie tout rapport du langage à ce dernier. De ce point de vue, la conception de l'atome est fausse, elle ne relève que d'une interprétation pour parler du réel, mais n'est qu'une duplique de celui-ci. C'est en gros la conception que prône Clément Rosset : rien ne peut être dit du réel, le réel échappe à toute représentation. Il y a, nous l'avons déjà évoqué, une inclusivité causale du réel. Et comme le souligne David Lewis, nous ne sommes jamais complètement séparés de l'étoile la plus lointaine, ni de l'événement le plus reculé dans le temps. Pourtant, il faut bien admettre certaines exclusions épistémiques, en ce sens que nous ne pouvons connaître qu'une petite partie de l'univers parce qu'il est impossible que la lumière émise depuis certaines régions provienne jusqu'à nous durant le temps de notre existence (individuelle, voire à l'échelle de l'humanité). En outre, et dans un autre contexte, nous pouvons estimer que ce que nous pensons est complètement incommensurable à ce que pensent d'autres personnes. En somme, il semble que la question d'une pluralité de points de vue sur le monde aboutisse inévitablement aux deux alternatives classiques du réalisme et de l'idéalisme. Soit le monde est unique et parler de mondes multiples s'avère être une métonymie plus ou moins pertinente, soit différents mondes cohabitent. Dans *Qu'est-ce qu'un monde ?*, Julien Rabachou considère cependant une troisième possibilité qu'il juge meilleure : qu'il n'existe qu'un seul monde, considéré comme totalité de ce qui existe, sans que cette totalité ne :

> parvienne à se constituer en *unité complète* et n'empêche d'autres totalités, d'autres mondes

> qu'elle contient, de venir s'articuler à elle, la concurrencer, ou du moins en contester l'exclusivité, voire l'unité. (RABACHOU, 2016, p. 20)

Pour appuyer son hypothèse, Rabachou cite les mondes physiques parallèles, les mondes possibles et les mondes construits qui entrent tous en concurrence avec le monde objectif. Mais ne risque-t-on pas ici de confondre l'acception ontologique et l'acception épistémique du terme de monde ? Peut-être ne s'agit-il finalement de ne rien dire de plus que ce qu'en a dit Kant, à savoir qu'il n'existe qu'un seul monde, mais qu'il y a une infinité de manières de le concevoir, autrement dit que nous avons bien accès à des représentations, mais jamais à la chose en soi.

Le nihilisme de Gorgias

Gorgias de Léontium est sans doute le premier partisan connu du pan-fictionnalisme. Selon lui, le langage ne possède pas de capacité référentielle car, et contre toute illusion, il n'y a rien. Ses principales thèses ont été rapportées par Sextus Empiricus, elles sont au nombre de trois :

> Premièrement et pour commencer, rien n'existe ; deuxièmement, même s'il existe quelque chose, l'homme ne peut l'appréhender ; troisièmement, même si on peut l'appréhender, on ne peut ni le formuler, ni l'expliquer aux autres [19]. (COLLECTIF, 2001, p.1023)

La première thèse est ontologique. Elle consiste précisément en un radical nihilisme ontologique : il n'y a rien. L'argument avancé est que si quelque chose existait, ce quelque chose relèverait soit de l'être, soit du non-être, soit d'un mixte des deux. Mais chacune des trois possibilités est rejetée par Gorgias. Premièrement, l'être ne peut exister, car dans le cas contraire, il serait soit éternel, soit engendré, soit les deux à la fois. Or il ne peut être éternel, car s'il l'était, il n'aurait pas de commencement. Or ce qui n'a pas de commencement est illimité, et ce qui est illimité n'est nulle part, donc inexistant. D'un autre côté, si l'être était engendré, alors il serait engendré à partir de l'être ou du non-être. Ce qui n'est pas possible car une chose ne peut ni s'engendrer elle-même, ni être engendrée par le non-être. Enfin, l'être ne peut pas non plus être à la fois éternel et engendré, car cela serait contradictoire. Ainsi, conclut Gorgias, l'être n'existe pas. Ensuite, il faut rejeter l'existence d'une chose qui relèverait du non-être car sinon, à la fois le non-être existerait, par cette chose, et il n'existerait pas, par définition. Troisièmement, il ne peut non plus y avoir à la fois l'être et le non-être, car si les deux coexistaient, ils seraient identiques du point de vue de l'existence. Or s'ils étaient deux, ils ne sauraient être identiques, et s'ils étaient identiques, ils ne pourraient être deux. Pour Gorgias, cela montre que ni le non-être, ni l'être, ni les deux à la fois n'existent en une chose. Autrement dit, rien n'est.

La deuxième thèse défendue par Gorgias est épistémique. Quand bien même existerait-il quelque chose, cette chose serait, de toute façon, inconnaissable :

[19] Soulignons que la forme de l'argumentation revêt ce que Freud appelle « logique du chaudron » dans *Le trait d'esprit et sa relation à l'inconscient* : A a emprunté à B un chaudron de cuivre et après l'avoir rendu, B l'accuse parce que le chaudron présente désormais un trou qui le rend inutilisable. Voici la défense de A : « Premièrement je n'ai absolument pas emprunté de chaudron à B ; deuxièmement le chaudron avait déjà un trou lorsque je l'ai reçu de B ; troisièmement je lui ai rendu le chaudron intact. » Comme le constate Freud, « Chacune des objections prise séparément est bonne pour elle-même mais, mises ensemble, elles s'excluent mutuellement. [...] On pourrait dire également : A met un 'et' à l'endroit où seul est possible un 'ou bien... ou bien' » (FREUD, 1992, p. 361).

> Que les pensées n'aient pas l'être pour objet, c'est évident : en effet, si les pensées ont l'être pour objet, tout ce qui est pensé existe, de quelque manière qu'on le pense. Ce qui est aberrant. En effet, de ce qu'on pense à un homme volant ou à un char qui court sur les flots, il n'en résulte pas pour autant qu'un homme vole ou qu'un char coure sur les flots. Par suite, les objets de nos pensées ne sont pas des êtres. En outre, si les objets de nos pensées sont des êtres, les non-êtres ne pourront pas être pensés : les contradictoires ont des attributs contradictoires et l'être et le non-être sont contradictoires. Par suite, et sans restriction, si l'être a l'attribut de pouvoir être pensé, le non-être aura celui de ne pouvoir être pensé. Or cela est absurde : Scylla, la Chimère et beaucoup de non-êtres sont pensés. Donc ce n'est pas l'être qui est objet de nos pensées. (COLLECTIF, 2001, p. 1025)

Puisque la coïncidence des pensées et des choses est impossible, c'est que les pensées ne portent pas réellement sur les choses. Sinon, les choses seraient par la pensée elle-même et celle-ci donnerait existence objective aux contenus de pensées, notamment aux choses qui n'existent pas actuellement. Par la même occasion, les choses qui n'existent pas ne pourraient pas être pensées.

Enfin, la troisième thèse formulée par Gorgias est sémantique. Elle affirme que non seulement le langage est incapable de porter sur les choses mais plus encore qu'il est incapable de communiquer à propos de ces choses. Le monde est ineffable. Somme toute, si rien n'est, alors le langage se suffit à lui-même. Il est, par nature, autotélique, puisqu'il ne réfère à rien. Son usage n'a pas pour effet de parler des choses pour ce qu'elles sont. Ainsi, les discours fictionnels ne peuvent être distingués des énoncés factuels car les mots contenus dans les putatifs énoncés factuels ne réfèrent pas plus que les mots contenus dans les énoncés fictionnels. Aussi, aucune différence sémantique ne sépare les deux types de discours. Cette évacuation radicale de la référence a d'ailleurs fait dire à Sextus Empiricus que Gorgias appartient à la catégorie des philosophes qui ont supprimé de leurs principes la notion de critère de vérité.

Si le langage est incapable de parler des choses qui, de toute façon, n'existent pas, il est en revanche capable de produire des effets. Il médiatise des actions par des dits. Comme le précise Philostrate, Gorgias est à la rhétorique ce qu'Eschyle est à la tragédie, car il a introduit l'art de la poétique à l'intérieur du discours philosophique. Diodore de Sicile cite par exemple les « antithèses, balancements, parisoses, rimes et autres procédés » (COLLECTIF, 2001, p. 1013) abondants dans ses discours. Ainsi, les arguments philosophiques de Gorgias empruntent les mêmes formes que celles de la *poiésis*, qu'Aristote décrit comme la production d'une action ayant une finalité qui lui est extérieure et qui s'oppose en cela à la *praxis* dont la finalité repose dans l'activité elle-même. Le langage permet d'agir indirectement sur les choses par l'exercice de la persuasion. La fonction principale du langage n'est, selon Gorgias, ni de représenter le monde, ni de dire le vrai. Il s'agit seulement de produire des effets externes. La véritable vocation du langage est donc pragmatique, idéologique et politique. Il s'agit d'inciter les hommes à produire l'action qui est attendue d'eux. Ce que Gorgias anticipe ici est donc en quelque sorte ce qu'Austin appellera, deux millénaires plus tard, l'aspect *performatif* du langage. Pour Gorgias, c'est donc le champ complet du langage qui est investi par la fiction.

On peut bien sûr critiquer cette réduction drastique du langage à une fonction purement performative, parce que si le langage n'avait d'autre usage que la production de certaines actions, il faudrait encore qu'il existât une possibilité de distinguer entre ces actions, entre celles attendues

sans être réalisées et celles qui adviennent. On voit que c'est là qu'achoppe la théorie de Gorgias. Car on peut difficilement prétendre à la fois que rien n'existe et que le langage ait une fonction de sélection entre l'actuel et le possible. Aussi, par leur radicalité, les thèses de Gorgias paraissent intenables.

La non-démarcation dans la théorie littéraire

Cette approche n'a cependant pas découragé certains auteurs, en particuliers ceux qui furent influencés par le structuralisme linguistique. Dans les années 70, émergea la thèse que les discours de fiction s'avèrent formellement indiscernables des discours factuels. Cette idée, soutenue par le principe kantien d'impossibilité de connaître la chose en soi, mena certains essayistes à affirmer l'inexistence des faits, en évacuant ainsi le réel par la bonde linguistique. Dans les travaux séminaux de Ferdinand de Saussure apparaissait déjà la thèse que le langage du pur sens doit être indépendant du langage de la pure référence pour qu'un nom n'ayant pas de référence puisse quand même signifier. Saussure rejette l'approche sémantique basée sur l'opposition externe fait/langage au profit d'une perspective purement linguistique autour de l'opposition interne signifiant/signifié. Le signifiant, forme linguistique de l'expression, est attribué de manière conventionnelle à un signifié (le sens). Ainsi, le sens est complètement défini par les structures formelles de l'expression, il est cloisonné dans son propre domaine d'usage. Cela a pour conséquence de rendre le sens complètement indépendant de la référence. Libéré de l'emprise de la référence extra-linguistique, le langage est alors célébré comme un système auto-référentiel. Les auteurs qui ont souscrit à cette thèse, et dont le collectif a été désigné du nom de « théorie littéraire », ont ainsi posé un « moratorium sur les questions de la référence » (PAVEL, 1988, p. 10).

Selon la théorie littéraire, la littérature reposerait donc non pas sur un principe mimétique de reproduction du réel, mais sur un principe sémiotique de signification par auto-référence. Toute l'importance consacrée à la littérature, et en particulier à la fiction, est ainsi déportée sur le signe. La notion de référence est remplacée par celle d'intertextualité, autrement dit, par les relations entretenues entre les textes.

La forme la plus radicale de la théorie littéraire soutient l'idée que la réalité n'est finalement qu'une pure et simple convention. Loin de réserver la thèse d'auto-référence sémiologique à la seule critique littéraire, ces théories ont progressivement étendu cette prétention aux discours prétendument factuels, la réalité étant ainsi ramenée à un simple « effet de réel » (Roland Barthes), voire, elles ont complètement dissipé la réalité dans ce que Jean Baudrillard appelle « le bruit des signes ». C'est par exemple ce que prétend Philippe Sollers qui écrit que le réalisme est « ce préjugé qui consiste à croire qu'une écriture doit exprimer quelque chose qui ne serait pas donné dans cette écriture, quelque chose sur quoi l'unanimité pourrait être réalisée immédiatement. Mais il faut bien voir que cet accord ne peut porter que sur des conventions préalables, la notion de réalité étant elle-même une convention et un conformisme, une sorte de contrat tacite passé entre l'individu et son propre groupe social » (SOLLERS, 1968, p. 236) On trouve aussi chez Barthes la même allégeance au refus de référentialité :

> La fonction du récit n'est pas de « représenter », elle est de constituer un spectacle qui nous reste encore très énigmatique, mais qui ne saurait être d'ordre « mimétique » [...] « ce qui se passe » dans le récit n'est, du point de vue référentiel (réel), à la lettre, rien ; « ce qui arrive »

> c'est le langage tout seul, l'aventure du langage dont la venue ne cesse jamais d'être fêtée. (BARTHES, 1985, p. 206)

Le langage n'imiterait donc jamais le réel mais serait comme retourné sur lui-même en une sorte de solipsisme sémantique : « le fait n'a jamais qu'une existence linguistique » (BARTHES, 2000).

Le point de départ de Barthes consiste à évaluer les points de divergence, s'ils existent, entre les récits historiques et les récits fictionnels. La méthodologie employée est la linguistique d'obédience structurale. Ainsi, dès les premières lignes du *discours de l'histoire*, Barthes confère à la sémiologie la tâche d'occuper la place de la rhétorique. Affairée à déterminer des invariants sémantiques, la sémiologie de Barthes indique la nécessité de réviser la putative objectivité des faits historiques en la rapprochant du discours de fiction.

> La narration des événements passés, soumise communément, dans notre culture, depuis les Grecs, à la sanction de la « science » historique, placée sous la caution impérieuse du « réel », justifiée par des principes d'exposition « rationnelle », cette narration diffère-t-elle vraiment, par quelque trait spécifique, par une pertinence indubitable, de la narration imaginaire, telle qu'on peut la trouver dans l'épopée, le roman, le drame ? (BARTHES, 2000, p. 153)

Barthes analyse ainsi les structures qui sont partagées entre récits historiques et récits de fiction. Il admet que le discours de l'historien procède par *shifters*, dont le concept est emprunté à Jakobson (JACKOBSON, 1963). Il s'agit de termes qu'on pourrait qualifier d'embrayeurs en ce qu'ils indiquent un passage de l'énoncé à l'énonciation. Il y a deux types de *shifters*. Les premiers sont dits testimoniaux quand le discours prétend rapporter à la fois l'acte de l'informateur et la parole de celui qui s'y réfère. Ils consistent en des marqueurs textuels comme « d'après mes sources », « un tel rapporte que », etc.

Le deuxième type de *schifters* relève de l'organisation du discours. Il s'agit des signes déclarés par lesquels l'historien structure son récit en le jalonnant de repères. Ce type d'embrayeur permet à l'auteur d'un récit de dédoubler la causalité temporelle des événements par une causalité discursive de l'ordre dans lequel ils sont racontés, « de compliquer le temps chronique de l'histoire en l'affrontant à un autre temps, qui est celui du discours lui-même et que l'on pourrait appeler par raccourci le temps-papier. » En effet, dans la nécessité de rapporter des faits qu'il connaît d'avance, l'historien peut agencer librement le fil de son récit. Autrement dit, celui qui discourt sur l'Histoire le fait nécessairement en racontant une histoire :

> C'est dans la mesure ou il sait ce qui n'a pas été encore raconté, que l'historien, tel l'agent du mythe, a besoin de doubler le dévidement chronique des événements par des références au temps propre de sa parole. (BARTHES, 2000, p. 68)

Encouragé par son propre mouvement, Barthes conclut que le discours historique ne suit pas vraiment l'ordre de causalité du réel, se limitant seulement à le signifier, « ne cessant de répéter c'est arrivé, sans que cette assertion puisse être jamais autre chose que l'envers signifié de toute la narration historique » (BARTHES, 2000). L'imaginaire de l'historien est ensuite rapproché de l'imaginaire psychotique et en ce que, dans l'un et dans l'autre, « personne n'est là pour assumer l'énoncé ».

Barthes constate ainsi la présence des embrayeurs à la fois dans les récits historiques et dans les récits fictionnels. Il en conclut alors qu'un évident parallélisme existe entre les deux types de discours. Plus encore, qu'ils sont mêmes *équivalents*. Le fait que ces marques rhétoriques soient également présentes dans les deux types de récits suffit donc à Barthes pour établir une équivalence générale entre eux. Mais de telles concordances structurales suffisent-elles pour conclure cette putative équivalence ? Une telle manière de voir les choses repose sur l'idée qu'une différence, si elle existait, entre la fiction et le factuel relèverait nécessairement d'une différence syntaxique, et qu'en l'absence d'une telle différence, il faudrait donc admettre l'identité entre faits et fiction. Les divergences entre faits et fictions s'évanouiraient par leur convergence structurale. De là, à en croire Barthes, tout discours historique ne serait finalement qu'un type particulier de discours fictionnel.

L'histoire comme description d'événements est soumise au soupçon de l'idéologie :

> Comme on le voit, par sa structure même et sans qu'il soit besoin de faire appel à la substance du contenu, le discours historique est essentiellement élaboration idéologique, ou pour être plus précis, imaginaire, s'il est vrai que l'imaginaire est le langage par lequel l'énonçant d'un discours (entité purement linguistique) « remplit » le sujet de l'énonciation (entité psychologique ou idéologique). On comprend dès lors que la notion de « fait » historique ait souvent, ici et là, suscité une certaine méfiance. [...] le fait n'a jamais qu'une existence linguistique (comme terme d'un discours), et cependant tout se passe comme si cette existence n'était que la « copie » pure et simple d'une autre existence, située dans un champ extra-structural, « le réel ». (BARTHES, 2000)

Le discours historique prétend à une forme d'énonciation performative, mais celle-ci est truquée, puisqu'en réalité, les constats établis ne réfèrent pas à des faits, ils ne sont que les signifiants d'actes autoritaires de parole.

Dans leur essence même, tous les discours relèveraient de fictions, et toute narration relèverait ainsi à une mythologie des événements. La théorie littéraire renoue ainsi avec l'aversion de Platon pour le mimétisme, mais en le conjurant :

> Le réalisme (bien mal nommé, en tout cas souvent mal interprété) consiste, non à copier le réel, mais à copier une copie (peinte) du réel [...]. C'est pourquoi le réalisme ne peut être dit « copieur » mais plutôt « pasticheur » (par une *mimesis* seconde, il copie ce qui est déjà copie (BARTHES, 1970, p. 61)

Comme Platon, Barthes soutient que ce n'est jamais le réel que l'œuvre artistique reproduit. Elle reproduit ce qui est déjà une autre reproduction, un autre texte. En ce sens, il y a bien une approche mimétique chez Barthes, mais elle doit être prise dans un sens non référentiel, et elle peut être qualifiée, selon l'expression de Julia Kristeva, d'« intertextualisme ». Ce n'est pas le réel, mais le texte qui est donné d'emblée, déjà là, avec ses présupposés idéologiques :

> L'artiste réaliste ne place nullement la « réalité » à l'origine de son discours, mais seulement et toujours, si loin qu'on puisse remonter, un réel déjà écrit, un code prospectif, le long duquel on ne saisit jamais, à perte de vue, qu'une enfilade de copies. (BARTHES, 1970, p. 173)

Finalement, tout cela conduit Barthes à une mise en question du réel lui-même. Car si aucun discours ne peut témoigner de la réalité, si les faits ne sont pas inscriptibles dans l'ordre du discours, c'est que la réalité relève d'une sorte de consécration religieuse illégitime. Le réel n'engagerait finalement qu'un acte de foi, et non un rapport paradigmatique entretenu par la référence des signes. Ce que la rhétorique ancienne nommait encore « réel » est finalement pour Barthes uniquement une illusion référentielle qui se pratique au sein des discours tant fictionnels que réputés factuels.

Si on admet cette forme épistémique d'anti-réalisme dont la thèse est qu'aucun discours ne réfère à autre chose qu'aux signifiants eux-mêmes, la question n'est donc plus de déterminer comment la fiction parvient à reproduire le réel, mais plutôt, comment la fiction parvient à nous faire croire qu'elle imite la réalité, alors que ce n'est pas le cas. D'où vient cet « effet de réel » dont parle Barthes ? Si le langage n'est plus du tout ancré sur le réel, d'où provient l'illusion référentielle qu'il produit ?

En assimilant les marqueurs textuels à des marques idéologiques, Barthes argumente encore l'identification des faits aux fictions selon le raisonnement suivant :

(1)	Le récit historique procède par significations.	hypothèse
(2)	Les significations sont porteuses d'idéologie.	hypothèse
(3)	L'idéologie a rapport à l'imaginaire.	hypothèse
(4)	L'imaginaire équivaut à la fiction.	hypothèse
(5)	le récit historique relève donc de la fiction.	conclusion

Cependant, force est de constater que Barthes n'a cependant étayé qu'une seule de ses hypothèses, à savoir, l'inclusion de l'idéologie dans l'imaginaire. Mais l'argument général reste peu convaincant, notamment parce que l'hypothèse (3) qui rend équivalents imaginaire et idéologie est très contestable, et semble indiquer une confusion entre l'irréel et l'illusoire.

De son côté, Michel Riffaterre a analysé les rapports entre discours poétique et discours factuel. On retrouve, comme chez Barthes, la même mise en question de la référence du langage au réel. Le langage poétique diffère certes du langage commun : « notre instinct nous avertit qu'un poème dit une chose et veut dire autre chose » (RIFFATERRE, 1978). C'est que la poésie n'exprime pas *directement* les idées et les choses, elle l'exprime de manière différée et transformée. Se proposant de déterminer en quoi consiste exactement cette différence, Riffaterre observe que le langage poétique produit une « illusion référentielle » : si les mots, dans leur usage commun, dénotent des choses, ce n'est pas le cas lorsqu'ils sont imbriqués dans un texte fictionnel. Ainsi, contrairement à ce qu'il paraît, les mêmes mots ne signifient pas la même chose lorsqu'ils sont utilisés dans un récit factuel ou dans un récit poétique. L'unité de signification des énoncés poétiques n'est pas le mot. Le sens d'un texte poétique est comme différé, il vient en second lieu. C'est le texte dans son intégralité qui devient, pour Riffaterre marqué de l'effet du trope : le sens des mots est ajourné puisque pour chacun des mots est mis à la place d'un autre. Cette métonymie généralisée se propage à l'intégralité du texte poétique.

En fait, Riffaterre distingue, dans le contexte poétique, trois effets qui agissent sur le sens des mots : le déplacement, l'altération et la création. Le déplacement du sens est celui produit par la

métaphore, un mot exprimant le sens d'un autre mot. L'altération manifeste une coïncidence de sens ambiguë et contradictoire. Enfin, la création désigne les sens particuliers produits dans le texte poétique, sens qui n'apparaissent jamais dans le langage factuel. Le texte poétique n'entretient donc qu'un rapport atermoyé avec la réalité. Privé d'un accès au réel, le texte poétique pourra donc voir multiplier ses interprétations.

Riffaterre concède que l'illusion référentielle est indispensable à la pratique du langage. La signification d'un texte peut avoir trois origines : il peut être situé dans l'intention de l'auteur, dans le texte lui-même ou dans l'intention du lecteur. Puisque la théorie littéraire rejette l'intention auctoriale, il ne reste que les deux dernières hypothèses. Or, pour Riffaterre, le texte ne détient pas, de lui-même, une signification. Le sens n'est pas dans les mots mais dans les interprétations des lecteurs :

> les critiques (...) mettent la référentialité dans le texte quand elle est en fait dans le lecteur, dans l'œil de celui qui regarde - quand elle n'est que la rationalisation du texte opérée par le lecteur. (RIFFATERRE, 1982, p. 93)

La croyance que les mots désignent directement les choses repose sur ce que Riffaterre appelle *la mythologie du signifié*, à savoir que les mots n'ont qu'une relation de convention avec leurs référents, arbitrairement noués avec des concepts dont la totalité représente le mythe du réel. Le signifié, qui exprime cette relation entre les mots et des regroupements de concepts, relève donc de la *doxa*, de la croyance sociale, mais non d'une objectivité impartiale. Pour des raisons pratiques, l'usage commun du langage oublie cette instance médiane, en faisant croire au caractère directement référentiel du langage.

> Tout comme l'illusion intentionnelle substitue à tort l'auteur au texte, l'illusion référentielle substitue à tort la réalité à sa représentation, et a tendance à substituer la représentation à l'interprétation que nous sommes censés en faire. (RIFFATERRE, 1982, p. 93)

Les mots sont tous déjà clivés par une instance régularisatrice, doctrinale, idéologique. Le signifié est donc une forme de totalitarisme idéologique qu'il convient de déconstruire. Pour cela, il faut admettre un usage du langage qui ne prétend pas référer, qui rend oblique sa signification. Riffaterre compense cette perte de signification par ce qu'il appelle la *signifiance*.

> Dans le langage quotidien, les mots semblent reliés verticalement, chacun à la réalité qu'il prétend représenter, chacun collé à la réalité qu'il prétend représenter, chacun collé sur son contenu comme une étiquette sur un bocal, formant chacun une unité sémantique distincte. Mais en littérature, l'unité de signification, c'est le texte lui-même. (RIFFATERRE, 1982, p.94)

Comme on le voit, en déplaçant l'unité de sens du mot à l'intégralité du texte, la théorie littéraire admet une prolifération des significations des textes : le langage n'est plus ancré à la réalité, les relations qu'il entretient sont purement réflexives. La signification ne relève pas d'un rapport entre les mots et les choses, mais seulement à un rapport entre les mots et eux-mêmes. En confinant ainsi la signification au langage tout en renonçant aux métalangages, la théorie littéraire renonce à l'objectivité, et mène à une prolifération de doubles sens. La volonté tant rabattue par les

tenants de cette théorie de fustiger, à raison, la croyance que *les mots sont les choses* a engendré l'idée d'une héméneutique infinie.

Le texte est ainsi ouvert à une infinité d'interprétations. Un texte aura autant de significations qu'il aura de lecteurs. Cela, on peut le concéder, si on admet toutefois qu'un texte ne dit pas n'importe quoi. Or, c'est bien le problème qui surgit de la théorie littéraire. L'échappée de signification, le fait qu'un mot en exprime un autre et non une chose du monde, induit que toute interprétation du texte demeure envisageable : qu'est-ce qui empêche alors d'interpréter la métamorphose de Grégoire Samsa comme une anticipation de la guerre civile espagnole ? L'herméneutique peut alors devenir prétexte à défendre n'importe quelle interprétation.

Les hypothèses constructionnistes d'une pluralité de mondes

Parallèlement au principe de non-référentialité défendu par la théorie littéraire, certains historiens et sociologues des sciences ne se contentent pas de considérer que des mondes épistémiques cohabitent ensemble ou sont inclus les uns dans les autres. Ils relativisent la notion de monde à des conceptions incommensurables. Selon eux, même si Ptolémée et Copernic exerçaient tous deux l'astronomie, et donc, partageaient le même *milieu*, le monde de Ptolémée est néanmoins radicalement différent du monde de Copernic.

D'après Thomas Kuhn, les modèles élaborés par la science sont des paradigmes qui vivent et meurent. Un paradigme qui émerge est érigé au rang de science normale quand il promet d'expliquer des observations jugées problématiques par les spécialistes du domaine concerné de recherches. Lorsque le paradigme est accepté, le fonctionnement pratique de la recherche entre alors dans la phase dite de « science normale », phase complètement consacrée à la réalisation des promesses du paradigme. Kuhn la décrit comme une période, généralement longue, de nettoyage destiné à rendre plus précis et plus convainquant encore le paradigme. Pour cela, les spécialistes tentent de concilier les observations aux prédictions en ajustant davantage le paradigme. Soulignons que durant la phase de science normale, un unique paradigme scientifique est admis par les spécialistes. Ainsi, Kuhn souligne que le mot de paradigme, qui doit être associé à l'idée de schéma et de modèle, n'a pas complètement la signification qui lui est généralement attribuée dans d'autres contextes. En particulier, il n'a pas la caractéristique de répétabilité qui est, ailleurs, associée à son acception :

> Dans une science, [...] un paradigme est rarement susceptible d'être reproduit : comme une décision judiciaire admise dans le droit commun, c'est un objet destiné à être précisé dans des conditions nouvelles ou plus strictes. (KUHN, 1983, p. 45)

Au cours de la phase normale, certaines observations nouvelles, ou certaines manières de considérer d'anciennes observations entrent en concurrence avec le paradigme. Cela arrive lorsque de telles observations échappent aux prédictions et aux explications, et que le groupe de spécialistes leur reconnaît une importance suffisamment aiguë pour ne pas les négliger. La science normale entre alors dans une phase de crise. Durant cette crise peuvent émerger d'autres paradigmes concurrentiels. Si l'un d'eux s'avère promettre des explications aux phénomènes qui ne sont plus supportés par le paradigme en cours, il suit alors une phase de révolution scientifique au bout de laquelle le nouveau paradigme succède à l'ancien. Le cycle de la science normale peut alors se

reproduire jusqu'à une nouvelle crise, et ainsi de suite. Comme on le voit, dans cette conception du monde, les paradigmes érigés au rang de science normale ne peuvent pas cohabiter, car ils ne sont pas compatibles les uns avec les autres. Ils sont successifs dans le temps. Bien sûr, on pourrait objecter par exemple que dans l'Antiquité, différentes écoles cohabitaient en proposant des modèles cosmogoniques différents, comme ceux de Thalès, de Parménide, d'Héraclite, etc. Mais selon Kuhn, de tels modèles ne relèvent pas réellement de paradigmes, quel que soit leur degré de pertinence par ailleurs, parce que dans son fonctionnement, la recherche n'était pas encore parvenue au stade de maturité scientifique.

La manière de concevoir ce qu'est un monde relève ici d'une sorte de modélisation holistique qui exclut certains faits ne trouvant pas d'explication dans le modèle considéré. On peut à nouveau rappeler en exemple le modèle géocentriste qui servit longtemps de représentation du monde, mais qui fut ensuite détrôné par le modèle héliocentrique. Les mondes de Ptolémée et de Copernic n'appartiennent pas à un certain *milieu scientifique*, ils sont des mondes radicalement différents et surtout *successifs*.

La question est essentiellement de savoir ce qui subsiste au delà des conceptions incompatibles du monde. De son point de vue de critique de la littérature, Pierre Bayard fait remarquer que la conception de l'incommensurabilité a également un impact sur nos interprétations des œuvres :

> Dès lors [...] que chacun d'entre nous perçoit des échos de certains univers parallèles - ou, si l'on préfère, vit pour une large part dans d'autres mondes -, il est logique de penser que l'incommunicabilité qui sépare les univers parallèles se réfléchit dans le nôtre et y perturbe les communications internes. Comment penser que pourraient dialoguer pleinement des êtres dont la vie psychique et intellectuelle se perd pour une part dans des mondes alternatifs ? On peut supposer que si les univers sont séparés par des différences de paradigmes, l'univers dans lequel vit chacun de nous est marqué par un *paradigme intérieur* qui rend difficile, voire impossible, la communication avec les autres. Cette notion de paradigme intérieur, qui complète la notion de paradigme proposée par Kuhn, permet de mieux comprendre les différences de lecture et d'interprétation que suscite une même œuvre. (BAYARD, 2014, pp. 135–136)

De même, dans son essai *Les Grecs ont-ils cru à leurs mythes ?*, Paul Veyne défend l'idée qu'il n'y a pas de véritable démarcation entre le monde du mythe et le monde réel. Dire d'un monde qu'il est « le monde réel », c'est seulement dire qu'il fait l'objet d'une croyance particulière :

> Un monde ne saurait être fictif par lui-même, mais seulement selon que l'on y croit ou pas ; entre une réalité et une fiction, la différence n'est pas objective, n'est pas dans la chose même, mais elle est en nous, selon que subjectivement nous y voyons ou non une fiction. [...] Einstein est vrai à nos yeux, en un certain programme de vérité, celui de la physique déductive et quantifiée ; mais si nous croyons à *L'Iliade*, elle sera non moins vraie, en son programme de vérité mythique. (VEYNE, 1983, p. 33)

La médiatisation du réel par une pluralité de mondes peut ainsi dégénérer en une renonciation pure et simple à l'idée de réalité. Le réel ne serait finalement qu'un monde fictionnel parmi d'autres. C'est aussi l'idée que défend Nelson Goodman, selon qui les mondes relèvent de constructions sociales. Mais plus encore, Goodman affirme que ces constructions ne sont ancrées sur aucune réalité objective, et rejette complètement la théorie mimétique.

La position que Goodman prétend occuper dans la discussion sur l'unicité du monde est à ce titre paradigmatique. Goodman affirme qu'il défend une thèse tierce entre réalisme et idéalisme : « je suis un anti-réaliste et un anti-idéaliste – donc un irréaliste » proclame-t-il dans la préface de *Manière de faire des mondes*. Dans *La Structure des Apparences*, il écrit :

> Mon irréalisme rejette l'idéalisme tout aussi certainement qu'il rejette le réalisme. Selon l'usage que je fais de ces termes, « réalisme » s'oppose à tous les systèmes idéalistes, « idéalisme » s'oppose à tous les systèmes réalistes, alors qu'« irréalisme » ne fait ni l'un ni l'autre. (GOODMAN, 2017, p. 204)

La théorie des mondes multiples de Goodman est fondée sur le rejet que la représentation serait une forme de relation de ressemblance. La ressemblance n'a pas de prise sur un monde, réel, qui se tiendrait au-delà des représentations, parce que, dit Goodman, un tel monde n'existe tout simplement pas. De fait, chez Goodman, l'interdépendance entre le monde et le langage est radicalisée, prenant la forme d'un constructionnisme de mondes, ce que Goodman appelle les *versions* de mondes. Ce n'est pas tant, comme le voudrait un relativiste, que la référence au monde ne peut se faire que dans un certain cadre épistémique. L'idée est plutôt de dire que c'est le cadre qui *constitue* le monde. Au delà du cadre, il n'y a rien à voir. En particulier, il n'y a pas de réalité *ready-made* sur laquelle serait ancré le cadre. Les mondes ne sont pas donnés tels quels :

> Si j'insiste pour que vous me racontiez comment est le monde indépendamment de tout cadre, que pourriez-vous dire alors ? Quoi qu'on ait à décrire, on est limité par les manières de décrire. À proprement parler, notre univers consiste en ces manières plutôt qu'en un monde ou des mondes. (GOODMAN, 2017, p. 11)

Il est certes possible de parler pour ne rien dire, autrement dit avoir « des mots sans mondes » mais en revanche, un monde n'existe que dans et par le langage, il n'y a « pas de mondes sans mots ou d'autres symboles » (GOODMAN, 2017, p. 15). Contre Wittgenstein, Goodman soutient que les limites du monde ne sont donc pas seulement celles de la logique. Elles sont essentiellement celles du langage. Différents sont les langages, différents sont les mondes.

En outre, pas plus qu'il n'y a de monde *ready-made*, il n'y a de sens à distinguer entre des mondes fictifs et des mondes réels. La démarcation est tenue pour nulle et non avenue :

> Il n'y a pas de mondes fictifs. L'écrivain qui en éprouverait du chagrin pourra ou ne pas pourra pas se consoler si j'ajoute qu'il n'existe pas davantage quelque chose comme le monde réel. Puisqu'il y a des vérités rivales, il y a de nombreux mondes, à supposer qu'il y en ait, mais il n'y a rien de tel que le monde. (GOODMAN, 2009, p. 35)

Il n'y a pas de mondes proprement fictifs parce que « la fiction, aussi fausse ou extraordinaire soit-elle, se rapporte à ce qui est réel, à supposer qu'elle se rapporte à quelque chose. Quoique les fictions soient toutes des contre-vérités littérales, ce qui, soit dit en passant, les distinguent des récits factuels, certaines, poursuit-il, sont néanmoins *métaphoriquement vraies* » (GOODMAN, 2009, p. 32).

La sémantique de Goodman admet deux types d'extensions : primaire et secondaire. La première consiste au sens usuel d'extension, à savoir, l'énumération des particuliers concrets qui remplissent la définition d'un concept. La deuxième correspond à ce que Goodman désigne comme

étant la vérité *métaphorique*, qui concerne particulièrement les termes dont l'extension primaire est vide, comme les noms fictionnels. Par exemple, Julien Sorel n'a pas d'extension primaire, mais il a cependant une extension secondaire : il représente la figure littéraire d'un ambitieux. La lecture de Stendhal peut ainsi élargir nos conceptions en délinéant finement un profil psychologique qui pourrait nous échapper dans d'autres circonstances. Le degré de réalisme qu'on accorde à une fiction ne réside pas dans la manière dont est racontée l'histoire mais dans le contenu de cette histoire. Considéré de manière littérale, le texte de Don Quichotte ne dénote rien. En revanche, dans son sens métaphorique, il « décrit un grand nombre d'entre nous qui se battent contre des moulins à vent (ou des moulins à paroles) [...] » (GOODMAN, 2009, p. 43).

Il y a ainsi, chez Goodman, une critique de l'usage esthétique du terme « réalisme », en ce qu'il occasionnerait un échelonnement des fictions selon le degré d'écart qu'elles entretiennent à la réalité. C'est une conséquence inévitable du rejet de la thèse de l'existence d'un monde réel prêt à être décrit. Déterminer le degré de vérité métaphorique exprimé dans une œuvre est une question difficile, reconnaît Goodman. Mais point n'est besoin d'y recourir pour constater que la vérité, littérale ou métaphorique, n'est en rien tributaire du putatif caractère « réaliste » de l'œuvre, car « un roman ou un tableau réaliste peut regorger d'erreurs, tandis qu'un tableau ou un roman fantastique peut être métaphoriquement vrai ou correct » (GOODMAN, 2009, p. 34). La question du réalisme n'est donc pas celle d'une similitude ou d'une proximité mondaine. Il s'agit d'une question concernant les modes de description, celle « de la familiarité des symboles utilisés dans le récit » et d'autre part, la question de la vérité est celle de « ce qui est raconté, que ce soit de manière littérale ou métaphorique, au moyen de symboles familiers ou fantastiques » (GOODMAN, 2009, p. 34). Le réalisme n'est donc pas une question de degré d'information préservée. Considérant l'exemple deux images représentant une même scène mais dont l'une la représenterait en perspective cavalière tandis que l'autre non, Goodman soutient que les deux images donneraient exactement la même information (GOODMAN, 2017, p. 144). Le réalisme ne relève donc pas d'un degré de similarité, mais d'un degré de conditionnement par certaines conventions établies de représentation.

Aussi, la question de l'usage de termes comme « impossible », « irréel » ou « fictionnel » ne relève pas de la philosophie mais, pour ainsi dire, d'une métaphilosophie :

> Le choix entre le monisme et le pluralisme tend à s'évanouir sous l'analyse. S'il n'existe qu'un seul monde, il doit embrasser une multiplicité d'aspects contrastés ; s'il y a plusieurs mondes, leur regroupement est un. Le monde unique peut être appréhendé comme multiple, ou les mondes pluriels comme un ; l'un et le multiple dépendent de la manière d'appréhender. (GOODMAN, 2017, p.)

De ce point de vue, l'apparente contradiction entre variétés de mondes et unicité du monde s'évanouirait à travers l'aspect pluriel des expériences que nous avons parfois du mal à intégrer en un tout unique qui pourtant les subsume. Il ne s'agirait au fond que d'une question d'échelle : les mondes construits sont comptés comme divers à l'échelle de représentations structurées, de systèmes indépendants de pensée mais aussi comme un monde unique à l'échelle supérieure de nos expériences. Bien sûr, un tel point de vue ne serait qu'un relativisme épistémique. Notre connaissance du seul monde qui existe dépend de nos manières plurielles de le concevoir.

Une telle manière de considérer le concept de monde enjoindrait une sorte de perspectivisme

affirmant que tout individu ou groupe d'individus est susceptible de constituer, par un langage propre, un point de vue sur le monde. Il n'y aurait pas lieu à discuter de la pertinence des représentations car chaque individu serait juge et partie de son monde. Un monde ne serait alors pas plus vrai qu'un autre. Goodman insiste d'ailleurs à de nombreuses reprises sur l'idée que le monde où la Terre est immobile au centre de l'univers est tout aussi vrai que le monde où la Terre se meut.

La construction des mondes est régie chez Goodman par cinq principes. Le premier est la composition/décomposition. Pour faire un monde, il faut d'un côté partitionner des genres en espèces, séparer des classes d'objets par des traits, établir des distinctions. D'un autre côté, il faut inversement remonter des membres aux classes, recombiner les parties en touts, établir des relations. C'est le travail des étiquettes linguistiques que sont noms, prédicats, images et gestes. Le monde de l'Esquimau, dit Goodman, diffère de celui de l'habitant de la Nouvelle Angleterre en ce que le concept général de neige demeure inconnu au premier et que la décomposition de ce concept, fragmenté en de multiples entités, est étrangère au second. Le deuxième principe de création de monde est la pondération. Les genres présents dans différents mondes n'y ont pas la même importance. L'accentuation explique par exemple, dit Goodman, les différences d'aspects entre des portraits de Daumier, Ingres, Michel-Ange ou Rouault. Les trois principes suivants sont l'agencement, la suppression/suplémentation et la déformation.

S'il n'y a pas un monde commun derrière les mondes construits, reste alors la question de la valeur de telles constructions. Sont-ils vrais, ou valables, ou corrects ? Y a-t-il des versions de monde qui seraient incorrectes ? Y a-t-il des contraintes pour faire une version correcte ou toute construction de monde peut-elle légitimement prétendre être correcte ? Selon Goodman, une version n'est correcte que si elle n'enfreint aucun de ses propres réquisits ni aucune croyance inébranlable.

Les ressources du langage étant larges, la conséquence d'un tel point de vue est que le Réel se voit multiplié en une pluralité de mondes : chaque monde relevant d'une construction de monde au moyen d'un langage particulier. Il n'y a pas de monde par delà le langage, pas de monde que le langage chercherait à exprimer : le monde n'est pas une donnée préalable qui serait exprimée par le langage, le monde est la construction d'un langage. Nous ne nous tenons donc pas devant le monde comme devant un tableau. Le monde n'est pas un point fixe archimédien, mais une condition d'inclusion dans le langage.

On peut objecter à la théorie des mondes construits de Goodman qu'il paraît extrêmement absurde de prétendre construire des mondes, au sens ontologique du terme, par le biais du langage. Qu'un langage et que des conventions de représentation soient nécessaires afin de parvenir à conceptualiser le monde est une chose admise. Mais que le monde n'existe pas en dehors de ces conventions, c'est difficile à croire. Israël Scheffler objecte à Goodman que sa caractérisation revient à affirmer qu'une étoile n'existe que par le langage, et que sans langage, elle n'existerait tout simplement pas. Goodman y répond ainsi :

> [...] Est-ce que des étoiles non-en-tant-qu'étoiles, des étoiles non-en-tant-que-mouvantes et non-en-tant-que-fixes, sont mouvantes ou non ? Sans une version, elles ne sont ni en mouvement, ni fixes. Et tout ce qui ne se meut ni n'est fixe, tout ce qui n'est ni comme-ceci, ni comme-cela, revient à n'être rien du tout [20]. (GOODMAN, 1993, p. 11)

[20] Ma traduction.

L'argument tend à dire qu'un objet est toujours conçu à partir d'une description, qu'il n'y a pas de désignation directe, mais toujours médiatisée, et ainsi que, sans le moyen de ce média, il n'y a rien que nous puissions atteindre. Pas d'étoile sans description d'étoile. Goodman constate que l'opinion commune considère bien que l'existence d'une étoile précède la version de monde qui la décrit. Or une telle manière de voir les choses relève justement d'*une version*. Mais il en est d'autres, la sienne par exemple, qui affirme qu'au contraire l'étoile naît *avec et par* la version qui la décrit comme étoile. Quelle version est correcte ? - les deux !, répond Goodman.

Il semble que le procédé dont fait usage Goodman soit plus sophistique que convaincant. Il consiste à prouver ce qui était déjà présupposé. En effet, Goodman ne répond pas vraiment à l'objection émise par Scheffler. En particulier, qu'on ne puisse strictement rien dire de la réalité ne peut pas servir de réfutation à l'existence d'une réalité indépendante et n'est pas non plus la preuve que des versions incompatibles puissent être simultanément correctes. Le monde ne cesserait pas d'exister avec la disparition du dernier être pensant. Et il n'a certainement pas commencé avec le premier.

Dans la belle thèse qu'il a consacré à Goodman, Alexandre Declos conclut avec Goodman qu'une réalité dont on ne peut rien dire serait une réalité bien pauvre. Précisément, Declos écrit que :

> Si cet appel à un monde aborigène conserve peut-être une certaine valeur affective et demeure en un sens inoffensif, il ne passe ni le test du rasoir d'Ockham, qui interdit la multiplication des entités sans nécessité, ni celui de la maxime pragmatiste, selon laquelle une dispute qui n'a pas de conséquences pratiques déterminables est tout simplement creuse. Si le Réaliste peut donc se prémunir de l'argument du monde perdu il remporte une victoire à la Pyrrhus, puisqu'il sauve le monde au prix de le rendre inconnaissable, imprédicable et surtout, inintéressant. (DECLOS, 2017, p. 421)

Il me semble cependant qu'il y ait là encore une confusion entre ce qui relève de l'épistémique et de l'ontologie. Que nous ne puissions connaître ni décrire les propriétés de la réalité n'a pas pour conséquence que la réalité en soit dénuée.

Le constructionnisme mondain de Goodman ne peut pas être écarté d'un nihilisme que l'auteur courtise dans de nombreux passages comme celui-ci :

> [...] le fait est bel et bien que monde n'existe plus et, avec lui, les stéréotypes ridicules de l'absolutisme [...] Où vivons-nous si quelque chose comme le monde n'existe pas ? On pourrait répondre « dans un monde », ou mieux, « dans plusieurs mondes » [...] on peut nier qu'il y ait un monde sans nier qu'il y ait des mondes. (GOODMAN & ELGIN, 1994, p. 49)

S'il n'existe pas de monde, mais des pluralités de mondes, c'est comme s'il n'en existait finalement aucun. Comme le montre Declos (DECLOS, 2017), il y a donc chez Goodman non seulement un anti-réalisme épistémique, mais aussi une véritable métaphysique basée sur un anti-réalisme ontologique.

L'hyperréel comme double du réel

L'âge d'une réalité partagée et commune à nos représentations est-il définitivement révolu ? Le réel serait-il devenu un présupposé humain, trop humain, une illusion sans lendemain ? Quoiqu'on puisse le regretter, Jean Baudrillard constate que la disparition du réel derrière l'écran du simulacre est l'une des principales caractéristiques des sociétés post-modernes. Baudrillard développe son ontologie à partir d'une sociologie analysant la circulation des objets, qui, selon l'auteur, est progressivement remplacée par une circulation infinie des signes. Les objets deviennent eux-mêmes des signes, mais des signes dont la dénotation est abolie, des signes qui se contentent d'une signification circulaire. Partant d'une analyse de la société de consommation, Baudrillard montre que l'interchangeabilité est passée des objets fongibles, comme le sont les céréales et la monnaie, aux objets abstraits que sont les signes. La société post-moderne serait ainsi caractérisée par l'extension hégémonique de la fongibilité. Tous les signes seraient advenus à une sorte d'équivalence hégémonique, une permutabilité totale.

En retraçant, de manière sémiologique, l'histoire de l'effacement du réel dans les sociétés post-modernes, Baudrillard distingue trois ordres de simulacres qui correspondent respectivement à trois étapes du développement des sociétés occidentales (BAUDRILLARD, 1976, p. 77).

La première occurrence de simulacres relève de la *reproduction*, simulacres qui sont « fondés sur l'image, l'imitation, la contrefaçon, harmonieux, optimistes, et visant à la restitution ou à l'institution idéale d'une nature à l'image de Dieu » (BAUDRILLARD, 1981, p. 177). Le simulacre de reproduction fait reposer la valeur économique sur une loi naturelle. Il correspond à l'époque classique qui s'étend de la Renaissance à la Révolution industrielle. Les sociétés du simulacre de reproduction présupposent un système social fortement inégalitaire et complètement hiérarchisé en valeurs aristocratiques où le mode de production des objets relève de la duplication des modèles aristocratiques. Par exemple, les meubles « Louis XVI » se veulent être les copies des meubles particuliers du roi. Dans la société féodale, la production des artefacts ressortissent à la reproduction d'objets originaux, singuliers, auquel la réplique réfère directement par un rapport d'authenticité ou de contrefaçon.

La *production* est le deuxième ordre de simulacres. Ceux-là sont « productifs, productivistes, fondés sur l'énergie, la force, sa matérialisation par la machine et dans tout le système de production - visée prométhéenne d'une mondialisation et d'une expansion continue, d'une libération d'énergie indéfinie » (BAUDRILLARD, 1981, p. 177). Ce type de simulacres ancre la valeur économique sur une loi marchande et correspond à l'ère industrielle. L'économie des sociétés du simulacre de production est encore proche de celles qui les précèdent en ce que la valeur marchande est encore ancrée à l'idée d'un modèle original. La voiture qui sort de la chaîne de production se veut encore être le double d'un prototype unique. En revanche, celui-ci n'est plus indexé par un système de références nécessairement aristocratiques. Ce deuxième type de simulacres peut donc être rapproché de ce que j'ai qualifié dans la première partie de ce mémoire de « répliques pures », modulo un degré d'imperfection dans la copie. Mais de manière au moins intentionnelle, les copies se veulent parfaitement indiscernables. Cette indiscernabilité des artefacts faisant d'ailleurs partie du cahier des charges de la production industrielle.

C'est différent avec le troisième type de simulacres, qui se rapporte à la *simulation*, et qui intervient dans les sociétés post-modernes. Les simulacres de simulation sont « fondés sur l'in-

formation, le modèle, le jeu cybernétique - opérationnalité totale, hyperréalité, visée de contrôle total » (BAUDRILLARD, 1981, p. 177). Avec le développement des sociétés de consommation, la prolifération des artefacts comme répliques tend à faire disparaître la dénotation au modèle. Les doubles ne sont plus hiérarchisés selon un ordre de ressemblance au modèle : une réplique vaut pour une autre. Plus encore, l'exemplaire unique vaut même pour toute la série. La fonction pratique d'un objet est complètement éclipsée par sa fonction de signe. Mais que dénote encore ce signe ? Ce n'est pas l'objet particulier dont on croit, en se procurant une copie, se doter du même mode de vie que son propriétaire. Dans les sociétés post-modernes, les objets deviennent leur propre messages, ils n'ont d'autre contenu de signification qu'eux-mêmes. Comme le dit Marshall McLuhan, le *médium est* le message.

Là se développe un processus d'effacement des référents au profit d'une équivalence généralisée des signes que Baudrillard qualifie d'« hyperréalité ». Le terme caractérise le rapport faussé au réel des sociétés qui assistent au renversement de la question métaphysique de Leibniz : « Pourquoi n'y a-t-il rien plutôt que quelque chose ? ». Contrairement à Goodman, chez Baudrillard, la disparition du réel n'est pas la conséquence logique d'une épistémologie anti-réaliste, mais un *événement* empirique et social, corrélé aux enjeux de la société de consommation. La conception de l'irréalisme est chez lui essentiellement sociologique. Le réel est absent non pas parce qu'il n'existerait pas, mais parce qu'on en a divorcé, comme s'il était devenu inutile, éclipsé par ses duplications médiatiques sans fin qui viennent à l'occulter sous le poids d'incessantes reproductions. Ainsi la guerre disparaît-elle - mais non ses effets dévastateurs - sous ses répétitions médiatiques (BAUDRILLARD, 1991) et ainsi la sexualité disparaît-elle dans la surenchère pornographique (BAUDRILLARD, 1979).

Selon Baudrillard, les deux tours du World Trade Center, en dépit de leur hauteur, ne représentaient pas la lutte vers les sommets, ce qui était encore l'objectif des tours précédentes qui les environnaient. Les tours manifestaient le paradigme de la société de simulation, symboles des pratiques de duplications visant à se débarrasser de la notion de référence. Chacune des Tours renvoyait en effet à l'autre, donc, par effet retour, à elle-même, dans une dénotation sans cesse renvoyée, et surtout, sans origine. Chacune des deux tours reproduisait l'autre, sans autre modèle. Ce dédoublement architectural marquait selon Baudrillard l'achèvement du siècle de la production de masse, de la réplique industrielle, de la société de consommation, au profit du nouveau paradigme de la simulation généralisée. Le dédoublement des tours exemplifiait le dédoublement incessant des signes dans les sociétés de simulation. C'est jusqu'à leur fin tragique que ces bâtiments sont les symboles du double, puisque *deux* avions les ont successivement percutés. L'effondrement de la tour sud fut suivi d'un second attentat conduisant à l'effondrement de la tour nord. Le second attentat constituait ainsi une sorte d'écho du premier, les deux événements étant encore démultipliés par le flux de leurs répétitions médiatiques incessantes.

Une formule permet à Baudrillard de comparer les sociétés encore ancrées sur la représentation du réel aux sociétés de l'hyperréalité : le réel est ce dont il est possible de donner une reproduction équivalente. Pour Baudrillard, cette définition a accompagné l'avènement de la science basée sur l'idée qu'une expérience, au sens d'expérimentation scientifique, est toujours reproductible. L'hyperréel, qui apparaît avec la société de consommation, a pour réquisit contradictoire un « système universel d'équivalence » où tout signe vaut pour tout autre. L'hyperréel est alors un réel qui n'est pas seulement reproduit, « mais qui est déjà toujours reproduit ». L'objet de consommation devient

signe autoleptique. Il n'intervient pas directement dans un rapport chose-représentation, mais dans une série d'équivalences entre des représentations de représentations. L'hyperréel, qui succède au réel, est alors vécu comme une hallucination esthétique produite par la prolifération des représentations, comme l'émergence d'une image dans des images, comme un phénomène rémanent, et non plus comme dénotation d'une chose par un signe.

Les trois ordres de simulacres encouragent respectivement trois types d'imaginaires successifs, soit, dans l'ordre chronologique, l'utopie, la science-fiction et enfin un nouvel imaginaire, qui n'a pas encore de nom chez Baudrillard, mais qui est caractérisé par la fin de la science-fiction, et dont les grands auteurs séminaux sont J.G. Ballard et Philip K. Dick. Ces trois types d'imaginaires peuvent être ordonnés selon la distance qui les démarque du réel. Dans l'utopie, cette distance est maximale, infinie. Là coexistent des mondes isolés les uns des autres, sans accessibilité possible. Dans la science-fiction, la distance au réel est largement abaissée, puisqu'elle ne dépend plus que d'une avance technologique et temporelle, une projection, une anticipation. Contrairement aux utopies, le monde de la science-fiction relève d'un univers unique, mais un univers en expansion spatio-temporelle et technologique qui permet encore une distance au réel. En revanche, dans les sociétés contemporaines, cette distance est annulée, elle est complètement abolie. La fiction coïncide alors avec le réel, et réciproquement. En cela, on peut rapprocher les simulacres des coïncidents purs : il n'y a plus aucun moyen de distinguer réel et irréel parce qu'il n'y a plus de relation spatio-temporelle permettant de séparer l'un de l'autre :

> Les modèles ne constituent plus une transcendance ou une projection, ils ne constituent plus un imaginaire par rapport au réel, ils sont eux-mêmes anticipation du réel, et ne laissent donc place à aucune sorte d'anticipation fictionnelle - ils sont immanents, et ne laissent donc place à aucune sorte de transcendance imaginaire. Le champ est ouvert à celui de la simulation au sens cybernétique, c'est-à-dire celui de la manipulation tous azimuts de ces modèles (scénarios, mise en place de situations simulées, etc.) mais *rien ne distingue cette opération de la gestion et de l'opération même du réel : il n'y a plus de fiction.* (BAUDRILLARD, 1981, p. 179)

En l'absence du réel, la volonté de démarquer la réalité de la fiction est par définition vaine, parce que le modèle est (vécu comme) le réel, le réel devenu un double de fiction.

L'hyperréalité, par le modèle et par son activité, la simulation, prétend supplanter la réalité par la représentation. Le virtuel pur, jadis élaboré par Aristote, comme une capacité d'instanciation, n'est ici plus vue comme destinée d'une actualisation différée. Il ne caractérise pas le statut de quelque chose qui serait en attente de réalisation. Il s'agit maintenant d'une substitution totale au réel, substitution qui est encore encouragée par les aspects maîtrisés, rationalisés et canalisés de la simulation. Baudrillard sacrifie ainsi le réel sur l'autel post-moderne :

> Faire advenir un monde réel, c'est déjà le produire et le réel n'a jamais été qu'une forme de simulation. On peut certes faire qu'il existe un effet de réel, un effet de vérité, un effet d'objectivité, mais, en soi, le réel n'existe pas. (BAUDRILLARD, 2000, p. 45)

Dans la simulation, le virtuel n'est plus vécu comme un mode de production du réel, mais comme un mode d'effacement, de renonciation au réel.

La réalité virtuelle

Chez Baudrillard, la simulation caractérise l'*absence* de tout original antérieur à sa duplication. Chez Umberto Eco (ECO, 1985), qui emploie lui aussi le même terme d'« hyperréalité », il s'agit de qualifier des situations culturelles spécifiques dans lesquelles la copie prédomine sur l'orignal, voire dans lesquelles la copie anticipe l'original.

C'est bien sûr dans le domaine des technologies du numérique que la notion connaît un regain d'intérêt, désignant la synergie des nouvelles technologies de la communication et de l'informatique avec le monde réel, donnant ainsi accès à une réalité « augmentée » en orientant les engagements intentionnels par des représentations d'espaces virtuels mêlés à l'espace physique :

> The concept of HyperReality (HR), like the concepts of nanotechnology, cloning and artificial intelligence, is in principle very simple. It is nothing more than the technological capability to intermix virtual reality (VR) with physical reality (PR) and artificial intelligence (AI) with human intelligence (HI) in a way that appears seamless and allows interaction. (TERASHIMA, 2001, p. 4)

L'imaginaire, le virtuel et le réel en viennent ainsi à se superposer.

Peut-on vraisemblablement parler de réalité à propos de mondes virtuels, comme l'engage l'expression de « réalité virtuelle » ? David Chalmers, dans un article de 2017 (D. J. CHALMERS, 2017), répond que la réalité virtuelle est véritablement une réalité, constituée de véritables objets et que ce qui arrive dans un monde virtuel est vraiment réel. Dans le domaine de la communication et de l'information, l'expression d'hyper-réalité désigne une aptitude technologique qui rend possible une « intégration transparente de la réalité matérielle, la réalité virtuelle, de l'intelligence humaine et de l'intelligence artificielle » (TERASHIMA, 2001, p. 4). John Tiffin et Nobuyoshi Terashima qualifient en particulier d'hyper-monde le monde constitué par le mixte entre la réalité physique et la réalité virtuelle. Un tel monde rend possible l'interaction entre des individus dispersés (en ce sens, un réseau social est un hyper-monde), mais aussi avec une forme de vie artificielle ou imaginaire générée par un ordinateur. Si elle n'est pas entièrement physique, il s'agit bien d'une réalité au sens où les intervenants co-agissent dans cet espace. Les applications sont nombreuses dans la recherche scientifique, dans l'enseignement, dans les milieux culturels (galeries et musées virtuels) et dans les loisirs (jeux vidéos), mais aussi dans l'industrie (manufactures à la demande) et même en politique comme en témoigne la conférence donnée simultanément à Lyon et à Paris par Jean-Luc Mélenchon [21].

Les mondes virtuels, s'ils sont seulement considérés comme des expériences immersives dans un monde différent du monde matériel, ne sont pas récents, puisqu'ils ressortissent à l'imaginaire :

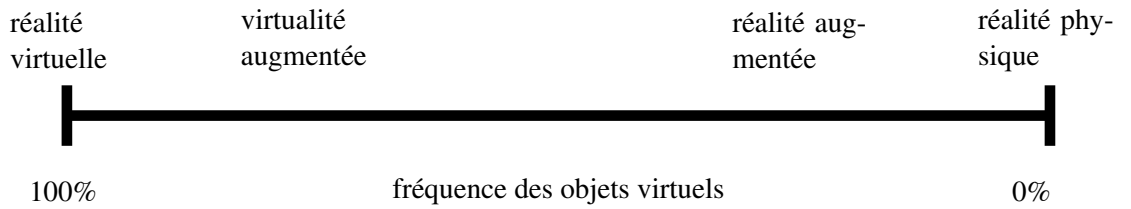
> Virtual worlds, or as we will now more broadly define them, immersive experiences delivered through the human imagination, have their origins in deep prehistory. Whether these varieties of nonphysical, dreamlike realities were communicated by our ancestors through the imitation of animals, the incarnation of spirits, the painting of scenes on the stone canvasses of caves, the holding of ceremonial rites in temples, or the elaboration of the human story through the

[21] La diffusion simultanée a été réalisée par un artifice comme ceux employés au théâtre, appelés « Pepper's ghost ». Il s'agissait en fait d'une double duplication alliant une projection vidéo à une projection spéculaire.

> fount of theater, humans have craved and crafted virtual-world experiences from the dawn of artistic and linguistic expression. (DAMER & HINRICHS, 2014, p.)

Quant à l'hyperréalité, elle se distingue des réalités virtuelles en ce que ces dernières sont exclusivement des mondes numériques. L'hyperréalité doit aussi être distinguée des réalités parallèles pour lesquelles l'usager choisit alternativement entre une représentation de la réalité physique ou une représentation entièrement numérique de cette même réalité physique [22]. Dans les réalités parallèles, le système fait cohabiter deux environnements que l'usager peut encore parfaitement démarquer. La notion d'hyperréalité se rapproche en revanche du procédé de réalité dite « augmentée » en ce que la démarcation naïve échoue, le monde virtuel coïncidant avec le monde matériel. [23] En outre, divers processus techniques peuvent accroître le réalisme de l'immersion dans le monde virtuel [24].

La réalité virtuelle [25] étant entièrement constituée d'objets qui sont par définition virtuels, sa démarcation avec l'environnement physique est nette. En revanche, la réalité augmentée relève d'un univers pour lequel la démarcation devient plus floue. On peut représenter les choses selon le diagramme le continuum réalité-virtualité de Paul Milgram et Fuimio Kishino (MILGRAM & KISHINO, 1994) :



Univers matériel et environnement virtuel figurent comme des pôles opposés, qui se croisent dans la réalité augmentée. Là où les univers virtuels demandent encore de rompre avec le monde matériel par un monde alternatif, la réalité augmentée semble consister en une juxtaposition d'environnements au sein d'un seule et même expérience, rendant caduque la séparation entre mondes fictifs et réalité matérielle. De ce point de vue, la fiction n'est plus une imitation du monde, une duplication mimétique, mais une co-construction d'un monde qu'on appelle réalité.

✤

Nous en sommes arrivés à l'idée que la fiction n'entretient pas un rapport de reproduction mimétique avec le réel. Je voudrais ici défendre l'idée que cette duplication est d'un autre ordre,

[22] On trouvera à l'adresse suivante une illustration des réalités parallèles : https://www.youtube.com/watch?v=UsDRPjDwr8A

[23] Le lecteur trouvera à l'adresse suivante une illustration de réalité augmentée https://www.youtube.com/watch?v=IIW3jsAiUyg

[24] Illustration d'une prothèse donnant l'impression de voler ou de nager dans un monde virtuel : https://birdly.com

[25] Le terme a été introduit par Antonin Artaud dans *Le théâtre et son double*. Au début des simulations virtuelles, dans les années 90, on utilisait également le terme « cyberspace », dont l'emploi est dû au *Neuromancien*, récit de science-fiction écrit par William Gibson.

métaleptique. La fiction n'entretient pas un rapport d'imitation à un modèle original, mais relève d'une forme de duplication par des représentations dans d'autres mondes. Le réel est alors caractérisé de manière épistémique comme un ensemble de mondes qui sont comme emboîtés, et non pas seulement comme des milieux qui se jouxtent en s'ignorant parfois. Cela n'implique aucunement de nier la réalité d'un monde, au sens ontologique, du terme. Nous prenons connaissance du monde par notre possibilité de nous en extraire, de nous affranchir de sa complexité infinie en nous projetant dans des mondes fictifs simplifiés. Ce sont les rapports d'accessibilité à des mondes épistémiques qui permettent de caractériser le monde réel par la médiation de rapports contrefactuels : nous nous approchons de ce qu'il y a par essais-erreurs.

Nous sommes néanmoins parvenus à une impasse. Cherchant à expliquer, comme l'invite Bachelard, la « filiation régulière du réel à l'imaginaire », nous sommes partis du constat que l'irréel, en particulier la fiction, pouvait relever du vrai, mais qu'en le qualifiant ainsi, nous butions sur la conception quinéenne de la vérité ancrée sur l'ontologie. Toutes les assertions fictionnelles sont, de ce point de vue, fausses. Comme nous l'avons vu dans les deux sections précédentes, le pan-réalisme et le pan-fictionnalisme se veulent être deux solutions contraires à ce problème, consistant à définir la fiction comme un double métaphysiquement indiscernable du réel respectivement par un élargissement de l'ontologie à la totalité des expressions signifiantes et par réduction drastique de la dénotation.

Quelle perspective avons-nous pour éviter de tels points de vue trop radicaux pour s'avérer compatibles avec nos intuitions ? Comment conserver l'idée que l'irréel soit un double épistémique du réel sans en faire un double ontologique ? Comment rendre compte des connaissances irréelles vraies tout en reconnaissant que les termes fictionnels sont hors d'atteinte du réel, c'est-à-dire sans souscrire à une conception purement mimétique ? Comment caractériser la fiction comme une contrepartie épistémique du réel, en conservant l'idée qu'elle se distingue des faits par l'absence de référence, mais en lui conférant toutefois la vertu d'un pouvoir explicatif, autrement dit, sans rendre équivalent l'irréel et l'illusoire ? Comment parvenir à faire ainsi coïncider absence de référence et teneur sémantique ? C'est ce que nous nous proposons de considérer dans la section suivante.

3.3 Les mondes possibles comme doubles épistémiques

Dans l'*Art poétique*, Horace écrit qu'« un poème est comme un tableau » (HORACE, 1944, P. 285, 361). De fait, la théorie mimétique des arts, sous le patron d'Aristote, a suivi une longue filiation s'étendant jusqu'à l'essai d'Erich Auerbach, *Mimésis. La représentation de la réalité dans la littérature occidentale* (1946) (AUERBACH, 1968) où l'auteur liste les avatars de la notion d'imitation. Selon Auerbach, la fiction relèverait à la fois de la dépiction et de la dénotation. De fait, de théories en théories, le concept de mimésis a bien évolué depuis Aristote. La question de la référence s'est largement complexifiée depuis l'Antiquité, comme le constate d'ailleurs Auerbach qui souligne que « la conception de la réalité qui se dégage des œuvres de la fin de l'Antiquité et du Moyen Âge diffère entièrement de celle du réalisme moderne » (AUERBACH, 1968, p. 551). Pour ce qui concerne notre propre point de vue, nous voudrions indiquer ici une conception basée sur l'usage des mondes possibles, favorisant l'idée que la fiction représente un double épistémique du monde réel, en relativisant la vérité aux mondes possible par une sorte de perspectivisme.

3.3.1 Le perspectivisme nietzschéen

Avant d'aller plus loin, je dois préciser pourquoi je ne place pas la pensée de Nietzsche dans la section consacrée aux philosophes irréalistes, mais bien à l'intérieur d'une section se rapportant à ceux qui veulent rendre au réel ses lettres de noblesse. La raison est qu'il me semble que chez Nietzsche, le réel est loin d'être réduit à une construction mentale. Le marteau nietzschéen n'est pas ontologique, mais épistémique. Que le langage transforme ce qu'il décrit en fiction n'invite nullement à croire que tout est fiction. Car il est toujours possible de référer au Réel (je viens de le faire), même si nous manquons bien vite de termes pour le décrire de manière objective. Bien sûr, il existe une tradition d'exégèse des œuvres de Nietzsche, véhiculée par les travaux de Bataille, de Blanchot et de Derrida, qui s'inscrit en faux contre ce que j'indique ici. Mais notons que ces auteurs ont tous été influencés par l'interprétation très discutable que Heidegger a tirée des écrits de Nietzsche. Selon ces philosophes, l'auteur du *Gai Savoir* défendrait une sorte de nihilisme pan-fictionnaliste qui annoncerait d'ailleurs le post-modernisme. Cela semble un évident contre-sens si l'on considère par là que Nietzsche remettrait en question l'existence du réel. On peut rappeler que toute sa philosophie éthique et pratique s'inscrit justement comme une invitation inconditionnelle à affirmer la réalité. Deuxièmement, le nihilisme fait justement l'objet de la critique de Nietzsche, qui le considère comme le soubassement de toute métaphysique qui nie la vie et dont les symptômes sont la mauvaise conscience et le ressentiment à l'égard du réel. Le but visé par la « transmutation des valeurs » est justement de libérer la pensée du nihilisme métaphysique.

Il apparaît certes des passages où l'existence d'une réalité en soi semble remise en question, et il est indéniable que Nietzsche conteste en particulier la notion de vérité objective. Ainsi, dans le *Crépuscule des Idoles*, passant en revue les étapes de la conception de la réalité, il accuse l'histoire d'être une « longue erreur », comme le spécifie le sous-titre du chapitre en question. Primitivement vécu comme vrai et accessible, le réel a progressivement été rendu inaccessible (Kant) puis complètement aboli (nihilisme) :

> Nous avons aboli le monde vrai : quel monde restait-il ? Peut-être celui des apparences ? Mais non ! *En même temps que le monde vrai, nous avons aussi aboli le monde des apparences !* (NIETZSCHE, 1974b, p.81)

En outre, dans *Humain, trop humain*, Nietzsche écrit de manière un peu plus obscure que :

> il est vrai qu'il pourrait y avoir un monde métaphysique ; la possibilité absolue n'en est guère contestable. Toutes les choses que nous regardons passent par notre tête d'homme et nous ne saurions trancher cette tête ; la question n'en demeure pas moins de savoir ce qu'il resterait du monde une fois qu'on l'aurait cependant tranchée. [...] de ce monde métaphysique, on ne pourrait rien affirmer sinon une différence d'être, être et différence qui nous sont inaccessibles, inconcevables ; ce serait une chose à qualités négatives - Et quand bien même l'existence de ce monde serait on ne peut mieux prouvée, il n'en serait pas moins certain que cette connaissance serait justement de toutes la plus indifférente : plus indifférente encore que ne l'est au marin menacé par la tempête la connaissance de l'analyse chimique de l'eau. (NIETZSCHE, 1988, p. 36, I, § 9)

Mais il est essentiel de distinguer ici entre une thèse qui affirme que, effectivement, la vérité



FIGURE 3.2 : Nietzsche et son double ?

est humaine, trop humaine, et une thèse que rejetterait la réalité en soi, thèse qui n'est pas soutenue par Nietzsche. Car de fait, la première n'implique pas la deuxième. Le prétendre serait faire une confusion entre irréalisme épistémique et irréalisme ontologique. La remise en cause d'une putative réalité chez Nietzsche l'est toujours dans son rapport à la vie, qui représente de fait la seule réalité et que Nietzsche ne remet jamais en question. Le philosophe défend, avec acharnement, une thèse vitaliste peu compatible avec le nihilisme. Le réel qui est conjuré dans les extraits précédents est bien celui des métaphysiciens de tous ordres et de leurs vérités d'arrière-monde, et qui s'opposent au « vrai » réel que constitue l'existence, au sens vital, de toute chose.

En dépit de quelques penseurs français qui ont rendu justice à l'auteur du *Gai Savoir*, dont Clément Rosset (ROSSET, 1983), la méconnaissance de l'auteur est certainement aussi grande que la part des discussions qui l'entourent. À titre d'anecdote éclairante et symptomatique, rappelons que le livre que Maurizio Ferraris prétendument consacré à Nietzsche, intitulé *Friedrich Nietzsche, Il re e il nichilista, un mistero coi baffi* (1999) arborait sur sa couverture le portrait du roi d'Italie Umberto I de Savoie (figure 3.2). Cette erreur répétait, de manière entêtée, celle déjà commise deux années plus tôt avec la publication du livre de Giorgio Penzo *Nietzsche allo specchio* (figure 3.2). Cela indique, il me semble, le juste degré de précision dont est affublé ce philosophe à qui on attribue non seulement des photographies, mais également des thèses qui n'ont jamais été les siennes. On peut se demander d'où vient une telle fin de non-recevoir qui tient manifestement d'un évident déni de réalité appliqué avec acharnement aux écrits de cet auteur. À ce sujet, Clément Rosset avance une hypothèse d'ordre psychologique : « La raison fondamentale de ce rejet me paraît résider en ceci que tout discours totalement affirmateur comme l'est celui de Nietzsche [...] est et a toujours été reçu comme totalement inadmissible » (ROSSET, 1983, pp. 33-34).

L'héritage kantien de la distinction entre monde en soi et monde des phénomènes, par la médiation de la double composante que Schopenhauer tire du monde comme représentation et comme volonté, se constitue chez Nietzsche comme une pleine accusation des métaphysiques. Selon Nietzsche, le tribunal instauré par les critiques de Kant n'a pas tenu son rôle, laissant intact les prétentions métaphysiques qui déferleront d'ailleurs, après Kant, avec l'idéalisme de Schelling et de Hegel. Aussi, Nietzsche accuse Kant de faux monnayage critique et d'avoir, *in fine*, fait triompher une théologie chrétienne pleine de ressentiment à l'égard de la vie, en instaurant un nihilisme tenace. Mais les putatives « réalités métaphysiques », qu'elles soient incarnées par une intériorisation du monde idéalisé (« Moi, Platon, *je suis la vérité* » (NIETZSCHE, 1974b, p. 80)), ou ancrées sur des apparences subjectives, sont toujours des mondes fabriqués pour fuir la

présence par trop menaçante du réel. La véritable critique ne doit pas se limiter à constater que les choses-en-soi demeurent hors de portée de la raison. Une véritable critique doit aussi s'atteler à démonter méthodiquement les illusions qui nourrissent encore les arrières-mondes, la facticité de ces choses-en-soi, l'artifice qu'elles représentent à l'égard du réel.

Quel est alors le but de la métaphysique ?

> La métaphysique explique en quelque sorte *pneumatiquement* le Livre de la nature, comme l'Église et ses docteurs faisaient autrefois la Bible. Il faut beaucoup d'intelligence pour appliquer à la nature le même genre d'interprétation rigoureuse que les philologues ont désormais établi pour les livres : en vue de comprendre simplement ce que le texte veut dire, mais sans y flairer, ni même supposer, un *double* sens. Pourtant, tout comme le mauvais moyen d'explication est loin d'être entièrement aboli même en ce qui touche les livres, et qu'on se heurte constamment, dans la meilleure société cultivée, à des vestiges d'interprétation allégorique et mystique : il en va de même en ce qui concerne la nature - et encore, c'est bien pis. (NIETZSCHE, 1988, p. 36 I § 8)

Nietzsche souligne ainsi que l'existence est interprétée à travers des grilles de lecture qui ne sont pas passives, mais qui refluent le chaos du monde en instituant des lignes directrices humaines, trop humaines. La répétition et la régularité sont des inventions qui n'ont pas de pendants dans un réel toujours singulier et à jamais neuf. Aussi, écrit Nietzsche, « nous nous sommes construit un monde dans lequel nous puissions vivre - en supposant des corps, des lignes, des surfaces, des causes et des effets, le mouvement et le repos, la forme et le contenu : sans pareils articles de foi, nul à présent ne supporterait de vivre ! » (NIETZSCHE, 1982, § 121) Les mondes alternatifs au réel, qui sont les mondes que nous construisons, sont des mondes *curatifs*, nous soignant du contact avec ce qui est, et qui est souvent de trop. Le double, comme ensemble de mondes de représentations, sert ainsi de pharmacopée contre l'effet psychologiquement délétère du réel. C'est dire que la réalité n'est pas seulement une construction mentale, mais une donnée que nous subissons, et cela, de mauvaise grâce.

Notons, à titre d'illustration, que, dans *Anywhere out of the world*, Baudelaire a donné une admirable description de ce désir de fuite hors d'une réalité *par nature* inacceptable :

> Cette vie est un hôpital où chaque malade est possédé du désir de changer de lit. Celui-ci voudrait souffrir en face du poêle, et celui-là croit qu'il guérirait à côté de la fenêtre. Il me semble que je serais toujours bien là où je ne suis pas, et cette question de déménagement en est une que je discute sans cesse avec mon âme. [...] Enfin, mon âme fait explosion, et sagement elle me crie : « N'importe où ! n'importe où ! pourvu que ce soit hors de ce monde ! » (BAUDELAIRE, 1961, p. 303)

Les mondes, qu'ils soient ceux de l'art, de la science et de la religion servent ainsi de providence dérisoire à celui qui veut s'arracher au Réel. Nous déménageons notre désespoir du réel dans une thébaïde provisoire. Mais ces duplications n'en sont justement, que des doubles, c'est-à-dire de vagues ombres, des reflets bien sommaires qui ne font pas le poids devant la réalité totalitaire et toujours excédentaire du réel.

À l'inverse de la mélancolie baudelairienne, il y a chez Nietzsche un accueil bienveillant à l'égard du réel, et, dirons-nous même, plus que cela, une sorte de radicalisation de l'hypothèse

leibnizienne du meilleur des mondes possibles. Dans *Le Gai Savoir*, au livre IV « *Sanctus Januarius* », l'auteur écrit « Maintenant, en effet, la pensée d'une providence personnelle se présente à nous de la plus envahissante façon, et elle a pour elle le meilleur porte-parole, l'apparence, dès lors qu'il nous est tangible que toutes choses, absolument toutes choses qui nous adviennent, *tournent constamment à notre avantage* » (NIETZSCHE, 1982, p. 190). Car ce ne sont pas que les choses agréables qui adviennent ainsi, mais absolument tout, « qu'il s'agisse de n'importe quoi, du mauvais comme du beau temps, de la perte d'un ami, de maladie, de calomnie, de la non-venue d'une lettre, d'un pied foulé, d'un coup d'œil dans un magasin, d'un contre-argument, d'un livre ouvert au hasard, d'un rêve, d'une tromperie : l'événement se révèle aussitôt ou bientôt après comme quelque chose qui ne pouvait pas ne pas se produire. » (NIETZSCHE, 1982, p. 190) Ce n'est pas la fatalité ou le destin que décrit Nietzsche ici, mais le hasard : « en réalité, çà et là, quelqu'un joue *avec nous* – le cher hasard : il mène notre main à l'occasion, et la providence la plus sage ne saurait inventer plus belle musique que celle qui alors réussit à notre main insensée. » (NIETZSCHE, 1982, p. 190). La thèse d'affirmation du réel consiste à considérer le monde réel comme *le meilleur possible*. Ce qui différencie ici Nietzsche de Leibniz, c'est que la sélection ne provient plus de l'interaction d'un esprit divin [26], mais de l'hypothèse de l'éternel retour basé sur une sélection volontaire. Plutôt que de céder aux doubles que sont les arrières-mondes, il s'agit de consacrer le hasard des choses en les affirmant. Le principe nietzschéen de l'éternel retour est fondamentalement une consécration de la vie, ce qui est, indépendamment de toute attribution de valeur à ce qui arrive. Ainsi, l'affirmation inconditionnelle de l'existence est sans doute encore la création d'un double, puisque la chose en soi demeure inconnaissable. Mais il s'agit ici d'un double de superposition, autrement dit, un double qui tente de faire coïncider l'existence des choses avec nos représentations. C'est ramener les ombres à leur plus courte portée, atteindre *le midi* de l'existence.

Le perspectivisme de Nietzsche pourrait ainsi nous encourager à développer un point de vue qui admet l'irréalité des fictions sans pour autant renier la place de la réalité. Si certains déménagements métaleptiques dans des mondes épistémiques ont, comme le rappelle Baudelaire, une visée purement curative visant à éclipser le réel, d'autres ont aussi une visée plus affirmative : nous connaissons le réel par ses ombres portées, par ses doubles de proximité que sont les mondes épistémiquement pertinents.

3.3.2 Relativiser la vérité aux mondes

Quand nous disons par exemple qu'« il n'y a pas un chat » nous voulons dire qu'il n'y a personne dans le lieu que nous occupons ici et maintenant, et pas du tout dans l'univers entier. Nos domaines de quantifications sont habituellement restreints par leur contexte d'énonciation. Littéralement, il est faux que Sherlock Holmes existe, cependant, il est vrai qu'il y a un personnage du nom de Sherlock Holmes dans l'œuvre de Conan Doyle. En revanche, il n'y existe pas de personnage du nom de « Bruce Wayne ». Le principe de contexte permet de médiatiser la vérité d'un énoncé.

[26] Au passage, et pour contrevenir encore une fois à une autre idée reçue et largement diffusée, Dieu n'a pas été « tué par Nietzsche », puisque s'en sont chargés bien avant lui les premiers chrétiens. Nietzsche, qui était, rappelons-le, philologue, insista sur les conséquences qui devaient être tirées du christianisme dès ses premières heures, à savoir, la mort de Dieu sur la croix. Le cadavre n'était donc pas frais à l'époque de Nietzsche, et il dégageait depuis longtemps cette épouvantable odeur que Nietzsche qualifiait de « nihilisme ».

Bien sûr, dans le cas de fictions, nous ne pouvons plus faire appel au contexte spatio-temporel de la réalité. Mais il demeure possible de faire intervenir d'autres mondes. La sémantique des mondes possibles permet en effet d'ancrer la vérité d'une proposition atomique à un monde. Cette solution qui contextualise les vérités aux mondes, permet de rendre compte du caractère objectif de nos énoncés fictionnels sans devoir faire l'hypothèse d'objets inexistants ou d'artefacts abstraits. Nous dupliquons ainsi le monde réel par des mondes possibles proposant des conditions de vérités différentes sur les énoncés atomiques, tout en conservant les liens de déduction logique par le calcul propositionnel appliqué aux énoncés complexes. La sémantique des mondes possibles nous sera donc utile pour contextualiser les vérités aux mondes.

Modalités épistémiques et doxastiques

Comme nous l'avons vu dans la première partie de ce travail, le cadre formel de la sémantique des mondes possibles permet d'exprimer la possibilité et la nécessité. Mais une logique de la fiction nécessite plus particulièrement que soient explicitées les relations entre possibilité et conception, autrement dit, que l'articulation entre la métaphysique et l'épistémique soit clarifiée.

Dans les années soixante [27], Jaakko Hintikka (HINTIKKA, 1962) a modélisé les attitudes de connaissance (attitudes épistémiques) et de croyance (attitudes doxastiques) à partir des structures de mondes possibles. Dans *Semantics for Propositional Attitudes*, Hintikka résume son modèle théorique en affirmant que, pour un agent donné, les mondes possibles peuvent être rangés en deux classes : ceux qui sont compatibles avec la construction conceptuelle de l'agent, et ceux qui ne le sont pas.

> Mon hypothèse de départ (un peu trop simplifiée) est que toute attribution d'attitude propositionnelle [...] implique une division de tous les mondes possibles (plus précisément, tous les mondes possibles qu'on peut distinguer dans la partie du langage que nous utilisons en produisant l'attribution) en deux classes : dans celle des mondes possibles qui sont conformes avec l'attitude en question et dans celle des mondes incompatibles avec elle. La signification de la division dans le cas d'attitudes telles que connaissance, croyance, mémoire, perception, espoir, volonté, effort, désir, etc. est suffisamment claire. Par exemple, si ce dont nous parlons sont les souvenirs de (disons) a, alors, ces mondes possibles sont tous les mondes possibles compatibles avec tout ce dont il se souvient [28]. (HINTIKKA, 1969b, p. 25)

Ainsi, le degré de pertinence que nous accordons à des mondes possibles relève d'un fait idéologique (au sens de Quine, autrement dit, opposé à « ontologique ») :

> Notre engagement "idéologique" envers des mondes possibles autres que le monde actuel n'est ni surprenant ni déconcertant. Si ce que nous traitons est ce que les gens font – plus particulièrement, les concepts qu'ils utilisent – afin d'être préparés à plus d'une éventualité, il n'est pas du tout étonnant que pour décrire ces concepts de manière complète, nous devions parler du cours d'événements autres que le cours actuel [29]. (HINTIKKA, 1969b, p. 29)

[27] Cette section s'inspire d'un cours de logique modale de Manuel Rebuschi.
[28] Ma traduction
[29] Ma traduction

Afin de formaliser de telles structures de mondes, Hintikka introduit, à côté des opérateurs aléthiques de nécessité et de possibilité, deux autres opérateurs, l'un de connaissance (noté « K »), et l'autre de croyance (noté « B »), qui permettent de conférer une signification aux propositions intentionnelles. Ensuite, on sélectionne une classe de structures parmi celles qui valident les formules qui corroborent nos intuitions pré-théoriques. Par exemple, si l'on opte pour une interprétation doxastique, il n'est alors pas souhaitable de retenir la réflexivité pour l'opérateur de croyance puisque celle-ci impliquerait que tout ce qui est cru est vrai, ce ne n'interprète pas correctement une attitude de croyance.

Admettons que Jean ne sache rien de Maupassant. On interprétera la connaissance de Jean selon une structure modale où tous les mondes possibles sont reliés les uns aux autres et à eux-mêmes, car, en termes de connaissance de Jean, ces mondes sont tous indiscernables. Autrement dit, le cas de mondes indiscernables correspond, en logique épistémique, à une absence de connaissance. Admettons ensuite que Jean apprenne si Maupassant est ou non un auteur, mais qu'il ne sache pas si Maupassant vit encore ou non. On interprétera la connaissance de Jean selon une nouvelle structure modale qui rend accessibles entre eux les mondes indiscernables selon la connaissance de Jean. Comme on peut le voir, les relations d'accessibilité s'appauvrissent au fur et à mesure que progresse la connaissance de Jean.

Précisions que la logique épistémique obéit aux trois principes suivants :

- Principe de clôture sous la dérivabilité : si $\mathrm{K}\Phi$ et si $\Phi \rightarrow \Psi$, alors $\mathrm{K}\Psi$.
- Règle de nécessitation : Si Φ est valide, alors $\mathrm{K}\Phi$.
- Principe de consistance : $\neg(\mathrm{K}\Phi \wedge \mathrm{K}\neg\Phi)$

Le principe de clôture affirme que toutes les conséquences logiques découlant d'une proposition connue sont elles-mêmes connues. La règle de nécessitation affirme que si une proposition est valide, alors elle est connue. Enfin, le principe de consistance affirme que nous ne pouvons pas connaître des choses contradictoires. On voit donc que cette modélisation traduit une forme très idéalisée de connaissance qui conduit au problème de l'omniscience logique, autrement dit, à l'idée que le sujet connaisse toutes les conséquences logiques de ce qu'il sait. Or connaître les axiomes de l'arithmétique n'implique manifestement pas que l'on sache si la conjecture de Goldbach est vraie ou fausse. Différentes stratégies peuvent néanmoins être mobilisées afin de rendre la modélisation plus fidèle à la connaissance humaine. L'une d'elles, due à Veikko Rantala (RANTALA, 1989), consiste à introduire des mondes impossibles dans lesquels les propositions complexes ne sont plus définies de manière récursives mais de façon atomique : les valeurs de vérité ne sont pas déduites, mais stipulées. La structure des mondes possibles est ainsi dotée d'au moins un monde impossible accessible depuis le monde actuel et contrevenant à la clôture, à la validité, et/ou à la consistance.

L'opérateur de fictionnalité

Les attitudes propositionnelles ne se limitent évidemment pas à la croyance et à la connaissance. Aussi, il est possible d'introduire également un opérateur Ⓡ qui sera interprété par « se représenter

que », « concevoir que » ou « imaginer que », en concevant ® comme une sorte d'opérateur $\Box$ réinterprété. En acceptant ceci, on détermine alors une logique ancrée sur la logique modale S5 et étendue à cet opérateur, en faisant jouer à ® le rôle joué par K en logique épistémique.

C'est dans *Truth in Fiction* que David Lewis a élaboré une sémantique pour les fictions basée sur des relations d'accessibilité à des mondes possibles dont le contenu réalise ce qui est dit dans le récit. L'objectif de Lewis est de proposer un modèle sémantique qui puisse expliquer nos usages fictionnels, en particulier lorsque ceux-ci attribuent une valeur de vérité à certaines expressions fictionnelles comme « Sherlock Holmes est un chauffeur de taxi » ou « Sherlock Holmes est un détective ». Il s'agit en particulier de rapprocher leur fonctionnement sémantique d'énoncés factuels qui portent sur des entités bien réelles.

Rappelons que, de leur coté, les thèses meinongiennes effectuent ce rapprochement en conférant une dénotation aux noms fictionnels par le recours à des entités non-existantes. Mais Lewis constate que ces thèses butent sur diverses contradictions qui émanent de simples énonciations comme celles-ci :

($\mathbf{P}_1$) Sherlock Holmes est un détective.

($\mathbf{P}_2$) Sherlock Holmes n'existe pas.

En effet, ce que nous voulons dire lorsque nous énonçons quelque chose comme la première des deux phrases précédentes, c'est que Sherlock Holmes est un humain pratiquant une certaine profession, et plus précisément qu'il est une personne en chair et en os, autrement dit, une entité biologique et non une entité exotique. En outre, bien que certaines fiction traitent de fantômes, de zombies, d'elfes ou de personnages qui se métamorphosent, ce n'est certainement pas le cas des fictions de Sherlock Holmes. Comme le souligne Lewis, Sherlock Holmes est un être de la même catégorie que Nixon, et non une entité à la biologie déviante qui pourrait s'affranchir impunément des lois physiques.

En outre, soulignons aussi que si chacune des deux propositions exprimées par les phrases précédentes semble vraie, la réunion des deux implique néanmoins une évidente contradiction. Pour éviter de telles difficultés, qui surgissent entre les usages factuel et fictionnel des termes, la stratégie de Lewis consiste à réduire la portée de la quantification par l'effet d'une paraphrase du type : « dans la fiction f, ... ». Cette expression fait alors office d'opérateur intensionnel afin de constituer une nouvelle proposition. Il arrive parfois que l'opérateur soit omis par abréviation, comme lorsque nous disons tout simplement « Sherlock Holmes est un détective ». Un tel énoncé intrafictionnel, quand il est considéré littéralement sans préfixe de fictionnalité, est alors traité à la manière de Frege ou de Russell, à savoir qu'il est ou dénué de valeur de vérité, ou faux. Lewis ne se prononce d'ailleurs pas sur le type de traitement qu'on doit lui réserver. En revanche, ces énoncés gagnent une valeur de vérité à partir du moment où ils sont explicitement ou implicitement préfixés par un opérateur de fiction. Ainsi, la phrase :

« Dans la fiction de Conan Doyle, Sherlock Holmes est un détective. »

exprime bien une proposition vraie. Et pour les énoncés métafictionnels comme « Sherlock Holmes n'existe pas », le traitement quantificationnel à la manière de Russell permet de rendre

compte de leur vérité. En outre, cela explique aussi la fausseté et la vérité respectives des deux propositions exprimées par les phrases suivantes :

« Dans la fiction de Conan Doyle, Sherlock Holmes n'existe pas. »

« Dans la fiction de Conan Doyle, Batman n'existe pas. »

Plus précisément, l'interprétation de l'opérateur intensionnel s'appuie sur une structure modale de Kripke. On dira alors que la proposition exprimée par la phrase « Dans la fiction f, Φ » est vraie si et seulement si Φ est vraie dans chaque monde possible d'un certain ensemble, cet ensemble étant déterminé par la fiction f. Les structures de Kripke permettent ainsi de relativiser la vérité au monde considéré. Seules les vérités logiques nécessitent d'être vraies dans *tous* les mondes. En revanche, une assertion qui n'est pas exclusivement logique peut s'avérer vraie dans un monde, tout en étant fausse dans un autre. Les mondes possibles sont alors décrits comme les mondes qui réalisent ce qui est écrit (ou dit) dans la fiction considérée. L'avantage de cette distinction est qu'elle rend compte à la fois du caractère littéralement faux des énonciations fictionnelles dans le monde actuel (Holmes n'existant pas, il ne peut être un détective) et à la fois de la possibilité d'attribuer le vrai à certains énoncés fictionnels en dehors du monde actuel.

Soulignons que l'un des bénéfices de la stratégie du préfixe intentionnel de fiction permet notamment de rendre compte de la clôture logique des énoncés fictionnels. En effet, si on a que :

$$\frac{\Psi_1, \ldots, \Psi_2}{\Phi}$$

alors il vient aussi, en préfixant par « Dans f » que :

$$\frac{\text{Dans } f,\ \Psi_1, \ldots, \text{Dans } f,\ \Psi_2}{\text{Dans } f,\ \Phi}$$

Évidemment, si, dans les arguments initiaux, on préfixe certaines prémisses mais pas toutes, alors en général, ces déductions ne sont plus valables.

Ainsi, la stratégie de Lewis a pour avantage de ne pas nécessiter d'entités abstraites ou inexistantes. Dans le monde de la fiction, Holmes est un homme au même titre qu'Emmanuel Macron. Son corps est matériel, il demande à être alimenté régulièrement, il respire par des poumons, il transpire et son rythme cardiaque s'accélère quand il se met à courir. Cette explication présente donc un sérieux avantage sur celles qui considèrent les personnages comme des objets abstraits ou inexistants, encodant des propriétés sans les exemplifier.

Avec l'opérateur de fictionnalité, Lewis re-dirige l'inactualité d'un récit fictionnel vers des mondes dans lesquels cette histoire est l'expression de faits authentiques. Il procède ainsi à un transfert indexical : si un nom fictionnel ne réfère pas à un existant dans le monde actuel, dans d'autres mondes il est néanmoins possible que ce nom réfère à un individu. Dans un premier temps, Lewis considère l'hypothèse naïve que les mondes possibles sur lesquels débouche l'opérateur de fictionnalité soient ceux où a lieu l'intrigue de la fiction, autrement dit, les mondes qui réalisent la suite des événements décrits dans l'histoire. Ce qui est vrai dans les histoires de Sherlock Holmes

correspond alors à ce qui est vrai dans tous les mondes possibles où les événements coïncident avec ce qui est dit dans les histoires de Conan Doyle. Bien sûr, cela fonctionne au bénéfice d'une reproduction mimétique du réel *ceteris paribus*, c'est-à-dire, modulo les changements assertés dans le récit.

Il y a pourtant diverses difficultés qui entravent l'explication naïve. La première est qu'il n'est pas toujours évident de déterminer les mondes possibles qui correspondent à l'histoire décrite dans la fiction, tout simplement parce que cette histoire peut être difficile à démêler, ou que le narrateur peut s'avérer non fiable. Comme le souligne Lewis, point n'est besoin pour cela de faire appel à quelque récit post-moderne : mêmes les histoires de Holmes ne ressemblent pas à de franches chroniques. L'extrapolation de l'histoire hors du texte, qui seule permet de configurer l'ensemble des mondes possibles accessibles par l'opérateur de fiction, ne se fait jamais sans difficultés. C'est ce qui fait dire à Lewis que :

> Peut-être que le lecteur l'accomplit seulement en comprenant ce qui est vrai dans les histoires – c'est-à-dire, seulement en exerçant sa maîtrise tacite du concept de vérité dans la fiction que nous sommes maintenant en train de chercher. (LEWIS, 2012, p. 40)

Mais dans ce cas, nous sommes malheureusement conduits à un cercle vicieux, puisque, rappelons-le, l'objectif de la théorie de Lewis est justement d'expliquer nos attributions de vérité *à partir* de la compréhension de l'histoire. Autrement dit, définir la vérité fictionnelle selon ce qui est vrai dans des mondes possibles présuppose de pouvoir comprendre l'intrigue racontée, mais cette compréhension passe justement par le fait de déterminer ce qui est vrai dans l'histoire.

Le deuxième problème sur lequel vient buter la conception naïve de l'opérateur fictionnel provient d'une objection de Kripke qui est relative à un cas d'indiscernabilité entre histoire factuelle et histoire fictionnelle. Celle-ci consiste à faire l'hypothèse que tout ce qui est décrit dans un récit de fiction est, incidemment, factuel. Mettons qu'une histoire de Conan Doyle ait véritablement lieu, à l'insu de l'auteur, dans notre monde [30]. Précisément, admettons qu'il existe, dans le monde réel, un individu du nom de Sherlock Holmes, détective londonien ayant vécu toutes les aventures du personnage de Conan Doyle et exactement celles-là. L'hypothèse est certainement improbable, mais pas impossible, comme en témoigne d'ailleurs la déclaration traditionnelle qui avertit les spectateurs d'un film que « les personnages et les situations de ce récit étant purement fictifs, toute ressemblance avec des personnes ou des situations existantes ou ayant existé ne saurait être que fortuite ». S'agirait-il alors de la même personne ? C'est ce que conteste Kripke :

> La simple découverte qu'un détective ayant accompli les exploits semblables à ceux de Sherlock Holmes a effectivement existé ne montrerait pas que c'est *de cet homme* que parlait Doyle. (KRIPKE, 1982, p. 146)

Il est évident que le nom utilisé par Conan Doyle, bien qu'homonyme, ne dénoterait pas le personnage réel, parce que l'intention de Conan Doyle ne serait pas telle. Conan Doyle veut parler de quelqu'un qui n'existe pas, qui est purement imaginaire. En outre, l'objection de Kripke ne s'arrête

[30] Les conséquences de l'hypothèse qu'un récit de fiction coïnciderait pleinement avec la réalité ont aussi été développées par H.N. Catstañeda (CASTAÑEDA, 1979) sans un angle réaliste des entités fictionnelles. L'auteur imagine qu'un récit de fiction précède un article de journal qui décrirait, par coïncidence, exactement la même histoire.

pas à l'hypothèse d'un individu réel indiscernable des descriptions fournies par Conan Doyle. Sa portée est modale et s'étend plus généralement à tout individu possible. En effet, Kripke souligne que, dans le cas plus probable où une telle personne n'existerait pas actuellement, on n'aurait de toute façon pas non plus pu dire d'aucune personne que, si elle avait existé, elle aurait été Holmes [31]

3.3.3 La vérité en fiction

Lewis concède à l'objection de Kripke que « notre monde ne peut pas être un tel monde, car si ce dont il s'agit est vraiment une fiction, alors l'acte de raconter dans notre monde n'était pas ce qu'il prétendait être » (LEWIS, 2012, p. 42). Pour éviter une éventuelle collusion avec le monde réel, Lewis en conclut que nous devons trouver quelque chose qui est vrai dans l'histoire de Conan Doyle mais faux dans la réalité. Or cela, la thèse naïve proposée précédemment est incapable de le faire. Il s'agit donc l'amender en une nouvelle version qui en conserve l'esprit. À cette fin, Lewis introduit un critère pragmatique lié aux énonciations fictionnelles : une fiction est un *acte* particulier de raconter une histoire. Une fiction n'est donc pas seulement un texte, une inscription. L'œuvre, la fiction, est donc identifiée à un acte, à une performance liée à un acte de communication, plutôt qu'à une seule inscription. Lewis donne alors, en première analyse, le critère de vérité dans la fiction suivant :

> [ANALYSE 0 :] « Dans la fiction f, Φ » est vraie si et seulement si Φ est vraie dans tout monde où f est racontée comme un fait connu plutôt que comme une fiction. (LEWIS, 2012, p. 44)

Ainsi l'acte de raconter l'histoire se produit certes dans le monde actuel, mais il doit être vrai dans des mondes possibles tout en étant faux ici, ce qui justifierait la différence entre des mondes réels et fictionnels où les histoires seraient indiscernables. Par l'entremise de cette précaution, Lewis parvient donc à éviter l'objection de Kripke.

> Le Holmes de la vie réelle aurait eu un Watson de la vie réelle qui aurait raconté des histoires vraies à propos des aventures dont il aurait été le témoin. Mais même si son autobiographie correspondait mot à mot à la fiction de Conan Doyle, ce ne serait pas les mêmes histoires, pas plus que le Quichotte de Cervantès n'est la même histoire que celui de Ménard. Par conséquent, notre monde ne serait toujours pas celui où les histoires de Holmes – les mêmes histoires de Holmes que Conan Doyle raconte comme fictions – seraient racontées comme un fait connu. D'un autre côté, tout monde où l'histoire est racontée comme un fait connu plutôt que comme une fiction doit se trouver parmi les mondes où l'intrigue de l'histoire est actualisée. Autrement son actualisation ne pourrait pas être connue, et nous ne pourrions rien en dire de vrai. (LEWIS, 2012, p. 43)

Mais en quoi consiste précisément l'acte de raconter une histoire ? Selon Lewis, raconter l'histoire f consiste à faire semblant que f est le cas. L'auteur prétend parler de personnages et des

[31] Kripke reconnaît là qu'il contredit des propos antérieurs : « Je ne pourrais donc plus écrire, comme jadis, que : 'Holmes n'existe pas, mais dans d'autres états du monde, il aurait pu exister.' [...]. L'assertion citée donne l'impression erronée qu'un nom fictif nomme un individu possible-mais-non-réel » (KRIPKE, 1982, p. 147).

actions et relations qu'ils entretiennent. Mais, s'agissant d'une fiction, quoique l'auteur prétende le faire, il ne le fait pas réellement. Cela ne veut évidemment pas dire que l'auteur ait l'intention de tromper quiconque. L'auteur *feint* d'accomplir des actes illocutoires plutôt qu'il ne les réalise réellement. Le discours fictionnel est donc un mode particulier d'acte de langage, une feinte qui se donne en tant que telle, c'est-à-dire sans volonté de leurrer son destinataire. L'auteur joue un faux rôle, consistant à raconter des histoires *comme si* celles-ci étaient vraies. Par exemple, dans le cas homodiégétique des récits de Conan Doyle, ce dernier prétend être Watson, le narrateur des histoires d'Holmes, personnage qui se veut être le témoin scrupuleux des aventures de Holmes et qui s'en fait le biographe [32]. Et dans la situation d'un récit à la troisième personne, l'auteur fait de même, en prétendant être le témoin d'événements qu'il relate.

En outre, en comparaison de celui de Lewis, le critère de fictionnalité proposé par Searle semble moins convaincant. Rappelons que ce dernier repose seulement sur l'idée que l'auteur feint des actes illocutoires, mais qu'il ne les exerce pas réellement. Or, lorsqu'un comédien clame par exemple « sortez ! », il ne fait pas semblant de le faire, il le fait *réellement*. Et lorsqu'un personnage de la *Comédie Humaine* fait une promesse à un autre, ce n'est pas Balzac qui feint de promettre, c'est bien le personnage qui promet. L'explication de Lewis permet de rendre compte de la différence externe/interne par laquelle ces actes sont feints ou réalisés. En outre, elle évite en particulier les problèmes qui apparaissent dans le cas où le narrateur fait partie des histoires, comme dans les récits homodiégétiques. Là, il n'est pas besoin de postuler de doubles comme ceux que Searle attribue à Conan Doyle qui « ne se borne pas à feindre de faire des assertions, mais feint d'être John Watson... en train de faire des assertions » (SEARLE, 1982, p. 112). Ce qui est feint, ce ne sont pas les actes de langage produits par les textes, mais c'est notre participation à ces actes. Comme l'a constaté Marion Renauld (RENAULD, 2014), entre la théorie pragmatique de la feinte encouragée par Searle et la théorie de Lewis, le caractère fictionnel transite des actes de l'auteur aux intentions des récepteurs. La sémantique des mondes possibles permet ainsi de conférer un sens aux actes illocutoires qui sont vrais mais déportés dans le monde de la fiction plutôt que feints et actuels.

Ancrer ainsi la fiction sur des mondes possibles via une accessibilité par l'imaginaire nécessite alors de déterminer les mondes possibles de la fiction. Pour cela il faut évidemment restreindre les mondes qui nous intéressent à ceux où les événements sont racontés comme des faits plutôt que comme des situations fictives.

Incomplétude des mondes fictionnels

Mais l'analyse 0 présente encore des inconvénients. John Heintz, dont Lewis affirme s'inspirer en partie, souligne que :

> A fictional world provides the answer to many questions about a character, but it does not specify his characteristics and relations in every detail. There are gaps. And any semantics for a fictional world must recognize the gaps. The assumption we all make about the real

[32] Soulignons au passage une difficulté dans le cas de fictions à plusieurs narrateurs, comme c'est par exemple le cas de *Rashomon* d'Akira Kurosawa, où les événements sont relatés de manière contradictoire selon les points de vue des personnages.

> world is that there are no gaps - perhaps that is one of two formal differences between fiction and reality. (HEINTZ, 1979, p. 93)

Constatons en effet que les fictions sont incomplètes. On pourrait objecter que dans notre propre monde, bien des connaissances nous échappent aussi. Par exemple, je ne sais pas si le nombre de feuilles actuellement présentes sur le houppier du bouleau et qui bruissent devant ma fenêtre est pair ou impair. Mais il n'est cependant pas impossible que je puisse y répondre. Je pourrais néanmoins parvenir à dénombrer les feuilles par l'exercice d'un recensement minutieux. De fait, comme le souligne Kant dans la *Critique de la raison Pure*,

> [...] toute chose, quant à sa possibilité, est soumise encore au principe de *détermination complète*, suivant lequel, de tous les prédicats possibles des choses, en tant qu'ils sont comparés à leur contraire, un seul doit lui convenir. Cette proposition : *toute chose existante est complètement déterminée*, signifie que non seulement de chaque couple de prédicats contradictoires donnés, mais aussi de tous les prédicats possibles, il y en a toujours un qui lui convient. (KANT, 2012, p. 415)

Si cela semble vrai pour la réalité, en revanche, il n'y a rien de tel pour les mondes fictionnels. Leur incomplétude n'est pas seulement épistémique, elle l'est également par « nature ». Dans une fiction, une infinité de caractéristiques restent inévitablement indéterminées, mais surtout, indéterminables. Don Quichotte a-t-il lu Homère ? Y a-t-il plus que vingt portes dans la Maison Usher ? Le grand-père de Sherlock Holmes était-il daltonien ? Nous n'avons aucun moyen de répondre. Et contrairement aux mondes possibles, toujours complets, ce déficit de connaissance n'est pas seulement dû à une information qui viendrait à manquer et qui pourrait éventuellement être pourvue ultérieurement.

Hector-Neri Castañeda considère que l'indétermination est justement le principal critère de démarcation entre les individus réels et les individus fictionnels :

> Les objets fictionnels [...] sont incomplets. Un objet fictionnel est ce qui est créé comme tel, et qui ne peut avoir d'autres propriétés que celles avec lesquelles il fut créé ou recréé. Or, dans la plupart des histoires, on suppose implicitement que les êtres humains dont traite l'histoire sont pleinement humains. [...] Mais, il y a une indétermination des propriétés : aucune nuance de couleur particulière, par exemple, n'est donnée ; pas de masse et de hauteur exactes, pas de longueur précise d'orteils, de pouces ou d'intestins. Pourtant, [...], nous devons comprendre qu'un héros de fiction n'est pas simplement [...] un véritable héros privé de propriétés ou de caractéristiques spécifiques et particularisées. Les objets de fiction demeurent totalement irréels, résidant dans un domaine séparé. La différence entre le réel et le fictif est abyssale [33]. (CASTAÑEDA, 1979, p. 35)

Selon Castañeda, l'indétermination est donc logée au creux de la fiction et surtout, elle doit le rester.

Remarquons néanmoins qu'il serait bien difficile de déterminer ce que serait au juste un individu incomplet, et qu'il est d'emblée assez étrange de prendre cette caractéristique pour une

[33] Ma traduction

preuve de l'existence d'entités fictionnelles. Robert Howell souligne que, contrairement à ce que prétendent les meinongiens, les personnages ne sont généralement pas imaginés en tant que tels. Car de deux choses l'une, soit l'on considère qu'Anna Karénine est un individu incomplet, et dans ce cas on se place d'emblée dans un récit de type fantastique parce qu'on s'écarte de la réalité dans laquelle les individus répondent bien au principe de détermination, soit on admet au contraire qu'Anna Karénine est bien un individu complet :

> Notons immédiatement qu'il existe au moins deux sortes importantes d'incomplétude [...]. D'un côté, s'il avait écrit une fiction fantastique et non une histoire réaliste, Tolstoï aurait pu préciser qu'Anna Karénine, comme elle existe avec les autres personnages de la situation fictive qu'il a créée, est explicitement dépourvue à la fois de la propriété d'être droitière et à la fois de la propriété de ne pas être droitière. Dans un tel récit fantastique, Anna Karénine serait ce que nous pourrions appeler un objet radicalement incomplet, une chose avec un trou dans sa liste de propriétés. Mais, d'un autre côté, le personnage réaliste d'Anna créé par Tolstoï ne s'avère pas radicalement incomplet. Bien au contraire, l'Anna créée par Tolstoï n'est pas radicalement incomplète, en ce sens que Tolstoï - à ma connaissance, ne nous a pas dit si, par exemple, Anna est ou non droitière. [34]. (HOWELL, 1979, p. 134)

Peut-on alors encore considérer qu'un opérateur fictionnel donne accès à un monde possible exclusif ? Selon Lewis, il est incontestable que la présence de nombreux éléments indéterminés dans une fiction a une importante conséquence sur l'accessibilité aux mondes possibles : plutôt qu'à un seul monde, une fiction correspond à tout un ensemble de mondes possibles. Contrairement à ce que pourrait laisser entendre une expression comme « *le* monde de Sherlock Holmes », l'opérateur de fictionnalité ne donne pas accès à un unique monde possible, mais bien à toute une multiplicité de mondes. Autrement dit, un texte conduit à stipuler une infinité de mondes différents qui, d'une part, satisfont tous aux descriptions auctoriales et qui, d'autre part, divergent entre eux quant aux propriétés qui ne sont pas déterminées par le texte. Par exemple, le même texte de Conan Doyle donne accès aux mondes possibles dans lesquels Sherlock Holmes possède toutes les propriétés attribuées par Conan Doyle dans les romans dont il est l'auteur plus celle d'avoir un grand-père daltonien, mais également aux mondes possibles dans lesquels Sherlock Holmes possède toutes les propriétés attribuées par Conan Doyle dans les romans dont il est l'auteur plus celle d'avoir un grand-père qui n'est pas daltonien. L'usage des mondes possibles n'est donc pas contradictoire à l'incomplétude des mondes fictionnels, mais justement cohérent avec le principe de sémiosis illimitée, c'est-à-dire avec l'interprétation multiples des textes. Umberto Eco constate ainsi que « si le texte est une machine paresseuse qui exige du lecteur un travail coopératif acharné pour remplir les espaces de non-dit ou de déjà-dit restés en blanc, alors le texte n'est pas autre chose qu'une machine présuppositionnelle » (ECO, 1979, p. 27).

En cela, la théorie de Lewis semble rejoindre, en somme, le constat de Nelson Goodman d'une multiplicité d'interprétations. Mais ce n'est vrai qu'à la différence près que Goodman refuse que ces interprétations soient exhaustives à la version de monde construite par l'œuvre. La théorie de Goodman stipule qu'une œuvre survit aux versions de mondes, elle n'est pas déployée à travers des doubles mondains. La multiplicité des interprétations est donc toujours ancrée à une unique version de monde, celle produite par l'inscription. Selon Goodman, certaines interprétations d'œuvre

[34] Ma traduction

sont correctes, d'autres non, mais elles sont toujours univoques, relatives à une seule et même version de monde. On voit donc l'avantage de la théorie intentionnelle de Lewis. Non seulement elle assure que nous puissions attribuer la vérité aux énoncés qui proviennent directement du texte, mais elle permet aussi de rendre compte de toutes les inférences qui en découlent. Goodman insiste pourtant sur l'idée que nous *bâtissons* des mondes, mais de quoi parle-t-il au juste si de telles choses n'ont pas d'ancrage sur nos raisonnements logiques les plus élémentaires ? A quoi donc servirait un langage s'il ne permettait pas d'extrapoler des conséquences ? Autrement dit, qu'est-ce qu'une version de monde si ce n'est pas un monde possible ? Le monde de Balzac n'est pas un ensemble d'inscriptions produisant une seule version de monde différemment interprétée. Il s'agit d'un univers au sens où l'on y rencontre des personnes de chair et d'os. Et contrairement à ce que moque Goodman, il n'est point besoin de résider à proximité de Disneyland pour affirmer que de tels mondes puissent être projetés par nos intentions lorsqu'elles sont stimulées par des textes ou des images. Il y a donc des mondes irréels, qui sont ceux que les récepteurs construisent à partir des œuvres.

Interpréter *ceteris paribus*

Bien sûr, si nous imaginons une infinité de mondes possibles pour chaque interprétation, il n'en demeure pas moins, que toute interprétation ne vaille pas, et qu'on doit alors expliquer comment contingenter la multiplication des mondes qui peuvent être tirés du texte. Or, puisque les textes ne peuvent pas tout dire, les interprétations que nous en faisons sont susceptibles de connaître des écarts plus ou moins grands avec le monde réel. Ainsi, du fait qu'elle ne s'intéresse pas à l'arrière-plan, Lewis constate que l'analyse n°0 risque de prendre en charge trop de mondes possibles. Bien que cela ne soit pas contre-indiqué dans les textes de Conan Doyle, il ne serait toutefois pas pertinent de considérer les mondes possibles dans lesquels Sherlock Holmes respirerait avec trois narines et poursuivrait des gnomes pourpres. Il faut donc s'accorder sur l'arrière plan fictionnel en autorisant ou en interdisant certaines interprétations. Par exemple, lorsqu'on consulte une carte de Londres, on peut se rendre compte que Baker Street est plus proche de Paddington Station que de Waterloo Station. Mais la carte de Londres ne fait pas partie des histoires d'Holmes et en outre, Conan Doyle ne mentionne pas cette caractéristique dans ses récits. En conséquence, les histoires de Conan Doyle déterminent autant de mondes possibles dans lesquels il est le cas que Baker Street est plus proche de Paddington Station que de Waterloo Station, que de mondes possibles dans lesquels cela n'est pas vrai. Or, entre les deux options, il est certainement préférable de considérer que le Londres d'Holmes ressemble le plus possible au Londres réel.

Notons qu'on pourrait toutefois être suspicieux à l'égard des lieux que désignent, dans le texte de Conan Doyle, des noms comme « Londres », « Baker Street », « Paddington Station » ou « Waterloo station ». S'agit-il de références aux lieux réels, ou s'agit-il seulement de créations imaginaires ? Pour certains auteurs, tout nom inscrit dans le contexte d'une fiction devient inexorablement un terme fictionnel. Le nom « Londres » cité dans les écrits de Conan Doyle est alors purement fictionnel et ne saurait donc être identifié au Londres réel. Gérard Genette admet en effet que les éléments fictifs d'un texte de fiction sont inextricablement mêlés à des éléments factuels, voire à des éléments indéterminés. Il devient alors difficile de considérer le caractère fictionnel d'un passage. Genette en conclut que ce qui détermine le caractère fictionnel d'un texte n'est pas local mais global. La Russie de *La Guerre et la Paix* n'est pas celle qui a été véritablement enva-

hie par les troupes napoléoniennes. Si la fiction ressemble à un patchwork de parties composites, factuelles et fictives, le tout demeure quant à lui d'un ordre radicalement fictionnel :

> Mais il faut au moins garder à l'esprit que le "discours de fiction" est en fait un patchwork, ou un amalgame plus ou moins homogénéisé d'éléments hétéroclites empruntés pour la plupart à la réalité. Comme le lion n'est guère, selon Valéry, que du mouton digéré, la fiction n'est guère que du réel fictionalisé, et la définition de son discours en termes illocutoires ne peut être que fluctuante, ou globale et synthétique : ses assertions ne sont clairement pas toutes également feintes, et aucune d'elles peut-être, ne l'est rigoureusement et intégralement – pas plus qu'une sirène ou un centaure n'est intégralement un être imaginaire. Il en est sans doute de même de la fiction comme discours que de la fiction comme entité, ou comme image : le tout y est plus fictif que les parties. (GENETTE, 2004, pp. 136–137)

Cependant, quand nous lisons que « Sherlock Holmes vit à Londres », nous identifions certainement une ville particulière et bien réelle. Nous nous attendons d'ailleurs à y trouver certains éléments, dont des bâtiments singuliers comme Big Ben, et des habitants respectant certains us et coutumes, s'alimentant selon un type particulier de régime et s'habillant selon une certaine mode vestimentaire, etc. Nous comprenons les agissements des personnages en fonction de ces attentes. Nous n'imaginons pas que Sherlock Holmes puisse sortir de chez lui entièrement nu pour s'aventurer dans une forêt tropicale résonnant du chant de perroquets multicolores. Nous nous attendons plutôt à ce qu'il soit revêtu d'un vêtement suffisamment imperméable pour braver l'humidité du *fog*, et qu'il traverse des rues pleines du brouhaha des passants et de la suie des machines à vapeur. Bien sûr, il arrive que nos attentes soient déçues, mais nous corrigeons alors, au fur et à mesure de notre lecture, en substituant aux éléments réels, ceux qui sont affirmés ne pas correspondre au texte de l'auteur.

Lorsque le lecteur doit effectuer des choix parmi des interprétations concurrentes et non contradictoires au texte, un principe de sélection doit alors être avancé afin de réduire les écarts imaginatifs qui seraient par trop fantaisistes. L'usage d'un tel principe est parfois jugé illégitime. Mais bien que certains philosophes admettent le principe de sémiosis illimitée par une forme d'équivalence généralisée des interprétations, ce n'est pas l'avis de Lewis, qui incite à comparer la sélection des mondes possibles à la sélection engagée dans les raisonnements contrefactuels. La vérité de la proposition exprimée par la phrase « si j'avais gratté cette allumette, elle se serait enflammée » demande à ne rien changer du monde actuel, sinon que l'allumette ait été grattée. De même, en fiction, nous n'avons aucun intérêt à modifier en pensée plus de choses que ne nécessite ce qui est explicite dans le récit. Autrement dit, il est pertinent d'introduire, à chaque fois que la fiction n'y contredit pas, des éléments d'arrière-plan qui sont tirés du monde réel. Aussi, Lewis propose-t-il d'amender son analyse, en introduisant une clause de similarité *ceteris paribus* :

> ANALYSE 1 : une phrase de la forme « Dans la fiction f, Φ » est vraie non trivialement si et seulement si un monde où f est racontée comme un fait connu et où Φ est vraie diffère moins de notre monde actuel, toutes choses égales par ailleurs, qu'aucun autre monde où f est racontée comme un fait connu et où Φ n'est pas vraie. Elle est trivialement vraie si et seulement si il n'y a aucun monde possible où f est racontée comme un fait connu. (LEWIS, 2012, p. 47)

Précisons aussi que l'intervention du narrateur dans la relation d'accessibilité a encore un autre avantage, à savoir celui de donner une explication aux fictions homodiégétiques, autrement dit, aux fictions où le narrateur, faisant partie de l'histoire, est lui-même fictionnel.

Cette nouvelle analyse prend donc en compte le fait que nos inférences ne sont pas toutes logiques, mais aussi conceptuelles, basée sur l'expérience acquise. À côté de la fermeture logique des fictions intervient ainsi ce deuxième principe que Marie-Laure Ryan appelle le « principe d'écart minimal (*minimal departure*) ». Ce principe stipule qu'en dehors des altérations explicitement décrites dans le texte ou de celles qui en dérivent logiquement, le monde fictionnel doit s'écarter le moins possible du monde réel. Quand Conan Doyle affirme que Holmes est un détective, le principe d'écart minimal implique qu'il est un homme, qu'il respire avec des poumons et non des branchies, que son sang véhicule des hématies plutôt que de minuscules billes de verre, et que ces hématies transportent des atomes d'oxygène plutôt que de l'uranium, ces atomes étant eux-mêmes constitués de particules élémentaires comme les bosons et les fermions, bien que Conan Doyle ignorât sans doute l'existence de telles particules.

Les récits sont d'ailleurs extrêmement sensibles à l'absorption du réel. Dans *Lector In Fabula* (ECO, 1979, pp. 15–17), Umberto Eco attire l'attention sur le fait qu'une histoire est comprise par certaines circonstances d'énonciations, mais que ces circonstances doivent aussi être complétées par le *contexte du texte*, c'est-à-dire les significations virtuelles que porte le texte. Le mot « lion » produit un contexte de signification, même en dehors de toute circonstance d'énonciation, par exemple parce qu'il implique certains lieux et en exclut d'autres, du moins dans les données culturelles socialement acceptées, comme la jungle, le zoo, ou le cirque. Le mot isolé est un « système d'instructions orientées au texte ». « Jungle, cirque et zoo sont des circonstances sémiotisées [...] où le lexème "lion" peut être produit » écrit Eco (ECO, 1979, p. 18). Ces différents contextes connotent eux-mêmes d'autres lexèmes comme « liberté » ou « captivité ». C'est donc que le mot développe déjà, à lui seul et en dehors de toute circonstance d'énonciation, des textes possibles. Lorsque le mot est effectivement associé, dans le texte, à l'un des termes de son contexte, c'est-à-dire lorsque cette occurrence se réalise, Eco le désigne par le terme de « co-texte ».

Cependant, importer l'ensemble de nos connaissances à l'intérieur de la fiction implique des conséquences qui peuvent paraître inattendues, voire fâcheuses. S'il est vrai que 1756 flèches furent lancées lors de la bataille d'Hasting, cela implique l'existence du même événement dans les mondes fictionnels de Sherlock Holmes. Or les histoires d'Holmes n'ont vraiment rien à voir avec le nombre de flèches tirées à la bataille d'Hasting. Serait-ce alors que la réalité contamine par trop la fiction ? Le critère d'écart minimal semble offrir à la fiction une béance sur le réel qui s'y immisce de manière illimitée et sans doute un peu forcée.

On peut toutefois répondre à cette objection. Si nous retrouvions un passage mentionnant la bataille d'Hasting - Watson pourrait par exemple en faire le compte rendu à Sherlock Holmes durant leurs heures de loisir, cela confirmerait évidemment l'importance de son importation dans le monde fictionnel. Or, tant que cela n'est pas le cas, le lecteur n'a pas à s'intéresser à elle. Mais il faut souligner qu'il y a une manifeste différence, pour un lecteur, entre le fait de refuser de se poser la question de cette bataille et le fait de ne pas se poser cette question. L'importation massive du réel dans les creux de la fiction ne semble pas être une objection contre le principe d'écart minimal, mais tout au plus la constatation que le monde est quelque chose de vaste, et qu'un monde fictionnel n'en est qu'une petite partie modifiée en pensée. C'est un fait que nous

ne remplissions pas complètement les mondes fictionnels à leur lecture, c'est-à-dire quand nous les reconstruisons. Mais nous ne le faisons pas plus lorsque nous nous représentons une situation factuelle. En revanche, si nous sommes amenés à une lecture critique, il est justement important de pouvoir raisonner sur la base de ce remplissage. Si l'on souhaite par exemple expliquer le comportement de tel ou tel personnage de Conan Doyle, on doit pouvoir faire appel aux structures idéologiques dominantes de l'Angleterre victorienne que Conan Doyle a consciemment ou non intégré à son récit. Les récits d'Agatha Christie ne prescrivent certes pas d'imaginer la défaite de Napoléon à Waterloo, mais les présupposés politico-idéologiques de la société dans laquelle agissent les personnages des romans d'Agatha Christie présupposent bien cette défaite. Ainsi, le principe d'écart minimal n'est pas une contrainte de lecture postulant que le lecteur doit tout imaginer du monde fictionnel qui est décrit. Il n'oblige pas à imaginer toutes les façons dont le monde fictionnel ressemble au monde réel, mais en revanche, il rend possible des raisonnements qui s'appuient sur cette ressemblance invasive.

Tout ce qui précède repose, comme nous l'avons dit, sur l'hypothèse d'un narrateur fiable, autrement dit, sur l'hypothèse que le narrateur est sincère, qu'il y a une complète adéquation entre ce qu'il dit et ce qui est vrai dans le monde fictionnel où il le dit. Mais il arrive que certains critiques refusent ce principe, et il arrive même que certains récits jouent justement sur son rejet. Le monde fictionnel absorbe alors des mondes possibles qui s'écartent de ceux qu'on obtient en considérant seulement la vérité des propos du narrateur. Dans son étude critique à propos du *Meurtre de Roger Ackroyd* d'Agatha Christie, Pierre Bayard fait l'hypothèse que les affirmations du narrateur dissimulent la vérité afin de disculper le véritable meurtrier. Le meurtre n'aurait tout simplement pas été élucidé et Bayard se propose d'en reprendre l'enquête. En remettant en question la fiabillité de l'auteur, Bayard libère ainsi une grande marge de manœuvre pour la sélection des mondes possibles afférents au récit d'Agatha Christie [35].

Mais il n'est même pas nécessaire de faire intervenir l'hypothèse d'un narrateur non-fiable pour faire émerger des difficultés dans l'analyse précédente. Lewis constate en effet qu'elle met en jeu une relation à l'arrière plan du monde qui est certainement bien trop idéalisée. Si l'on concède, comme nous l'avons fait jusqu'à présent, que les mondes sont construits par des processus d'imitation basés sur la manière dont le monde est, certains problèmes apparaissent immédiatement. D'abord, constatons qu'auteurs et lecteurs de différentes époques ne disposent pas des mêmes ressources, et que ce qui relève de connaissances pour les uns tombera dans le régime des croyances erronées pour les autres. Homère fait intervenir Athéna dans son récit. Or, si elle était certainement assimilée à une entité bien réelle pour un Grec antique, elle ne l'est sûrement pas pour un lecteur contemporain. Il n'est même point besoin de recourir à des expressions dont la dénotation est d'emblée suspecte. L'expression « baleine », n'a pas la même signification pour un lecteur médiéval et un lecteur contemporain informé, puisque pour l'un, il s'agit d'un poisson alors que pour l'autre il s'agit d'un mammifère.

Il arrive en outre que les croyances des auteurs et celles des lecteurs entrent en conflit de manière synchronique. Comme l'a souligné Carl Ganz, la vipère de Russell que mentionne Conan Doyle dans *Une étude en Rouge* n'est en réalité pas un animal constricteur. En conséquence,

[35] A ce sujet, citons aussi le cas du roman de Philippe Doumenc, *Contre-enquête sur la mort d'Emma Bovary*, dans lequel l'auteur élabore une réquelle du roman de Flaubert en partant de l'hypothèse que celui-ci a menti à propos du présumé suicide d'Emma Bovary (DOUMENC, 2007).

l'animal est incapable de produire les mouvements dont Conan Doyle écrit qu'ils sont la cause du décès de la victime de son récit. En suivant l'analyse 1 de Lewis, c'est donc que, soit Holmes s'est trompé et donc que l'énigme n'est pas résolue, soit que la vipère décrite ne correspond pas à celle du monde réel. Comment choisir ? Il n'est jamais dit par Conan Doyle que Holmes est infaillible, et quel homme le serait ? En outre, il n'est pas dit non plus qu'on a affaire à un serpent fantaisiste. D'ailleurs, le fait même d'en spécifier le genre inciterait à imaginer le contraire. Aussi, ce qui relève vraisemblablement d'une erreur a pour effet, selon l'analyse de Lewis, l'échec de Sherlock Holmes. Les mondes possibles où Holmes se trompe diffèrent moins du nôtre que chacun des mondes possibles où Holmes a raison et où la vipère de Russell est un animal constricteur.

Si l'on souhaite alors contingenter l'importance des faits peu ou mal connus pour la sélection des mondes possibles qui ont rapport au monde fictionnel donné, il devient alors nécessaire d'améliorer encore l'analyse précédente. Pour ce faire, on peut substituer la croyance à la connaissance. Bien sûr, il ne doit pas s'agir de nos propres croyances, sans quoi l'ensemble des mondes possibles de la fiction serait en constant changement, ce qui impliquerait d'adhérer au pan-fictionalisme :

> C'est étrange à dire, mais l'historien ne serait pas en meilleure position pour savoir ce qui était vrai dans les fictions de sa période que le profane ignorant. Ceci ne peut pas être vrai. Ce qui fut vrai dans la fiction quand elle fut racontée pour la première fois doit rester vrai en elle pour toujours. C'est notre connaissance de ce qui est vrai dans la fiction qui peut croître ou décroître. (LEWIS, 2012, p. 50)

L'arrière-plan fictionnel doit donc être celui des croyances fortes de la communauté d'origine de la fiction.

Constatons toutefois qu'en substituant la croyance à la connaissance, un nouveau problème surgit : certaines croyances peuvent s'avérer inégalement partagées. Cela pousse Lewis à ajouter une dernière sophistication à son analyse en introduisant la notion de croyance prévalente, autrement dit, une croyance dont chacun penserait qu'elle est partagée par la majorité des autres individus de sa communauté à un moment donné. L'aspiration du réel à travers les filtres de la fiction doit ainsi être relativisée aux connaissances et aux croyances des récepteurs. Autrement dit, il est nécessaire de tenir compte des croyances de la communauté d'origine de la fiction à propos des faits qui y sont décrits. Aussi, Lewis propose-t-il un dernier amendement à son analyse en y introduisant le monde des croyances collectives de la communauté d'origine de la fiction.

> ANALYSE 2 : une phrase de la forme « Dans la fiction f, Φ » est vraie non trivialement si et seulement si, si w est un des mondes de la croyance collective de la communauté d'origine de f, alors un monde où f est racontée comme un fait connu et où Φ est vraie, diffère moins du monde w, toutes choses égales par ailleurs, que n'importe quel monde où f est racontée comme un fait connu et où Φ n'est pas vraie. Elle est trivialement vraie s'il n'y a pas de monde possible où f est racontée comme un fait vrai. (LEWIS, 2012, p. 47)

Cette dernière analyse, qui est celle que nous retiendrons, est en particulier compatible avec un anti-réalisme épistémique appuyé sur un réalisme métaphysique. Elle permet de considérer que le monde comme état de choses ne coïncide pas strictement avec les descriptions que nous en donnons. Ainsi, ce n'est pas le monde lui-même que nous importons comme arrière-plan de la fiction, mais nos croyances à son sujet.

Lewis précise enfin que son analyse pourrait être poussée encore un plus loin. Les précédentes analyses font reposer la vérité dans la fiction à la fois sur le contenu explicite du texte et sur l'arrière plan tiré du monde actuel, qu'il soit rapporté par des connaissances (analyse 1), ou par des croyances (analyse 2). Or nous devrions concéder au pan-fictionnaliste que la fiction s'enrichit de la fiction. Les importations ne devraient pas provenir exclusivement des connaissances ou des croyances du monde actuel, mais également des stimulations imaginatives suscitées par d'autres mondes fictionnels. Il arrive qu'une fiction mentionne un dragon, et il est généralement attendu de l'imagination du lecteur que les dragons sont susceptibles de cracher du feu, sans même que cela ne soit stipulé dans le texte. Dans ce cas, l'arrière-plan *ceteris paribus* que le lecteur est invité à importer dans la fiction est, en partie, celui des croyances partagées par la communauté de réception de la fiction, mais en partie aussi, les imaginations partagées par cette communauté, imaginations corrélées aux fictions où il est explicite que les dragons crachent du feu.

La théorie anti-mimétique de Doležel

Parmi les critiques concernant l'usage des mondes possibles en théorie de la fiction, Arthur Danto fait remarquer qu'un lecteur ne devrait éprouver qu'un moindre intérêt pour *Don Quichotte* si le livre de Cervantès ne concernait qu'un personnage non actualisé et vivant dans un ensemble de mondes possibles complètement isolés du nôtre (A. DANTO, 1993, p. 183). Mais nous pouvons aussi considérer que les mondes possibles ne sont que des constructions qui se rapprochent de modélisations heuristiques, des scénarios alternatifs dont l'importance cognitive est difficilement contestable.

Une autre objection qui est généralement retenue contre l'utilisation des mondes possibles en théorie de la fiction est qu'elle est contraire à une théorie mimétique de la représentation. Lubomír Doležel prétend d'ailleurs que sa propre théorie fait usage des mondes possibles, justement afin d'éviter la théorie mimétique :

> Mes travaux de recherche autour d'une sémantique non-mimétique de la fiction ont été guidés par la prise de conscience que les difficultés de la théorie mimétique découlent de ce que les fictions sont attachées exclusivement au monde réel. Toutes les fictions, y compris les plus fantastiques, sont interprétées comme faisant référence à un seul et même "univers de discours", à savoir, le monde actuel. La fonction mimétique est une formule destinée à intégrer les fictions dans le monde réel. La sémantique mimétique est située dans le cadre d'un unique modèle. Une alternative radicale à la mimésis sera une sémantique fictive définie dans le cadre d'un modèle multi-monde. La sémantique mimétique sera remplacée par la *sémantique des mondes possibles de la fiction*. [36] (DOLEŽEL, 1988, p. 481)

En particulier, le refus d'une activité mimétique dans le processus de construction des mondes conduit Doležel à considérer qu'ils sont uniformément fictifs de part en part, et donc qu'il ne peuvent inclure des individus importés.

> En tant que possibles non-actualisés, toutes les entités fictives sont ontologiquement homogènes. Le Napoléon de Tolstoï n'est pas moins fictionnel que son Pierre Bezuchov et le

[36] Ma traduction.

> Londres de Dickens n'est pas plus réel que le Pays des Merveilles de Lewis. Le principe d'homogénéité ontologique est une condition nécessaire à la coexistence et à la compossibilité des particuliers fictionnels. Il explique pourquoi des individus fictifs peuvent interagir et communiquer les uns avec les autres. Un point de vue naïf qui représente les individus fictifs comme un mélange de 'vrais' personnages et de personnages 'purement fictifs' est ici explicitement réfuté [37]. (DOLEŽEL, 1988, p. 483)

Un peu plus loin dans le même passage, Doležel admet toutefois que le monde actuel participe à la construction des mondes possibles :

> Le monde actuel participe à la formation des mondes fictionnels en fournissant des modèles de sa structure (y compris l'expérience de l'auteur), en ancrant l'histoire fictionnelle à un événement historique, en transmettant des "faits bruts" ou des "réalèmes" [38] culturels. [...], etc. (DOLEŽEL, 1988, p. 485)

De fait, nous avons vu que les mondes possibles permettaient de conférer la vérité aux assertions auctoriales. Ce qui n'est pas dit dans le texte, et qui n'est pas déterminable de manière logique à partir des propositions atomiques l'est toutefois par imitation du monde réel. Les théories des mondes possibles ne sont donc pas des théories complètement anti-mimétiques : elles ne sont certes pas mimétiques en ce qu'elle distribuent différemment les valeurs de vérité des propositions atomiques, mais en revanche, elles le sont par l'action du principe d'écart minimal, qui est bien un principe de similarité comparative. En outre, soulignons encore une fois que ce principe n'est pas seulement réservé au monde matériel, mais aussi à nos croyances, et surtout à nos pratiques culturelles. Pour reprendre encore une fois cet exemple, l'occurrence du mot « dragon » dans un texte implique le co-texte « cracher du feu » et « voler », même lorsque rien dans le monde actuel ne correspond à un dragon, et même si ces possibilités ne sont pas directement évoquées dans le texte considéré.

Comme le souligne Doležel, (DOLEŽEL, 1998, p. 202), les phénomènes que représentent l'intertextualité et les productions successives de mondes fictionnels confirment le besoin d'une sémantique à deux branches qui puisse rendre compte des œuvres de fiction à la fois sur le niveau intensionnel et sur un niveau extensionnel des inscriptions matérielles. Doležel englobe ces deux lignes sous le nom de *transduction*. Ce concept ne serait qu'une excroissance du phénomène beaucoup plus large de communication. Ainsi, il pourrait être présenté, dans un premier temps, selon le schéma classique de communication de la figure 3.3.

Mais Doležel constate que ce premier schéma possède deux défauts. Le premier est qu'il modélise l'acte de lecture comme un déchiffrement passif du texte. Le deuxième, c'est évidemment la place manquante du monde fictionnel qui est généré par cet acte. En conséquence, Doležel corrige le modèle en proposant un second schéma (représenté sur la figure 3.4) qui fait apparaître

[37] Ma traduction

[38] La notion provient d'Itamar Even-Zohar (EVEN-ZOHAR, 1985). L'auteur définit d'abord la notion de *répertoire* comme ensemble d'éléments gouvernés par des relations de nature systémique, puis définit les réalèmes comme des répertoires culturels de réalités : « [...] les "éléments de la réalité" (tels qu'êtres humains et phénomènes naturels, voix et meubles, gestes et visages), lors même qu'ils sont bien "là" dans le monde extérieur, constituent dans toute expression verbale qui s'y réfère des éléments du répertoire culturel, [...] des "réalèmes" » (EVEN-ZOHAR, 1985, p. 111).



FIGURE 3.3 : Schéma de communication

l'interaction entre l'auteur qui construit un monde en écrivant un texte, et le lecteur qui reconstruit ce monde à l'aide du texte. Ainsi, c'est à la fois l'auteur et le lecteur qui performent un acte de communication. Ces actes ne sont pas équivalents, mais complémentaires.

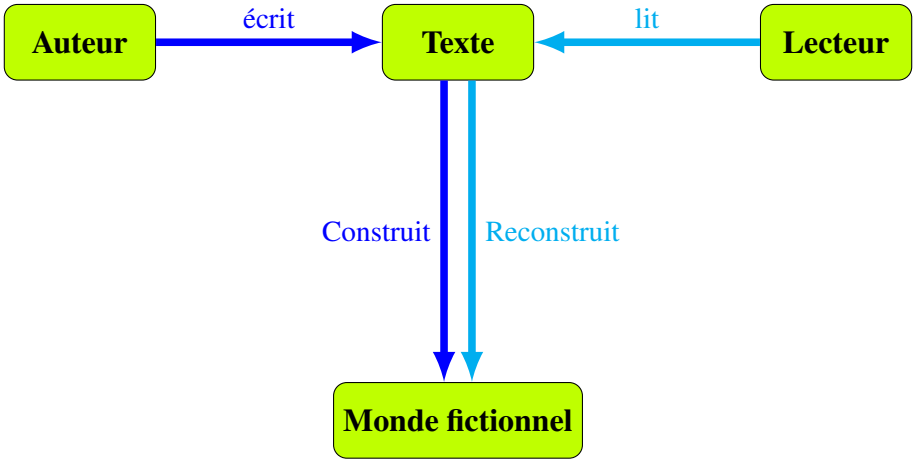


FIGURE 3.4 : Schéma de Doležel

Doležel observe qu'on peut encore améliorer le schéma en soulignant l'aspect transductif de la production de mondes, en reliant différents actes de communication entre eux selon la nouvelle figure 3.5.

Pour ma part, je propose d'intégrer à ce schéma la notion de mondes possibles selon la version suivante, qui inclut un niveau culturel en définissant les mondes fictionnels comme des ensembles de mondes possibles relatifs à un même texte comme représenté sur la figure 3.6.

3.3.4 La théorie métaleptique des fictions

Avant d'en finir avec cette section, je souhaite indiquer encore une proposition en faveur d'une théorie métaleptique des fictions.

C'est que la modélisation des fictions par des ensembles de mondes possibles possède un sérieux avantage sur les théories pan-réalistes et pan-fictionnalistes. En effet, elle permet de justifier le relief de fictionnalité. Dans les théories néo-meinongienne, un personnage est un objet inexistant attaché à des propriétés considérées comme des universaux. On est alors contraint à n'admettre qu'un seul niveau de démarcation entre le réel, exemplifié, et le fictionnel, encodé. Mais la fiction nécessite sans doute d'être vue comme non comme une structure plane : elle exige de constater des degrés différents de profondeur fictionnelle.

Cette notion de profondeur est sans doute ce qui motive Wolfgang Iser à abandonner la tradi-

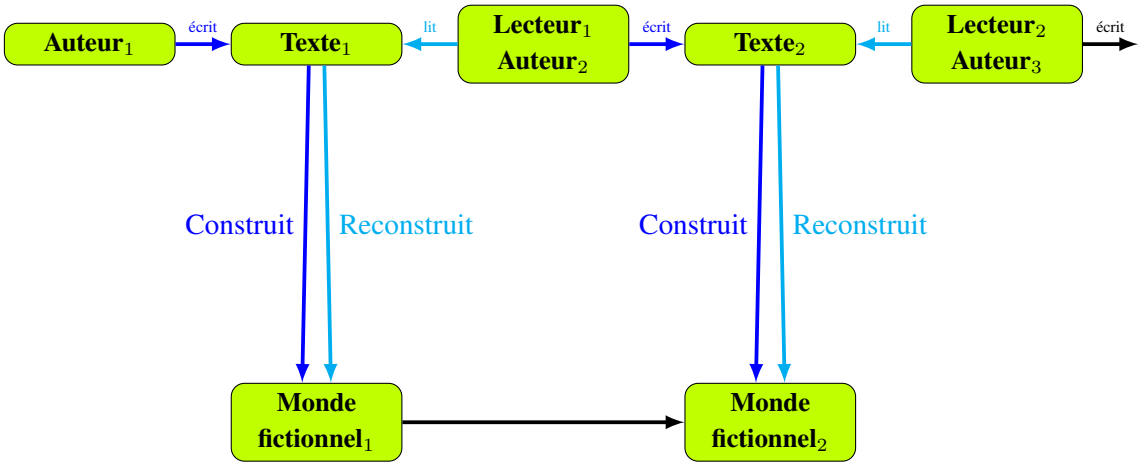


FIGURE 3.5 : Schéma de Doležel

tionnelle dichotomie réalité *vs.* fiction au profit du triplet réalité, fiction et imaginaire :

> The literary text is a mixture of reality and fictions, and as such it brings about an interaction between the given and the imagined. Because this interaction produces far more than just a contrast between the two, we might do better to discard the old opposition of fiction and reality altogether, and to replace this duality with a triad: the real, the fictive, and what we shall henceforth call the imaginary. It is out of this triad that the text arises: Just as the text cannot be confined to those of its elements which are taken from referential reality, so it cannot be pinned down to its fictional features. For these fictional features do not constitute an end or an entity into themselves. Rather, they provide the medium through which a third element emerges. This is the element I have called the imaginary, whose salient features will evolve as we proceed. (ISER, 1993, p. 1)

Mais l'usage des mondes possibles a l'avantage de ne pas céder à cette distinction au demeurant peu claire entre fiction et imaginaire, et permet de rendre compte du caractère relatif de la fictionnalité.

De fait, un cas pose réellement problème aux conceptions néo-meinongiennes et créationnistes des *ficta*. Il s'agit de la situation dans laquelle la fiction engage un jeu de croyance-feinte dans laquelle on n'a pas affaire véritablement à deux personnages, mais à un seul qui croit à l'existence d'un autre individu qui en réalité, s'avère être une illusion. Dans William Wilson, le narrateur hétéro-diégétique raconte qu'il est depuis toujours confronté à un rival homonyme et indiscernable. A la fin de la nouvelle, le lecteur apprend qu'il s'agissait d'une pure illusion. Il n'y avait qu'un seul William Wilson. Mais l'apparition du double devrait impliquer, du point de vue du réaliste des *ficta*, la création d'un objet fictionnel. En outre, la narration permet de distinguer cet objet fictionnel de William Wilson$_1$ lui-même, parce que la focalisation narrative exercée sur Wilson$_1$ permet toujours de le distinguer de Wilson$_2$ (par exemple par la propriété d'être victime de son Doppelgänger). La fusion des deux qu'occasionne la fin du récit devrait donc annuler le deuxième objet fictionnel, ce qui est impossible d'un point de vue meinongien parce que les objets fictionnels ne sont ni créés, ni détruits, mais seulement sélectionnés. Et dans le cas du créationnisme, le

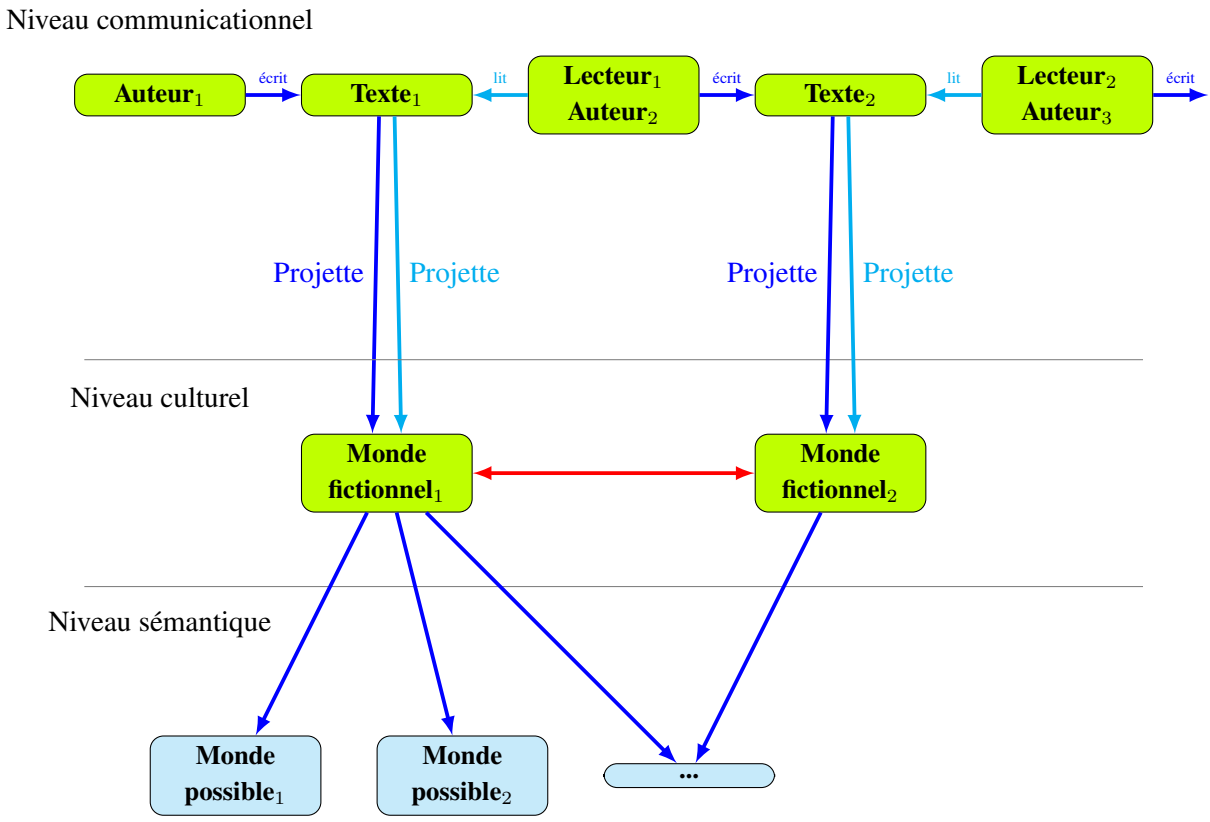


FIGURE 3.6 : ma proposition

rabattement final de Wilson$_2$ sur Wilson$_1$ pose également des problèmes : a-t-on affaire à un ou à deux personnages ?

Le caractère étrange qui ressort des discussions sur les objets indiscernables semble témoigner du caractère très irrationnel des objets abstraits, dont les caractéristiques tendent vite à l'absurde quand on l'applique à des cas littéraires dont l'enjeu ressortit aux intentions des personnages. Le cas de fictions post-modernes dans lesquelles l'auteur affirme mentir est encore plus complexe et inexplicable pour le réaliste, parce que les fictions ne sont pas considérées comme des imitations, elles ne signifient rien d'autre qu'elles-mêmes.

Il semble d'ailleurs que le problème mis en exergue ici provienne de ce que les théories réalistes ne peuvent pas empêcher, et parfois souscrivent à, un aplatissement des niveaux de fictionnalité qui structurent un récit. En revanche, les stratégies qui recourent aux jeux de croyance-feinte sans engagement envers des entités fictionnelles ont les moyens de discerner des fictions à l'intérieur de fiction, par exemple en relayant les accès possibles à d'autres mondes possibles.

Lorsque le texte littéraire fait mention des croyances fausses, des illusions, des mensonges et de manière générale des intentions des personnages, le réaliste est pris d'emblée dans des difficultés qui s'avèrent inextricables [39]. En revanche, la théorie sémantique des mondes possibles asso-

[39] A ce sujet, Anthony Everett propose l'objection suivante dans une note de bas de page de *The Nonexistent* : «

ciée à l'opérateur de croyance feinte permet de résoudre ces cas par la dynamique des relations qui peuvent ou non donner accès à certains mondes doxastiquement possibles (voire impossibles) ou à des mondes épistémiquement possibles. Ainsi, on pourra désigner de niveau 1-fictionnel le niveau de ce qui, dans la fiction, doit être considéré comme un fait. Par exemple, dans *Le Double* de Dostoïevski, il est 1-fictionnel que Goliadkine est un fonctionnaire de Saint-Pétersbourg, et dans *A rebours* de Huysmans, il est 1-fictionnel que Des Esseintes fasse une sortie dans Paris. On désignera ensuite de niveau 2-fictionnel ce qui, dans la fiction, doit être considéré comme non factuel. Ainsi, le voyage de Des Esseintes à Londres est-il 2-fictionnel, puisqu'il procède exclusivement de l'imagination du personnage. On peut ainsi expliquer les différents niveaux de profondeur fictionnelle par des relations entre mondes possibles.

3.4 Conclusion

Nous sommes passés, au fil de ce chapitre, d'une théorie mimétique des représentations fictionnelles à une théorie métaleptique basée sur des relations d'accessibilité épistémiques et doxastiques entre mondes possibles construits par des actes illocutoires.

Deux écueils menacent une théorie des fictions : d'une part le pan-réalisme qui admet que tout est réel, y compris notre imagination de licornes broutant sur la face cachée de la lune, et d'autre part, le pan-fictionnalisme qui rejette l'idée d'une réalité objective par delà la fiction.

L'enjeu était de naviguer entre ces deux pôles en proposant de conserver une signification aux propositions fictionnelles sans pour autant céder au pan-réalisme. Nous avons distingué préalablement l'illusoire (toujours faux) de l'irréel, afin d'introduire l'idée d'une signification qui ne serait pas ancrée à l'ontologie. Nous avons vu que la sémantique des mondes possibles pouvait relever ce défi, en désolidarisant la signification de l'ontologie.

En outre, la théorie métaleptique que je propose ne considère pas que le réel soit absorbé dans la fiction (théorie pan-fictionnaliste), ni que la fiction soit un double mimétique du réel (théorie mimétique) mais que la fiction absorbe à la fois nos connaissances et nos croyances à propos de la réalité, mais aussi nos croyances imaginatives importées de fictions antérieures. Les mondes fictionnels se soutiennent les uns les autres par partage d'imagination, et leur degré de fictionnalité n'est pas uniformément le même, mais répond à une profondeur s'échelonnant en degrés dans l'irréel (ce qui, bien sûr, n'implique pas qu'une proposition fictionnelle soit plus ou moins vraie).

Le dernier schéma de la section précédente ouvre sur la question de la discernabilité des mondes possibles en rapport à l'identité du texte. Un texte implique-t-il une multitude d'interprétations ? Faut-il alors admettre que différentes œuvres puissent être générées par un seul et même texte ? Nous le voyons, nous glissons maintenant à la question de l'identité des fictions, que nous affinerons dans le chapitre suivante. Le problème est de déterminer à quel niveau placer ce qu'on appelle l'œuvre textuelle. Doit-on la placer au niveau de l'inscription physique, au niveau culturel, ou au niveau des mondes possibles ?

Consider, for example, "both Holmes and Soames are admired by their colleagues." This seems a perfectly fine thing to say and lacks any forced or punny feel to it. But, since it seems natural to take the property of being admired by one's colleagues to be an intentional property, on Parsons's account Soames should have the extranuclear version of that property while Holmes should merely have its watered-down nuclear correlate. » (A. EVERETT, 2013, p. 175)

Il s'agira, dans les deux chapitre qui suivent, de donner une réponse à cette question en considérant successivement les problèmes qu'occasionnent respectivement l'individuation des fictions et l'individuation en fiction.

Chapitre 4

Le texte et ses doubles

> Copiez, et si en copiant, vous restez vous-même, alors c'est que vous avez quelque chose à dire.
>
> Conseil de Ravel à ses élèves.

L'objet de ce chapitre est de considérer les rapports qui nouent un texte à ses reproductions, et de considérer la place qu'y occupe la notion d'œuvre textuelle. L'enjeu global est ici de rejeter l'idée que l'œuvre est un double ontologique du texte, et de considérer l'œuvre sous son rapport fonctionnel. A cet effet, j'introduis la notion de dupliques, sorte de doubles par similarité fonctionnelle.

Le premier chapitre de l'essai de Nelson Goodman *Langages de l'art* (1968) s'ouvre par une citation attribuée à Virginia Woolf : « L'art n'est pas une copie du monde réel. L'un des deux suffit largement. ». La référence n'est pas, insiste Goodman, une question de ressemblance. Si un signe pictural dénote une chose du monde, ce n'est pas par la médiation d'une imitation de forme, puisque, au fond, en ce qui concerne les signes, « presque tout peut valoir pour presque n'importe quoi d'autre » (GOODMAN, 1976, p. 35). Dans son essai, Goodman centre la question du double en la faisant porter sur le problème de déterminer le juste degré de résistance des œuvres à la duplication. La question du double est ainsi déplacée de celle, traditionnelle, de la représentation mimétique que nous avons commentée dans le chapitre précédent, à celle de la duplication des supports physiques des œuvres. Précisément, Goodman partitionne les arts selon leurs aptitudes respectives à l'identification à travers les copies. D'un côté, il qualifie d'arts *autographiques* ceux dont l'identité offre une résistance nulle, autrement dit, dont l'apparition de toute nouvelle occurrence constitue une contrefaçon de l'original. D'un autre côté, il appelle arts *allographiques* ceux pour lesquels cette résistance est complète, autrement dit, ceux qui regroupent les œuvres dont l'identité persiste à travers leur reproduction, comme les textes littéraires ou les partitions musicales.

Les textes sont des supports pour la pensée. Or, pour un texte donné, ce support peut bien sûr être reproduit en une multitude de copies. Par exemple, en mentionnant *simpliciter* l'*Odyssée*, nous identifions un seul et même texte, indifféremment du fait qu'il s'agisse de l'édition de la Pléiade, de celle de la collection Budé, d'un livre illustré de gravures ou d'une édition de poche. Vous et

moi pouvons probablement situer l'*Odyssée* sur une certaine étagère de nos logements respectifs. Les reproductions d'un texte sont, en termes de support pour la pensée, parfaitement équivalentes à l'original dont elles sont tirées. Notre question, dans ce chapitre, sera donc la suivante : comment un texte préserve-t-il son identité à travers sa reproduction ? Qu'est-ce qu'une copie, en quoi se différencie-t-elle d'une réplique ou d'une duplique ?

Déroulons le plan du chapitre. Dans la première section, nous introduirons la notion de duplique pour qualifier les reproductions d'un texte original. Nous comparerons les dupliques aux répliques et aux copies. Il s'agira ensuite, dans la section suivante, de situer l'œuvre textuelle par rapport au texte. Outre le problème de concilier l'identité numérique à la diversité des supports, il faut également tenir compte de certaines pratiques qui perturbent la présumée unicité du texte. Un même texte peut avoir différentes variantes auctoriales, être diversement traduit, et finalement, recevoir différentes interprétations. L'enjeu de la troisième section sera donc aussi de traiter la question des versions, des traductions et des interprétations, phénomènes qui semblent, là encore, diffracter l'unité du texte en multiplicités éparses. Enfin, dans la quatrième section, nous terminerons ce chapitre par l'examen du problème d'identification des mondes fictionnels que Jorge Luis Borges a popularisé dans sa courte nouvelle intitulée *Pierre Ménard, auteur du Quichotte* (1939). N'en déplaise à la nature allographique que Goodman attribue aux textes, les arguments de Borges tendent à montrer qu'une copie indiscernable d'un texte peut néanmoins relever d'une œuvre numériquement différente de son original.

4.1 Le texte et ses dupliques

4.1.1 Répliques, copies et dupliques

Commençons par considérer de manière générale les doubles matériels comme des répliques physiques d'objets en élaborant quelques distinctions de terminologie qui serviront à analyser plus finement les reproductions de textes. Dans un premier temps, il s'agira de différencier les notions connexes que sont les copies, les répliques et les dupliques.

La notion de copie

Le terme de copie est générique, associé à celui de reproduction. Comme le note Corinne Maitte (MAITTE, 2015), l'étymologie révèle une signification méliorative issue du mot latin *copia* signifiant abondance, sens que l'on retrouve dans l'adjectif « copieux ». C'est que la copie est d'abord l'indice d'une profusion sans restriction. À partir de 1250, la signification du mot change, il perd son sens premier au profit d'un usage servant à désigner, de manière technique, la reproduction d'un manuscrit. Cette acception se généralise ensuite à l'imprimerie, puisqu'à partir du XVIIe siècle, le mot y est employé pour faire référence au texte définitif d'une composition. Et c'est seulement à partir de là que sa signification vient à couvrir aussi celle de contrefaçon. Ce nouvel emploi suggère un aspect imitatif, voire négatif, qui prend peu à peu le dessus sur la notion de reproduction. De nos jours, le terme sert autant à désigner l'acte de reproduction et l'exemplaire issu de cet acte. La législation du droit d'auteur définit la reproduction comme la « fixation matérielle de l'œuvre par tous procédés qui permettent de la communiquer au public d'une manière

indirecte. Elle peut s'effectuer notamment par imprimerie, dessin, gravure, photographie, moulage et tous procédés des arts graphiques et plastiques, enregistrement mécanique, cinématographique ou magnétique » [1].

Comparaisons entre copies et répliques

Nous pouvons noter certaines caractéristiques qui distinguent les copies de termes concurrents. Nous disons généralement d'un artefact x qu'il est une copie d'un autre artefact y si, lors du processus de production de x, intervient l'intention de manifester une similarité avec cette autre chose, y. L'intervention d'une intention dans le mode de production permet de distinguer les copies parfaites des indiscernables. Deux galets d'un rivage pourraient parfaitement se ressembler, et donc constituer des répliques pures, ils ne seront toutefois pas considérés comme des copies, en raison que leur mode de production ne fait par intervenir d'intention.

Notons également que des différences logiques différencient les notions de répliques pures et de copies. Contrairement à la relation de réplique, la relation de copie est irréflexive : nous ne dirions pas d'une chose qu'elle est une copie d'elle-même. En outre, elle est aussi asymétrique : si x est une copie de y, alors y ne peut être une copie de x, l'asymétrie logique de la relation de copie résultant de l'asymétrie temporelle des artefacts. Enfin, il s'agit d'une relation qui n'est généralement pas transitive : un artefact x peut être une copie d'un artefact y et y une copie d'un artefact z sans que x soit une copie de z. De telles différences proviennent, encore une fois, du caractère intentionnel des copies.

Comparaisons entre copies et dupliques

Il vaut la peine de considérer la notion qu'Athanassios Tzouvaras (TZOUVARAS, 1993) recouvre sous les appellations multiples de *copies*, de *replicas*, et de *spare parts*, que j'appellerai pour ma part, *dupliques*. Cherchant à circonscrire les conditions d'identité temporelle d'artefacts composés de parties atomiques substituables, Tzouvaras introduit une relation binaire, l'ajustement $\mathcal{F}$ (*fitness*) et une fonction d'assemblage $\square$ (*assembly operation*) :

$\mathcal{F}$: Pour deux artefacts x et y, $x\mathcal{F}y$ si et seulement si x et y s'ajustent l'un à l'autre et peuvent être assemblés en un nouvel objet.

$\mathcal{A}l$ **:** $x\square y = z$ si et seulement si z est l'assemblage de x, y, quand ces derniers s'ajustent.

Ces deux relations répondent aux conditions suivantes :

(C1) $x\mathcal{F}y \equiv (\exists z)x\square y = z$

(C2) $x\square y = y\square x$

(C3) $x\square y = x'\square y' \rightarrow (x, y) = (x', y')$

[1] Code de la Propriété Intellectuelle du 1er juillet 1992, http://www.cfcopies.com/juridique/textes-de-loi.

(C4) Chaque objet est analysable en termes de nombre fini de parties atomiques.

Remarquons que le premier axiome est discutable parce qu'il implique qu'un objet démonté est encore un artefact. Une montre dont les diverses parties du mécanisme seraient complètement éparpillées aux quatre coins de l'univers serait encore une montre. En revanche, contrairement à l'universalisme méréologique, l'axiome ne prescrit pas de considérer que n'importe quelle somme de parties matérielles constitue un objet. Il s'agit seulement des parties matérielles qui s'ajustent matériellement les unes aux autres dans un assemblage, c'est-à-dire, dont la configuration structurelle permet l'ajustement respectif. Notons que le dernier axiome sert de principe de fondation méréologique, les parties atomiques étant définies comme des parties n'ayant pas de parties propres.

Après avoir introduit ces deux axiomes, Tzouvaras définit la relation de duplique : deux artefacts x et y sont des dupliques respectives si chacun d'eux s'ajuste précisément là où l'autre s'ajuste également, autrement dit, si, dans n'importe quel assemblage de parties qui contient l'un d'eux, l'autre peut le remplacer. La duplique relève ainsi de l'interchangeabilité et non de la similarité parfaite ou quasi-parfaite des répliques. Un double de clés n'a pas nécessité d'instancier exactement les mêmes qualités que son modèle. Par contre, il doit s'ajuster avec la même serrure que son modèle afin de remplir le même usage que lui. Les dupliques ne nécessitent pas de ressemblances qualitatives en dehors d'une mêmeté structurelle garantissant leur interchangeabilité à un emplacement donné d'un assemblage. L'identité temporelle d'un objet à travers le remplacement de ses parties est ainsi assurée par une relation de duplication des parties. Le bateau de Thésée entièrement rénové est le même bateau que celui qui fut baptisé sur le port du Pirée parce que ses parties remplacées sont des dupliques de ses parties originales. Dans cette conception, c'est une certaine notion de double attribuée aux parties propres d'un artefact qui permet d'assurer l'identité de l'assemblage.

Synthèse des différences

Selon les axiomes de Tzouvaras, la relation de duplique est une relation d'équivalence, donc elle n'est pas une variété de la relation de copie, puisque cette dernière n'est pas une relation d'équivalence. Ainsi, produire un double de clé ne conduit pas à une copie mais à une duplique, à un objet jouant un rôle identique dans un assemblage donné. On peut finalement résumer les différences entre les notions de copies, de répliques et de dupliques selon le tableau suivant :

relations	rapport intentionnel à un original	réflexivité	symétrie	transitivité	similarité
répliques pures	contingent	réflexive	symétrique	transitive	parfaite et nécessaire
copies	nécessaire	irréflexive	asymétrique	généralement non transitive	quasi-parfaite et nécessaire
dupliques	nécessaire	réflexive	symétrique	transitive	contingente

Notons en particulier que la notion de duplique est relative à l'égard du contexte : une duplique dépend d'un système plus grand dans laquelle elle prend place pour fonctionner (elle s'y articule). Une copie, quant à elle, s'avère suffisante à l'égard d'elle-même, elle n'a, si je puis dire, de comptes à rendre qu'à l'égard de sa ressemblance qualitative à l'original qu'elle reproduit.

Applicabilité de la relation de duplique aux textes

L'importance de la notion de duplique à l'égard de notre sujet de réflexion, c'est qu'elle est applicable aux textes matériels. Imaginons une situation dans laquelle une page d'un roman aurait été arrachée puis remplacée par une reproduction manuscrite du texte manquant. Nous pouvons bien admettre que la page manuscrite serait une duplique de la page arrachée dans l'assemblage que constitue le roman, en ce qu'elle « s'ajusterait » à son complément pour fonctionner comme la même œuvre textuelle. En outre, nous pouvons imaginer toutes sortes d'amputations, et comme dans le cas du bateau de Thésée, la totalité des parties matérielles du roman pourrait bien s'avérer remplacée, par exemple par des pages manuscrites. On peut même imaginer qu'on relie ensuite dans le même ordre, la totalité des pages arrachées, de manière à reconstituer le roman mutilé [2]. Il se pourrait même que l'intégralité du roman soit recopié à la main, ou photocopié, ou réimprimé, sans qu'il subisse par ailleurs la moindre altération de ses parties matérielles. Dans cette situation, il reste à savoir si nous reconnaîtrions que l'objet obtenu serait ou non une duplique du premier.

Dans ce qui suit, je donnerai une réponse positive à cette question, en décrivant différents types de relations d'ajustement. Mon objectif est donc, lorsqu'on parle de « différents exemplaires de la même œuvre textuelle », d'échanger la notion d'identité entre copies similaires par celle de duplication fonctionnelle. J'assimilerai en effet l'œuvre textuelle à une fonction d'assemblage sous la relation d'ajustement syntaxique.

4.1.2 Caractériser les reproductions pertinentes d'un texte

Nous avons remarqué plus haut que la nature d'un texte semble osciller entre une singularité liée à sa particularité matérielle et une pluralité ressortissant à la diversité numérique de ses supports. Nous aurons donc à éclairer la confusion, dans nos rapports au texte, entre d'une part le texte vu comme un type, une abstraction, autrement dit une collectivité de choses dispersées dans l'espace et le temps, et d'autre part le texte vu comme une occurrence physique concrète (*token*). La notion de duplique devrait nous aider à supprimer le type (comme entité abstraite) au profit des seules occurrences matérielles, c'est en tout cas notre objectif.

La thèse inscriptionnaliste

Si on définit un texte comme un ensemble de marques physiques, il peut alors sembler inutile de faire appel à la notion d'intentionnalité pour déterminer ses conditions d'identification. On recourt ainsi à une caractérisation purement extensionnelle des textes, qu'on désigne habituellement sous l'appellation d'inscriptionnalisme. L'avantage de cette définition est qu'elle permet d'individuer

[2] Bien sûr, il s'agit ici ni plus ni moins de la version du paradoxe du bateau de Thésée imaginée par Hobbes où le bateau original est reconstruit progressivement.

un texte par l'application du principe d'identité des indiscernables. On dira en effet qu'un texte x est identique à un texte y si et seulement si x et y partagent exactement les mêmes inscriptions. Une telle conception de la reproduction d'un texte semble d'emblée plus proche de la notion de réplique que de celles de copie ou de duplique en ce qu'elle fait complètement fi de l'intention de reproduction.

D'un point de vue historique, une telle conception axiomatique fut utilisée par David Hilbert afin d'élaborer son programme de reconstruction formaliste de la géométrie euclidienne. Chez Hilbert, les points, les droites et les plans n'ont pas de signification par eux-mêmes, ils sont exclusivement associés à des caractères alphabétiques, à savoir, respectivement des lettres latines majuscules et minuscules et des lettres minuscules grecques. Les objets géométriques sont ainsi caractérisés de manière purement graphique, sans aucune considération sur la question de leur nature intrinsèque. En conséquence, ces « objets » sont ré-interprétables à volonté, le développement épistémique des caractéristiques du système demeurant entièrement indifférent au choix des interprétations.

L'inscriptionnalisme a également été défendu par Willard Van Orman Quine et par Nelson Goodman qui, dans « Steps Toward a Constructive Nominalism » écrivent que :

> It may naturally be asked how, if we regard the sentences of mathematics merely as strings of marks without meaning, we can account for the fact that mathematicians can proceed with such remarkable agreement as to methods and results. Our answer is that such intelligibility as mathematics possesses derives from the syntactical or metamathematical rules governing those marks. (GOODMAN & QUINE, 1947, p. 111)

Parmi les successeurs de Quine et Goodman, mentionnons encore Israël Scheffler (SCHEFFLER, 1954) et Donald Davidson (DAVIDSON, 1968) qui ont appliqué le traitement inscriptionnaliste aux citations, aux métaphores et aux discours indirects.

Ma position sera la suivante : je vais tenter de suivre la conception inscriptionnaliste en l'appliquant au cas des œuvres textuelles. Pour commencer, ma caractérisation des textes s'alignera donc sur la conception que Quine et Goodman résument en écrivant que « the language elements in the concrete world are rather inscriptions or marks, the shaped objects rather than the shapes » (GOODMAN & QUINE, 1947, p. 106). Je vais considérer ainsi qu'un texte consiste en un enregistrement d'inscriptions physiques, quelle que soit la nature de ces marques et quel que soit son mode d'enregistrement. Aussi, un manuscrit et un enregistrement numérique seront des exemplaires de ce qui sera nommé ici « texte ». La question sera de déterminer à quel moment nous devrons renoncer à la thèse inscriptionnaliste.

Soulignons tout de suite que l'inscriptionnalisme implique une conception extrêmement large de la notion de texte. Car, sous cette définition, on pourra considérer comme des textes les traces d'un pagure anachorète sur le sable humide de la grève normande, les veines d'un marbre de Carrare et les cernes d'un noyer alsacien. Provisoirement, je laisserai de côté la question de savoir si le fait d'être produits par des humains est ou non une caractéristique essentielle des textes, mais nous aurons de toute façon l'occasion d'y revenir amplement au cours du chapitre.

Le Texte et ses doubles numériques

De fait, le terme « texte » se dit d'abord d'une occurrence physique particulière, comme par exemple la lettre que Friedrich Nietzsche adressa le 27 mai 1876 à Peter Gast, qui est un certain morceau de papier présentant des marques disposées d'une certaine manière, colorées d'une certaine teinte, tracées avec une certaine calligraphie, et dont le support peut être maculé d'empreintes provenant d'un certain individu. Autrement dit, un texte est tout d'abord un objet matériel unique, et pour cette raison, parfois convoité par des collectionneurs. Je désignerai cette clause d'originalité du texte par la graphie « Texte », marquée d'une majuscule afin de renforcer l'attention portée au caractère singulier de la chose, en l'écrivant de la même manière que n'importe quel autre nom propre. Caractérisé ainsi, le Texte est essentiellement une entité concrète qui s'étend dans le temps et dans l'espace.

On pourra d'emblée se demander si enrégimenter ainsi le Texte dans le domaine des inscriptions matérielles est vraiment pertinent. N'existe-t-il pas de Texte qui soit complètement dématérialisé ? Dans l'hypothèse où il serait tapé sur un clavier d'ordinateur, l'enregistrement d'un Texte consiste en une série de migrations d'électrons au sein de différents transistors assortis de combinaisons de pixels mis dans un certain état électronique. En ce sens, on peut certainement encore parler de la matérialité du Texte, bien que celui-ci mène une existence sans doute plus précaire que celle d'un texte écrit à l'encre sur du papier. Nous pourrions poser la question du commencement du Texte. Quand le Texte numérique démarre-t-il son existence ? Instantanément, lors de la frappe du premier caractère, ou bien lors de son enregistrement en dehors de la RAM, sur un disque dur matériellement présent ou à distance (cloud) ou sur un support externe (disquette, CD-ROM, DVD, cartouche, carte SD, etc.) ? La question est sans doute complexe, mais de toute façon, on la retrouvera aussi dans le cas d'une inscription sur papier avec de l'encre. Elle n'est donc pas propre aux inscriptions électroniques. Passons à la question de savoir quand le texte cesse d'exister. Sans doute est-elle, dans sa version électronique, complexifiée par le dispositif technique qui tend à accroître l'effet de palimpseste. En outre, la carrière d'un Texte électronique peut notamment dépendre des conditions de rafraîchissement de l'écran. Nous nous retrouvons alors face à des problèmes d'identité : sous quelles conditions peut-on affirmer que l'état de l'écran au moment t_1 est encore le même au moment ultérieur $t_1 + \mathrm{T}$ où T représente la période de rafraîchissement ? Entre deux périodes, s'agit-il encore du même Texte ? Ce problème peut aussi apparaître au niveau du stockage dans la RAM selon la manière dont les transistors persistent dans leur état physique. Nous pourrions alors considérer que, sous sa forme électronique, la carrière d'un Texte est extrêmement fugace, que son inscription électronique est un phénomène momentané aussitôt remplacé par des dupliques quasi-instantanées. Dans ce contexte, les questions d'identité sont sans doute nombreuses et complexifiées, mais il paraît toutefois indéniable que nous pouvons conférer le statut d'inscription matérielle à un quelconque enregistrement numérique.

Les modes de reproduction du Texte

Revenons à l'exemple de la lettre de Nietzsche. On dit qu'on peut se la procurer à moindre coût dans une édition de poche ou par téléchargement numérique. Dans ces situations, nous sommes encore en présence d'objets matériels, qu'il s'agisse de livres brochés ou de moniteurs d'ordinateur, sans toutefois être en « accointance » (pour le dire à la manière de Russell) avec l'inscription

originale. Ces nouvelles occurrences du Texte sont, en quelque sorte, libérées du support matériel unique qui est la prérogative du Texte. En ce nouveau sens, les textes peuvent s'affranchir des limites spatio-temporelles que leur impose *a priori* l'unicité d'un support dont l'étendue est inévitablement restreinte dans le temps et l'espace. On désigne habituellement sous le nom de « tirage » la reproduction du Texte en exemplaires multiples, le terme servant également à désigner le nombre d'exemplaires produits. Différents exemplaires relèvent du même Texte, mais bien sûr, chaque exemplaire constitue une nouvelle occurrence unique.

En tant qu'inscriptions matérielles, les tirages existent, de même que les originaux. Mais où situer le lien qui les corrèle ? Si nous disons de tel tirage qu'il relève de la *même œuvre textuelle* que tel original, sur quoi faisons-nous reposer l'identité présumée ? Devons-nous ajouter une entité exotique en plus des supports matériels ? Afin d'expliquer une putative identité du texte en dépit de la différence numérique des supports, doit-on inférer que le texte est quelque chose « en plus », derrière ou au delà de son seul support matériel ? S'agit-il alors d'une entité abstraite ? Il semble qu'il ne soit pas nécessaire d'hypostasier de telles entités, comme l'a indiqué notamment Goodman dans *Languages of Art*. En évitant d'introduire des *types* abstraits ou des entités platoniciennes, Goodman propose de considérer que le mode d'identification des textes ne relève pas de l'identité numérique de chacune des occurrences, mais de leur capacité à la reproduction en tant qu'art allographique, autrement dit, art pour lequel toute copie parfaite possède une authenticité équivalente à l'exemplaire original. Certes, il existe des contrefaçons de l'édition originale d'un texte. Mais en revanche il n'existe pas de contrefaçons des exemplaires eux-mêmes. On pourrait objecter que certains exemplaires pourraient avoir une valeur plus grande du fait d'être particularisés par certaines propriétés identifiantes, par exemple d'avoir appartenu à une personne notoire, de comporter des marques manuscrites intéressantes ou d'avoir été acquis dans certaines conditions. Mais une telle particularisation de l'exemplaire fait qu'il échappe à son mode commun d'existence au sein du tirage. Aussi, parler de « contrefaçon » à propos de le reproduction d'un Texte consiste au mieux à désigner improprement ce qui relève en réalité d'un plagiat, c'est-à-dire d'une attribution volontairement erronée [3].

Il reste alors à connaître les conditions de reproduction par lesquelles on peut authentifier différents textes en les identifiant les uns aux autres. Si nous prétendons qu'il existe une corrélation entre le Texte et chacune de ses copies, il s'agit alors de déterminer de quel type de relation il s'agit précisément. Insistons sur le fait que nous pouvons d'emblée éliminer la relation d'identité : une chose matérielle est identique à elle-même, et à rien d'autre qu'elle. Au sens littéral, une reproduction de la lettre de Nietzsche *n'est pas* la lettre de Nietzsche. Sur quoi fonder alors la relation de reproduction ?

On pourrait d'abord prétendre que, faute d'identité, l'identification doive être rapportée à la ressemblance. Cependant, considérés comme des ensembles de marques, deux textes imprimés peuvent encore être discernés par la taille des caractères, par les fontes employées, par la couleur de leur encre, par leur remplissage plus ou moins complet et par la netteté de l'impression.

[3] Il est à noter que les textes de loi concernant les droits de reproduction parlent bien de contrefaçons dans le cas de reproductions illicites d'un texte, en ce sens que la contrefaçon y désigne moins l'objet que « le fait d'exploiter, de faciliter l'exploitation ou d'utiliser une œuvre sans le consentement de l'auteur ou de ses ayants droit » (Code de la Propriété Intellectuelle du 1er juillet 1992, http://www.cfcopies.com/lexique/c). En revanche, le terme de plagiat n'y est pas employé.

En outre, différentes éditions d'un même texte peuvent également se différencier par la mise en page, autrement dit par le nombre de caractères par page et par ligne, par le nombre de lignes sur une page, par les espacements interlignes et marginaux et par les césures des mots. Elles peuvent également se distinguer par la qualité, la coloration et le grain du papier, ainsi que par le format de pagination. D'autres différences peuvent encore être ajoutées à ce constat, comme la présence d'illustrations, sous la forme de lettrines, de gravures hors-textes ou d'inclusions graphiques dans le texte. Mentionnons aussi les différences relevant du paratexte, comme par exemple des notes de bas de page dues à l'éditeur, ajoutant des commentaires, des références ou des précisions historiques au texte de l'auteur. La reproduction d'un Texte ne relève donc pas nécessairement d'une relation de similitude parfaite ou quasi-parfaite. Dans la terminologie décrite plus haut, nous ne devrions donc pas l'enrôler sous la notion de copie, mais bien plutôt sous la notion de duplique, moyennant diverses relations d'ajustement que nous allons décrire maintenant.

Différentes relations d'ajustement

Identifier des reproductions d'un Texte par une relation de similarité qualitative plus lâche ne conviendrait pas mieux. En effet, la marge d'identification que nous concédons est dépendante de nos intérêts. L'attention portée à la reproduction d'un Texte doit en conséquence s'accompagner d'une attention portée à l'acte même de reproduire, autrement dit, à la fabrication des reproductions. Or l'emploi de telle ou telle méthode de reproduction dépend de ce que nous considérons être une reproduction correcte. Et de fait, ces méthodes, qui dépendent de nos attentes, sont diverses. En considérant les reproductions d'un Texte comme des dupliques, ces méthodes de reproduction déclinent différentes relations d'ajustement dans l'assemblage que constitue la fonction d'œuvre textuelle, relations que nous allons illustrer par différents exemples.

Une première manière de dupliquer un Texte consiste à en recopier la chaîne des caractères, comme par exemple quand on recopie une citation sur un cahier. Cela revient à n'en retenir que les caractéristiques syntaxiques. La reproduction obtenue affranchit le Texte de toutes ses autres caractéristiques, comme la qualité du papier, la teinte de l'encre, la finesse de la plume, la mise en page, etc. qualités qui sont jugées non pertinentes dans l'acte de reproduire. Bien sûr, nous aurons à préciser les réquisits qui régissent de telles reproductions purement syntaxiques, mais poursuivons d'abord notre enquête sur les différents processus de reproduction d'un Texte.

Certains textes exigent de leurs dupliques qu'elles retiennent un nombre de ressemblances plus élevé que la seule pertinence syntaxique. C'est par exemple le cas du calligramme d'Apollinaire intitulé « Paysage » (APOLLINAIRE, 1965a), en partie reproduit à la figure 4.1 dont les seules composantes syntaxiques ne suffisent pas aux copies qui doivent encore prendre en compte la mise en page et sans doute aussi, dans une certaine mesure, la typographie [4]. Une copie qui reproduirait seulement la concaténation des caractères selon la chaîne « Cet arbrisseau qui se prépare à fructifier te ressemble » serait en effet inapte à restituer le sens intégral du poème qui dépend étroitement de sa forme calligraphique.

Pour d'autres types de calligrammes, en particulier pour les cryptogrammes, la duplication syntaxique, c'est-à-dire la restitution du texte sous la forme d'une concaténation des marques, au-

[4] A moins de considérer une acception plus élargie de la syntaxe qui tienne compte de ce genre d'organisation du texte.

CET
ARBRISSEAU
QUI SE PRÉPARE
À FRUCTIFIER
TE
RES
SEM
BLE

FIGURE 4.1 : Apollinaire, *Paysage*

rait pour effet néfaste d'en détruire inévitablement la signification. Car elle annulerait l'effet de codage dû au sens de lecture commun. Donnons en exemple le cryptogramme de Paul Nougé que Jacques Morizot cite dans *Le Problème de Borges* reproduit à la figure 4.2 et dont la forme dissimule un message que nous laissons au lecteur le soin de décrypter, et dont la chaîne concaténée « UEEVOSDHEMEUCMSSCUEEREONDEVBTTGIASEAASDSRJED » perd la signification.

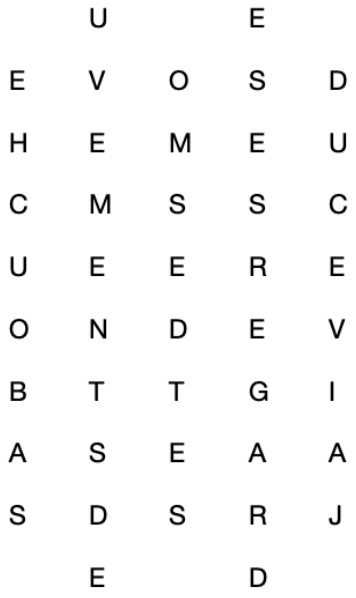


FIGURE 4.2 : Paul Nougé

À d'autres occasions, il arrive que la reproduction nécessite de préserver un nombre encore plus grand de caractéristiques. D'Apollinaire encore, on peut citer le poème intitulé « La Méduse », tiré du *Bestiaire* (APOLLINAIRE, 1965b), reproduit à la figure 4.3. Comme on peut le voir, le poème est accompagné d'un bois, estampe obtenue à partir d'une technique de gravure dont l'exécution suit deux phases. La première phase consiste à élaborer une matrice de l'image en gravant, de manière inversée, le motif sur une planche de bois. La deuxième phase consiste à prélever, sur du papier, l'empreinte de cette matrice à l'aide d'une encre uniformément répartie à sa surface. Le bois est une technique de gravure apparentée à l'eau-forte et à l'aquatinte. Ces différentes tech-

niques ne se distinguent que par le matériau employé pour produire la matrice (principalement du bois ou du métal) et par le type de solution chimique (de l'encre ou de l'acide) appliquée sur la surface de la matrice afin d'effectuer le prélèvement de l'empreinte. Dans ce type de reproduction, le Texte est bien dupliqué, mais chaque exemplaire est considéré à la fois comme une reproduction authentique en rapport avec son modèle matriciel et comme un original en tant que tel.



LÀ MÉDUSE

MÉDUSES, malheureuses têtes
Aux chevelures violettes
Vous vous plaisez dans les tempêtes,
Et je m'y plais comme vous faites.

FIGURE 4.3 : Apollinaire, *La Méduse*

Il pourrait être objecté une certaine absurdité à dire de plusieurs exemplaires qu'ils sont tous des originaux issus du même modèle, en ce sens que leurs apparitions se sont de toute évidence succédées dans le temps. Mais ce serait croire qu'une reproduction découle d'une autre, ce qui n'est évidemment pas le cas ici. En outre, la copie est différente de la matrice, ne serait-ce que parce que son dessin est inversé. En ce sens, on peut dire que le modèle est virtuel, qu'il n'existe pas matériellement parlant (ce qui n'est bien évidemment pas le cas de la matrice). En manque de modèle, chaque exemplaire vaut pour original. Les tirages sont d'ailleurs souvent numérotés afin de préciser le rang temporel d'un exemplaire, rang qui a une importance pour le collectionneur en raison de l'usure progressive de la matrice au gré des impressions. La valeur marchande d'un exemplaire de rang inférieur est jugée plus grande que celle d'un exemplaire de rang plus élevé parce que son dessin est estimé être plus précis. Mais cette occurrence n'est pas pour autant jugée plus authentique qu'une autre.

On pourrait contester que les poèmes à gravures [5] ne relèvent pas vraiment du régime des arts allographiques, mais plutôt de celui des arts autographiques car bois et gravures comportent une densité graphique qui échappe à la notation. Arts autographiques donc, mais aussi, arts à occurrences multiples, en raison que leurs œuvres sont constituées d'une multiplicité d'originaux spatialement et temporellement séparés. Avec la technique de la gravure *numérotée*, le Texte échappe inévitablement au régime de l'allographique, puisque la copie d'un exemplaire perd sa valeur d'authentique. La raison en est que la matrice n'est pas reproductible sans perte d'une qualité esthétique. Il en serait sans doute autrement si la matrice pouvait être reproduite d'une manière indiscernable, et dans ce cas, nous devrions sans doute admettre les arts à matrice dans le domaine des arts allographiques, valeur marchande mise à part.

La technique de la gravure numérotée est surtout employée pour l'impression d'éditions coûteuses d'œuvres textuelles. D'autres procédés réduisent considérablement les frais d'édition, comme la reproduction photographique d'un Texte et sa numérisation. Celles-ci peuvent retenir non seulement les composantes syntaxiques du Texte, mais également ses formes et ses proportions, la graphie et la mise en pages. Elles peuvent également indiquer les éventuelles ratures, les taches et les froissures du papier. L'intérêt de préserver autant de traits du Texte à travers les copies réside dans son exploitation historique et philologique.

Pour achever ce tour d'horizon des techniques de reproduction d'un Texte, mentionnons encore celles ayant pour objectif la reproduction d'un nombre maximum de caractéristiques, notamment les caractéristiques matérielles de son support, comme l'aspect et la qualité du papier. C'est évidemment le but que se donne la *contrefaçon* d'un original. Si l'intention n'est pas frauduleuse, autrement dit, si la reproduction n'est pas considérée comme un original, on parle alors de *fac-simile*, la qualité de celui-ci variant selon le nombre de caractéristiques rapportées de l'original à la duplique.

Pour résumer, s'il n'y a qu'une seule manière pour la copie d'un Texte d'être fidèle, il existe en revanche différentes manières d'en être des dupliques correctes, en fonction de sélections de ressemblances. Nous devons donc dissocier la notion de fidélité de celle d'authenticité : des dupliques peuvent s'avérer être toutes authentiques dans leur rapport à l'original qu'elles reproduisent tout en mobilisant des degrés de fidélité très variés, autrement dit, des relations d'ajustement plus ou moins contraignantes. Contrefaçons et *fac-simile* sont les plus exigeantes des dupliques, poussant la fidélité à son degré optimal. La similitude, quant à elle, est correcte si et seulement si elle conserve l'aspect visuel du texte, sa mise en page, sa calligraphie. Enfin, la « recopie » se contente de ne retenir que les caractéristiques syntaxiques de l'original. Selon leur degré de fidélité, les reproductions d'un Texte peuvent ainsi s'avérer authentiques tout en perdant certaines significations de l'original. Le *fac-simile* retiendra les diverses ratures d'un Texte que négligera sa reproduction commerciale. De l'un à l'autre, certaines significations seront perdues, précieuses pour le chercheur, mais certainement sans intérêt pour le simple lecteur.

On le voit, la copie comme relation d'imitation, basée sur la ressemblance qualitative, n'est pas une notion pertinente pour l'authentification des reproductions d'un texte. Mieux vaut lui substituer une notion de duplique par degrés de fidélité, autrement dit, une substituabilité d'ordre

[5] Notons qu'il existe des cas d'œuvres gravées où la matrice est restreinte à des caractères typographiques, comme par exemple les productions de Marion Renauld (http://marionrenauld.com), tapées à la machine à écrire, instrument utilisé comme un outil d'impression pour rédiger des poèmes de circonstance.

fonctionnel qui dépende de nos intentions à l'égard du Texte.

4.1.3 Équivalence des marques dans un système linguistique

Comme nous l'avons constaté, les situations où les copies nécessitent plus que la conservation de propriétés syntaxiques semblent manifester un type particulier de reproduction du Texte, parce qu'elles lui attribuent un caractère hybride entre art allographique et art autographique. Nous limiterons maintenant notre étude aux cas de duplications syntaxiques, c'est-à-dire aux cas où ce qui s'avère pertinent pour authentifier un texte ne ressortit qu'à ses composantes syntaxiques.

Si tout texte est un type d'inscription, n'importe quelle inscription n'est bien évidemment pas nécessairement un texte. Constituer un système d'annotations nécessite certaines conditions. Il faut bien sûr qu'un texte soit bien plus qu'un ensemble de marques, sans quoi une tache et une éraflure seraient indiscriminables d'un mot. Pour être conçues comme des éléments d'un langage partagé, les marques doivent instancier la qualité de symboles, c'est-à-dire qu'elles doivent être à la fois reproductibles et reconnaissables. Une marque qui n'aurait qu'une seule occurrence serait un *hapax*, une sorte de décoration ou de salissure, dans tous les cas, un bruit de fond du texte qui n'exprimerait rien. Cela ne veut pas dire pour autant qu'une tache soit toujours insignifiante. Elle pourrait traduire une certaine négligence de son auteur, c'est-à-dire fournir une interprétation psychologique indépendamment de l'intention consciente visée par celui qui l'a produite. Mais elle pourrait aussi véhiculer une intention de communication, par exemple servir de signal, pour coder un message destiné à n'être compris que par une personne initiée. La tache servirait alors d'indice, mais ne serait néanmoins pas une annotation au sens systématique du terme, puisqu'elle viserait une forme de dissimulation.

Ainsi, pour constituer un signal, c'est-à-dire une intention de communication entre un émetteur et un récepteur, un ensemble d'inscriptions présuppose d'abord un langage auquel il appartient, autrement dit, un système de signes conçu comme un ensemble fini de classes d'équivalence de marques qui répondent à une syntaxe, c'est-à-dire à un ensemble de règles d'agencement. La syntaxe prescrit, entre les marques, une relation d'équivalence qu'on pourrait qualifier de « substituabilité *salva syntaxe* ». Les classes d'équivalence des symboles *modulo* cette relation sont appelées « caractères » du langage. Tout texte peut alors être considéré comme une suite particulière de *caractères* comme par exemple « Longtemps, je me suis couché de bonne heure » ou comme « $z^2 = \dfrac{1+i}{1-i}$ ». De ce point de vue, un texte peut être considéré comme un enregistrement de marques dans un système linguistique. Un mot comme « Canada » est ainsi décomposable en sous-unités $\langle \text{C} \rangle, \langle \text{a} \rangle, \langle \text{n} \rangle, \langle \text{a} \rangle, \langle \text{d} \rangle, \langle \text{a} \rangle$ comprenant 4 caractères pour 6 signes. L'intérêt que présentent les signes identifiés sous des classes d'équivalence d'annotations est bien sûr leur capacité à être reproduits, à être mémorisés et à être diffusés. C'est le système d'annotation qui confère aux textes leur haute capacité à la duplication.

Les conditions de reproduction d'un texte sont ainsi prescrites par celles de la reproduction des caractères. Or nous allons retrouver, à l'échelle des marques syntaxiques, les problèmes de ressemblance que nous avons déjà évoqués dans le cas des textes. Comme n'importe quel texte, les marques d'un caractère peuvent varier non seulement par leur distribution spatiale, mais également par leurs formes et par leurs tailles. La condition scripturale est d'abord une question de la

reconnaissance d'une forme à travers un dessin. Mais en tant que telle, l'équivalence de caractères exclut tout ce que la marque possède d'inessentiel à son égard, comme peuvent l'être les traits dépendants de l'habilité plus ou moins grande du rédacteur. Bref, la reconnaissance du caractère exige d'expurger, comme le dit Michel Serres, tout ce que le dessin présente de *cacographie*, c'est-à-dire le bruit de fond du texte. Serres concède néanmoins que la cacographie est partout présente dans le texte, et donc que tout lecteur est un épigraphiste qui s'ignore :

> si l'on se demande quelles sont, dans l'écriture, la part de la forme et la part de la cacographie, il faut bien admettre que le bruit l'emporte, certains diront, exhaustivement. (SERRES, 1968, p. 43)

L'élimination du bruit de fond est toutefois nécessaire à l'aboutissement de la communication du message. La question n'est pas seulement liée aux textes manuscrits, puisque tout procédé de reproduction mécanique témoigne également de différences aspectuelles entre différentes marques d'un même caractère :



Comme en témoigne cette variété de formes, la relation d'équivalence qui corrèle les marques à l'intérieur d'une classe de caractères n'est pas une relation de ressemblance naturelle. Il faut donc souligner la nature conventionnelle des classes de caractères, qui ne dépendent pas exclusivement de leurs ressemblances mutuelles. Aussi diverses soient-elles en apparence, les différentes marques présentées ci-dessus sont uniformément identifiables au premier des caractères alphabétiques. Avec Goodman, on peut en particulier souligner que la comparaison des marques de la figure 4.4 indique que les deux premières sont moins ressemblantes entre elles que la première et la troisième, en dépit du fait qu'elles n'appartiennent pas aux mêmes classes de caractères [6].



FIGURE 4.4 : similitude entre caractères issus de classes différentes

Précisons en outre que, comme le montre l'exemple de la figure 4.5, les considérations topologiques ne sont pas plus pertinentes que les considérations isométriques pour assurer la relation d'équivalence de caractère. En effet, le premier « B » n'a pas la même configuration topologique que le deuxième, qui présente, quant à lui, la même configuration topologique que le « O ».



FIGURE 4.5 : ressemblances topologiques

[6] On peut notamment citer l'exemple de l'ancienne marque « f » en usage dans la typographie des textes du XVIIe siècle, marque qui est apparentée à l'actuel classe du caractère « s ».

L'appartenance d'un signe à un caractère relève parfois d'une pure stipulation, d'une relation d'indifférence de forme au point de ne ressembler que très vaguement à leur modèle de départ, comme dans le cas des lettres gothiques illustrées par la figure 4.6.



FIGURE 4.6 : lettres gothiques

Constatons qu'on peut retrouver dans le dénombrement des caractères la distinction établie par Charles Sanders Peirce entre *type* et *token*, qui peut s'interpréter autant comme une distinction entre un universel et des particuliers qui en dépendent, ou comme une classe de ressemblance et les éléments qui lui appartiennent. Mais Goodman préfère parler de classes et de répliques plutôt que de *types* et de *token* :

> Bien que je considère dans le présent texte un caractère comme une classe de marques, il s'agit pour moi d'une façon de parler peu rigoureuse qui n'est admissible que parce qu'on peut facilement la traduire dans un langage plus acceptable. Je préfère rejeter complètement le type et traiter lesdits « tokens » d'un type comme des *répliques* l'une de l'autre. Une inscription n'a pas besoin d'être un double exact d'une autre pour être une réplique, ou copie vraie ; de fait il n'y a en général aucun degré de similarité qui soit nécessaire ou suffisant pour qu'il y ait réplique. (GOODMAN, 1976, p. 209)

Goodman souligne ainsi que la reproduction n'est pas une copie puisqu'elle ne nécessite pas la quasi-similarité. Reste à savoir si, dans la terminologie exposée plus haut, il s'agit ou non d'une sorte de duplique, basée sur l'assemblage que constitue le texte par ajustements respectifs des caractères.

On peut juger à l'aune de sa déclaration la place éminente que Goodman confère au double dans sa philosophie de l'Art. La relation entre différentes copies d'un caractère n'est pas, dit Goodman en substance, une relation naturelle de ressemblance, d'imitation, mais une relation instituée, une forme de représentation. L'une des motivations pour renoncer à la ressemblance est que cette relation n'est, en général, pas transitive. Il faut néanmoins nuancer cette affirmation dans les cas où l'aspect des lettres est lui-même signifiant, comme pour certains messages publicitaires ou pour certains logos commerciaux, pour lesquels la substituabilité dans la classe d'équivalence syntaxique n'a plus cours. L'explication réside en ce que le logo est à la fois lu comme un texte et comme un dessin, ce qui revient à le considérer à nouveau comme un type de calligramme.

La ressemblance entre les reproductions et leur modèle doit ainsi toujours prendre appui sur des points de comparaisons relatifs à des intentions comparatives. Nos classifications ne sont pas indépendantes les unes des autres, et les jugement comparatifs sont toujours des constructions humaines, et non des vérités de faits. Ce que nous appelons une reproduction correcte d'un Texte n'est donc pas un fait du monde, mais une construction sociale enrégimentée dans des consensus culturels. Comme l'écrit Jean-Marie Schaeffer dans un passage de *Pourquoi la fiction* inspiré du relativisme de Quine à l'égard de la ressemblance, « toute similarité posée entre entre deux ou plusieurs choses résulte d'une hiérarchie attentionnelle qui nous fait privilégier, mettre en évidence certaines propriétés plutôt que d'autres » (SCHAEFFER, 1999, p. 86). S'il est juste de dire que les

classes de caractères ne sont pas entièrement définies par des ressemblances factuelles, il semble néanmoins faux de prétendre que leurs appariements y échappent complètement. Certaines personnes, atteintes de dyslexie, confondent les lettres « p », « d », « q ». Cette confusion s'explique par leur incapacité à distinguer les différences que ces formes entretiennent entre elles par rotations et réflexions, ce qui tend manifestement à montrer ici une évidente association des caractères au moyen de ressemblances.

Dans l'*Analyse des sensations* (MACH, 1991b), Ernst Mach prescrit une différence entre l'espace géométrique et l'espace optique. Il distingue ainsi les figures géométriques (*Gebide*) des formes optiques (*Gestalt*). Il donne les exemples suivants, reproduits sur la figure 4.7. Sur la figure de gauche, la même lettre « N » est reconnue deux fois en dépit des différences de perceptions des éléments sensoriels. La différence de couleur n'est pas un trait pertinent dans la reconnaissance de la lettre. En revanche, la partie droite de la figure 4.7 présente deux figures indiscernables (deux carrés) constituant néanmoins des formes différentes en raison des dissemblances qu'elles manifestent du strict point de vue rétinien. Cet écart entre figures géométriques et formes optiques permet de rendre compte de différences syntaxiques, comme par exemple celles qui distingue les figures indiscernables que sont les lettres « d » et « p ».

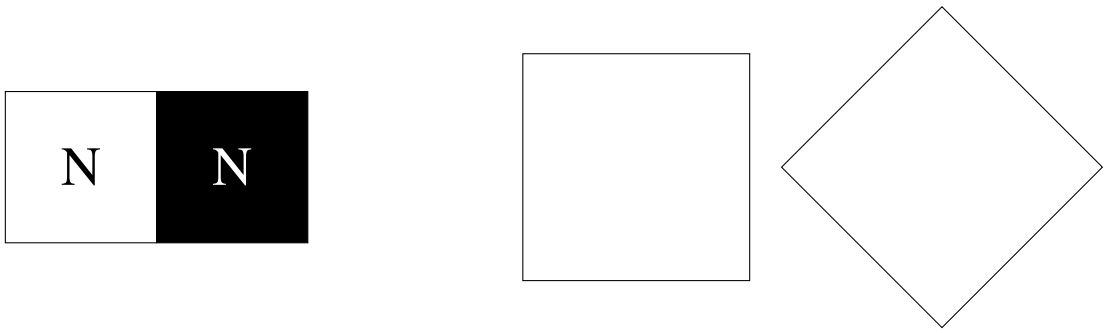


FIGURE 4.7 : figures et formes

Si les ressemblances relèvent toujours de constructions culturelles, elles n'en sont pas moins contraintes de manière factuelle. Si les entités ne manifestent pas leurs ressemblances réciproques en les pointant de l'extérieur par des flèches indicatives, et qu'en ce sens, elles dépendent évidemment de sélections proprement humaines, il est tout aussi vrai que seules celles qui sont opératoires peuvent faire l'objet d'une sélection pertinente. Ainsi, « reconnaître la fonction cognitive centrale de la relation de ressemblance n'est donc nullement incompatible avec une épistémologie constructiviste » (SCHAEFFER, 1999, p. 87). Ce qui fait d'une reproduction d'un Texte une duplique, ce n'est donc pas une quasi-similarité naturelle, mais une sélection fonctionnelle de ressemblances.

Quoi qu'il en soit, c'est la notion centrale de classe d'équivalence de caractères qui fait l'enjeu de la caractérisation des reproductions d'un Texte :

> être des exemples d'un caractère dans une notation doit constituer une condition suffisante pour que des marques soient des « copies vraies » ou des « répliques » l'une de l'autre ou soient orthographiées de la même manière. (GOODMAN, 1976, p. 168)

L'équivalence doit alors assurer l'univocité des caractères. Chacune des marques doit être duplicable par n'importe quelle autre appartenant au même caractère. Il reste alors à définir ce qui régit l'équivalence, ou précisément, la relation d'indifférence de ressemblance qualitative. Pour Goodman, cette relation doit être construite comme une équivalence de caractères dans un système de notation.

Dans *Languages of Art*, Goodman a montré que constituer un système de notation compatible avec une relation d'équivalence de caractères nécessite de répondre à deux réquisits syntaxiques et à deux réquisits sémantiques. Les réquisits doivent assurer, autant que faire se peut, la consistance et la décidabilité du système de notation.

Une première contrainte syntaxique est que chaque marque d'un système d'annotation doit pouvoir être authentifiée sans ambiguïté comme appartenant à un caractère particulier. C'est le premier réquisit syntaxique, qui concerne la disjonction des classes de caractères. Le point essentiel est que, dans un authentique système de notation, « des marques dont on juge correctement qu'elles sont co-membres d'un caractère seront toujours des copies vraies les unes des autres. » (GOODMAN, 1976, p. 172). Pour assurer la consistance du système de notation, aucune marque ne doit appartenir à plus d'un caractère, sans quoi l'équivalence aurait pour conséquence que des classes se chevauchant par une marque viendraient à fusionner complètement. Bien sûr, le chevauchement interdit aux classes doit être distingué de l'éventuel chevauchement physique des marques. Ce dernier ne présente, quant à lui, aucun danger d'inconsistance pour le système de notation, tout au plus une difficulté de déchiffrement du message dans le système.

Admettre l'équivalence qui existe entre des marques issues de différentes typographies conduit à les considérer elles-mêmes comme des sortes de dupliques, c'est-à-dire, à les envisager selon leurs ajustements respectifs dans l'assemblage que constitue un texte. J'appellerai C-équivalence (pour équivalence de caractère) la relation binaire qui relie deux marques interchangeables dans un texte donné. Plus précisément, un caractère c_1 est C-équivalent à un caractère c_2 et si et seulement si, dans n'importe quel assemblage de parties textuelles qui contient l'un d'eux, l'autre peut le remplacer en validant le même ajustement.

Sous les conditions d'ajustement que nous avons appelées *salva syntaxe*, la marque « a » frappée à l'aide d'une machine à écrire est considérée comme C-équivalente à la marque « a » frappée avec la même machine à un autre moment (en dépit de l'usure de la matrice). Cette marque est également C-équivalente à la marque « a » frappée avec une autre machine munie de caractères appartenant à une typographie différente. En revanche, la marque « a » est généralement considérée comme non C-équivalente à « A » et comme non C-équivalente à la marque « a ». La relation d'ajustement dont est tributaire la C-équivalence n'est cependant pas unique, et elle est dépendante du contexte. Elle est beaucoup plus souple dans le cas d'une retranscription manuscrite d'un Texte, par exemple parce qu'elle tend à négliger l'inclinaison des marques.

Le deuxième réquisit syntaxique relève de l'articulation des caractères, ou différenciation finie. Un système de notation doit permettre de décider, en un nombre fini d'étapes, si une marque appartient à tel ou tel caractère, et ainsi exclure les cas ambigus comme celui présenté par la figure 4.8 où la deuxième marque peut nous laisser hésiter entre choisir qu'elle appartient au caractère « a » ou au caractère « d ». La différenciation finie garantit ainsi la décidabilité du système de notation : pour tout couple de caractère ($\mathcal{C}$, $\mathcal{C}'$) et pour toute marque m disjointe, c'est-à-

dire n'appartenant pas à deux caractères à la fois, il doit exister un processus fini permettant de déterminer soit que m n'appartient pas à $\mathcal{C}$ soit que m n'appartient pas à $\mathcal{C}'$. S'il reste du vague dans l'identification d'un caractère, c'est seulement de manière contingente, par exemple en raison des conditions de sa réalisation.

add

FIGURE 4.8 :

Dans le cadre de ses recherches, la volonté de Goodman est de caractériser des systèmes de notation qui puissent rendre compte autant des marques scripturales que des marques picturales. Par son approche des œuvres conçues comme des ensembles de symboles, Goodman est amené à élargir la notion d'équivalence de marques afin d'y inclure aussi leur fonctionnement référentiel. Ainsi, Goodman ne s'arrête pas aux conditions formelles, et ajoute encore deux réquisits sémantiques qui viennent compléter les réquisits syntaxiques de disjonction et d'articulation déjà cités. Les contraintes qui régissent le niveau syntaxique valent d'ailleurs aussi sur le niveau sémantique, et on parlera ainsi de relations de concordance pour qualifier les réquisits de disjonction et d'articulation sur le plan sémantique. Ces niveaux reflètent ce que nous avons appelé, dans la terminologie des dupliques, la relation d'ajustement.

On peut maintenant formuler une condition nécessaire et suffisante d'identification des reproductions correctes d'un Texte :

Critère d'identification des copies un texte x est, sous la fonction d'œuvre textuelle, une duplique d'un Texte y si et seulement si l'ensemble des marques de x est C-équivalent à l'ensemble des marques de y énoncés strictement dans le même ordre.

On peut également définir la classe d'équivalence d'un Texte comme l'ensemble des dupliques de ce Texte.

4.1.4 Tracer l'authenticité des dupliques

Authentifier des dupliques matérielles

Avec un tel critère, nous avons maintenant un modèle d'authentification des dupliques par rapport au Texte. Reste à savoir si cette authentification est aussi valable entre les dupliques elles-mêmes. Dans le cas simplifié, illustré sur la figure 4.9, où un Texte fournirait un exemplaire original qui serait la source unique d'une chaîne ramifiée de diffusion de copies, plusieurs difficultés apparaissent encore.

Premièrement, l'identification des copies au Texte doit être valable non seulement directement à partir de la source, par exemple entre C_1 et Texte, mais également entre les copies appartenant à des générations successives comme C_1 et $C_{1;1}$. En outre, puisque l'authentification des copies doit intervenir sur différentes générations, l'identification des textes doit conséquemment reposer

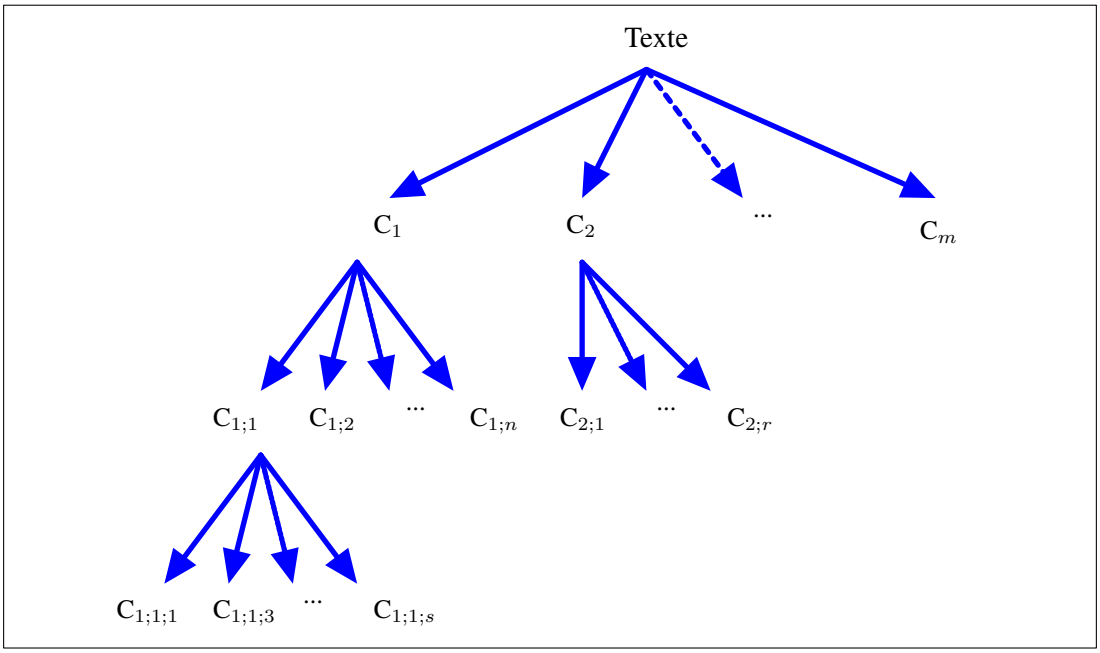


FIGURE 4.9 : schéma d'authentification des copies d'un texte

sur un principe récursif, puisqu'on doit pouvoir, en sautant une génération, identifier le texte de la copie $C_{1;1;1}$ au texte de la copie C_1 comme on le voit sur le diagramme proposé.

Un deuxième problème vient de ce que des copies d'une même composition mais provenant de tirages différents doivent être également authentifiées. La copie $C_{1;1}$ doit être authentifiée à la copie $C_{2;1}$ nonobstant que ces copies soient elles-mêmes issues de copies différentes (à savoir, C_1 et C_2 sur le diagramme proposé).

Puisque la C-équivalence est construite comme une relation d'équivalence, la relation qui corrèle les copies est transitive. En conséquence, l'authentification des copies est étendue non seulement à toutes les générations de copies mais aussi entre les différentes sources de copies.

Authentifier des textes digitaux

Nous venons de considérer un schéma de reproduction des copies qui découlent d'un modèle matériel, à savoir le Texte, dont l'authenticité peut être caractérisée en termes spatio-temporels. Mais ce modèle ne se laisse pas toujours aussi bien caractériser, notamment dans le cas où l'on partirait d'un fichier numérique. Les objets digitaux peuvent être vus comme des types de textes, ramenés, en langage machine, à une chaîne binaire exclusivement constituée de 0 et de 1. Ces objets sont non seulement très facilement reproductibles, mais en outre, ils le sont de manière parfaitement indiscernable. De fait, ils posent la question de l'authenticité, et plus particulièrement, celle de leur rareté : est-il possible de conférer de la rareté à un objet digital ?

Le domaine des objets esthétiques fournit des exemples intéressants d'objets supposés rares et singuliers, et certains d'entre eux sont justement des objets numériques. Pour les partisans de l'art

conceptuel, l'œuvre ne réside pas dans une réalisation matérielle mais dans l'intention de l'artiste. L'objet esthétique est donc dématérialisé, ramené au strict statut d'idée, idée qui peut être formulée par un ensemble de contraintes, de règles, autrement dit, par un programme. D'un point de vue historique, le premier objet d'art digital reconnu est sans doute *Oscillon 1*, réalisé en 1950 par l'informaticien américain Ben Laposky (1914–2000) à partir de courbes de Lissajous. Dans une rétrospective de 1953 à Cherokee (Iowa), Lapovsky désigna ses œuvres sous le nom d'abstraction électroniques (*electronic abstractions*).

Considérons quelques exemples d'objets esthétiques digitaux. En 1965, Michael Noll a reproduit un exemplaire numérique de *composition avec lignes* (1917) de Pietr Mondrian à l'aide d'un ordinateur (voir figure 4.10). Plutôt que de reproduire avec exactitude l'aspect du tableau de Mondrian, Noll a produit un exemplaire différent, mais respectant les mêmes contraintes formelles. Le principe qui organise l'œuvre originale est de disposer à l'intérieur d'un grand cercle blanc et de manière plus ou moins aléatoire des segments noirs horizontaux ou verticaux de longueurs variables. Noll a généré la disposition des segments en simulant de l'aléatoire avec un ordinateur et en conservant la densité graphique et les dimensions de l'œuvre originale. De fait, Noll, intrigué par le jeu de l'imitation de Turing, en proposait sa propre version : « une machine peut-elle produire une peinture convaincante et, le cas échéant, quelle différence y aurait-il entre cette peinture et celle d'un véritable peintre ? ». Autrement dit, là où Turing demandait si les ordinateurs peuvent penser, Noll demandait si les ordinateurs peuvent produire des œuvres esthétiques. En outre, l'œuvre de Noll étendait celle de Mondrian tout en l'absorbant : là où Mondrian n'avait réalisé qu'une possibité de répartition parmi d'autres, celle de Noll la dépassait par la totalité des possibilités que recouvrait son programme informatique. Pour autant, il se trouve que, bien que l'œuvre possède de troublantes ressemblances formelles, elle n'est pas organisée exactement par le même principe. Les nouvelles œuvres sont organisées par des agencements plus souples que l'originale, ce qui a incité différents juges à la préférer à celle de Mondrian [7]. C'est pour cette raison que Florence de Méredieu écrit :

> Michael Noll intègre dans l'ordinateur les données relevées dans les tableaux de Mondrian, classe celles-ci et demande ensuite à la machine de fabriquer de nouveaux « Mondrian », des Mondrian « possibles » et que Mondrian n'a cependant pas réalisés. Tout ceci est dans la ressemblance, dans le *comme si*. (MÉREDIEU, 2003, p. 105)

Meredieu constate d'ailleurs que :

> Au vu des « nouveaux tableaux » réalisés par Noll et plébiscités, de fait, par le public, on est en droit de s'interroger : ceux-ci sont-ils véritablement des tableaux que Mondrian aurait pu réaliser ? La question est complexe, car la répartition même des traits et des points ne correspond pas au fort géométrisme du peintre. La nouvelle œuvre apparaît comme plus pointilliste, plus impressionniste. Elle rompt donc largement avec l'esthétique de Mondrian. Signe qu'une

[7] Noll commente ainsi les réactions des spectateurs de son œuvre : « Xerographic copies of the two pictures were presented, side by side, to 100 subjects with educations ranging from high school to post-doctoral; the subjects represented a reasonably good sampling of the population at a large scientific research laboratory. They were asked which picture they preferred and also which picture of the pair they thought was produced by Mondrian. Fifty-nine percent of the subjects preferred the computer-generated picture; only 28 percent were able to identify correctly the picture produced by Mondrian » (NOLL, 1967, p. 92).

apparence formelle ne suffit pas à rendre compte d'une œuvre dont un processus purement mécanique ne parvient finalement à créer qu'une simple caricature. (MÉREDIEU, 2003, p. 105)



FIGURE 4.10 : Michael Noll, *Computer Composition With Lines*, 1964

De son côté, en 1974, Véra Molnar, artiste numérique, a créé une œuvre intitulée *javà de 24 carrés* qui présente des séries de superpositions de carrés aux longueurs aléatoires et dont les centres sont aléatoirement décalés les uns par rapport aux autres. L'œuvre de Molnar ne relève pas de l'achèvement d'un produit lié à des contraintes artisanales ou industrielles. Il s'agit d'un programme qui permet d'obtenir les tracés par la médiation d'une interface graphique comme un écran ou une imprimante. Le programme est l'œuvre, et les instanciations graphiques sont les témoignages matériels pointant vers cette œuvre, immatérielle. En dehors des aspects optiques et plus généralement esthétiques soulevés, c'est en fait la somme méréologique des occurrences de la série qui permet la compréhension globale de l'œuvre. Car celle-ci dépasse les occurrences imprimées par Molnar. En fait, l'œuvre de Molnar indique, à travers les fragments imprimés, l'ensemble complet des multiples agencements possibles que pourraient occuper les carrés. L'œuvre n'est donc pas délimitée par l'extension de l'ensemble des exemplaires imprimés, mais relate les gravures qui *pourraient* ou auraient pu être tracées, sans qu'elles le soient actuellement.



FIGURE 4.11 : Vera Molnar, *Javà de 24 carrés*, 1974

Une question que posent les objets digitaux est celle du droit d'auteur et plus généralement de l'authentification. Qui est le propriétaire de *composition en lignes* ? Si le tableau de 1917 peut être attribué à Mondrian, il n'en va sans doute pas de même pour l'idée exemplifiée par ce tableau, que Noll interprète comme un programme.

Les nouvelles technologies de l'information et de la communication permettent une duplication des sources et une facilité de modifications telles qu'il devient souvent possible à tout utilisateur d'altérer un document original. Les bases de données, qu'il s'agisse de transactions financières, d'archives d'histoire ou de pédigrée d'œuvres d'art, peuvent ainsi être remplacées par des faux.

A la fin des années 80, deux scientifiques du New Jersey, Stuart Haber, cryptographe, et W. Scott Stornetta, physicien, s'inquiétèrent de l'aisance avec laquelle peuvent être dupliqués et modifiés les fichiers numériques. Ils posèrent les deux questions suivantes : « comment, désormais,

saurons-nous ce qui est vrai du passé ? » et « comment pourrons-nous encore faire confiance en ce que nous savons du passé sans être tributaire d'une autorité centrale détenant les archives ? » :

> In many situations there is a need to certify the date a document was created or last modified. For example, in intellectual property matters, it is sometimes crucial to verify the date an invent or first put in writing a patentable idea, in order to establish its precedence over competing claims. (HABER & STORNETTA, 1991, p. 99)

L'idée des deux scientifiques fut de construire un registre d'archives ne dépendant pas d'une autorité centrale. Pour cela, ils utilisèrent la capacité de duplication des fichiers digitaux de la manière suivante : produire une chaîne de copies des documents, tout en rendant chaque occurrence de cette chaîne dépendante des autres occurrences de telle manière que l'altération d'un élément produise inévitablement la destruction de tous les autres documents de la chaîne. Leur idée fut donc de créer un registre crypté, horodaté, partagé et diffusé. Ils ont ainsi posé les fondements de la technologie des chaînes de blocs (*blockchains*). Il s'agit de processus facilitant l'enregistrement des transactions et de suivi des actifs, que ceux-ci soient des biens tangibles comme des biens immobiliers, ou intangibles, comme des droits de propriété intellectuelle. Un registre de *blockchain* permet ainsi de garantir le suivi dans le temps de n'importe quel bien doté d'une valeur, tout en réduisant les risques de faux. Cette chaîne permet d'assurer l'authenticité des transactions en conservant un enregistrement permanent et surtout immuable, donc infalsifiable, des données. Le principe général est le suivant : chaque participant dispose de la capacité de partager le registre qui est mis à jour, à chaque transaction, par réplication de pair à pair. De cette manière, toutes les parties possèdent la même copie du registre. Ainsi, l'enregistrement de la transaction est non seulement partagé par toutes les parties, mais surtout, visible par tous. Aucune modification du registre ne peut être apportée sans laisser une trace visible par tous. Les données sont stockée dans un registre sur le réseau, l'enregistrement prenant la forme de blocs de transactions reliés entre eux par une chaîne. Les blocs sont horodatés par un code, une empreinte numérique appelée *hash*. Chaque bloc contient le *hash* du bloc précédent, ce qui permet ainsi de les chaîner en évitant qu'un bloc ne soit modifié ou inséré entre deux blocs existants. La seule manipulation possible est donc l'insertion d'un bloc à la suite de la chaîne. Ainsi, la chaîne s'étend au fur et à mesure des transactions. Tout objet numérique contient deux composantes : l'une, en quelque sorte immergée, des méta-données, l'autre, celle-là émergée, de son corps d'inscriptions. Tout fichier est duplicable à volonté, mais la technologie des chaînes de blocs permet d'assurer une stabilité des authentifications à travers des empreintes numériques sur des enregistrements réseaux. Ces technologies ont donc permis le développement de réseau d'identification des objets numériques créés.

Les crypto-monnaies, comme le Bitcoin lancé en 2009 par le collectif Satoshi Nakamoto, sont des applications particulières basées sur des chaînes de blocs s'inspirant des méthodes de Haber et Stornetta. De fait, l'économie bitcoin échappe à toute autorité centrale, cette monnaie étant traitée par une chaîne de blocs à l'intérieur de registres stockés sur des réseaux informatiques fonctionnant de pair à pair.

Contrairement aux crypto-monnaies, qui par définition, sont fongibles, les *non fungible tokens* (NFT) sont des objets non fongibles, uniques, rares, et indivisibles. Il s'agit de sortes de certificats dont l'enregistrement est basé sur des chaînes de blocs utilisées comme source unique d'authentification et de pedigree pour l'achat ou la vente d'objets numériques, voire d'objets matériels.

L'unicité est codée par une empreinte numérique à l'intérieur des méta-données, ce qui lui assure une sorte de certificat d'authenticité. Bien sûr, certains objets numériques, bien qu'uniques en terme d'identité numérique encodée, sont indiscernables du point de vue de leurs apparences. Les développeurs peuvent cependant brider des aspects afin de les rendre « rares ». On a donc affaire à une rareté marchande qui n'est pas basée sur une rareté naturelle, comme celle d'un matériau, ni sur la rareté d'une exécution, comme celle d'une œuvre graphique, mais d'une rareté complètement intentionnelle [8]. Les NFT ouvrent ainsi la possibilité d'authentifier les objets digitaux.

Les NFT contribuent donc à désolidariser l'appropriation de l'objet numérique de l'appropriation d'un enregistrement numérique. Tout utilisateur d'Internet peut acquérir une reproduction de fichier informatique en l'enregistrant sur sa propre machine, mais il n'en sera pas pour autant le propriétaire. La technologie des FTP permet ainsi d'authentifier le véritable propriétaire d'un objet numérique tout en laissant circuler à loisir ses reproductions. Alors que la libre reproduction des fichiers digitaux laissait présumer, à l'aube de l'informatique, une disparition plus ou moins rapide de l'authenticité attribuable aux fichiers numériques, celle-ci est maintenant promue par la capacité à la reproduction des blocs de chaînes à travers les réseaux pair-à-pair. C'est donc, assez paradoxalement, la facilité de reproduction des objets digitaux qui en assure, aujourd'hui, la propriété privée.

Plus encore, la singularisation que permet la technologie des chaînes de blocs via les NFT transforme en donnée marchande l'appropriation de n'importe quelle chose, de n'importe quel événement, de n'importe quel état de choses, un peu à la manière inaugurée par l'explosion du marché de l'art contemporain. Le premier tweet envoyé a été vendu 2,9 millions de dollars par son inventeur. Le marché des NFT est ainsi devenu une bulle spéculative aux surenchères concurrençant sérieusement le marché de l'art, cela avec des objets immatériels n'ayant aucunement affaire à l'esthétique, et sans même prétendre à les transférer dans le domaine de l'art comme c'était encore le cas pour les *ready made*. C'est, en quelque sorte, une forme de revanche des objets ordinaires contre le traitement que leur ont consacré Duchamp et Warhol, objets qui peuvent dès lors accéder à une valeur marchande hypertrophiée sans nécessairement passer par une consécration esthétique conférée par un individu ayant autorité esthétique.

Les technologies des chaînes de blocs proposent donc de se passer d'un modèle, d'un original temporel dont seraient tirées les copies, en instituant un fichier de transactions dans lequel son historique commercial apparaît. Ce qui assure l'authentification et la rareté du fichier, ce ne sont pas les données numériques du fichier lui-même, ni même ses méta-données, mais son inscription dans la chaîne de blocs qui en constitue le pédigrée. C'est ainsi la possibilité même d'être reproduit qui confère au texte numérique son authenticité.



Voici notre conclusion pour cette section : en ce qu'elle dépend du contexte de fonctionnement, la reproduction d'un texte n'en délivre ni une réplique pure, ni une copie : elle n'est pas

[8] On peut citer le cas des figurines à collectionner *Rare Pepes* représentant des images de grenouilles diversement accoutrées : en 2019, une carte digitale représentant le personnage Homer Simpson en Pepe The Frog a été vendue à 38 500 dollars. Autre exemple, un éditeur de jeux vidéos a mis sur le marché des icônes de chats, les *cryptokitties*, dont la vente de l'un d'eux a atteint la valeur de 170 000 dollars. A ce sujet, voir .

ontologique, mais pragmatique. La notion de duplique, basée sur des relations qui ne sont pas exclusivement intrinsèques, mais aussi extrinsèques, permet en revanche d'authentifier entre eux les textes remontant au même Texte, non par des ressemblances qualitatives, mais, nous l'avons vu, par des intentions pratiques complexes qui visent à admettre ou non qu'un texte donné fonctionne comme duplique d'un autre au sein de l'assemblage que constitue l'œuvre textuelle.

Dans la section suivante, nous allons confronter la notion de duplique à diverses altérations que peut ou non tolérer une œuvre textuelle.

4.2 L'œuvre et ses doubles

4.2.1 La question de la nature des œuvres littéraires

En cherchant ce qui nous permet d'identifier une œuvre, nous sommes naturellement conduits à faire usage de l'unicité des inscriptions. Dès lors, la question qui se pose est de savoir s'il est possible d'identifier la notion d'œuvre à celle de classe d'équivalence des textes, ou si la notion d'œuvre « déborde » celle-ci. Cette question recouvre évidemment celle, ontologique et plus générale, de la nature et du mode d'existence des œuvres. Les positions sont diversifiées selon selon le classement suivant :

1. Une œuvre textuelle est une entité réelle et autonome, autrement dit, sa réalité est indépendante de l'expérience qui peut en être faite. On peut décliner cette réponse affirmative selon différentes conceptions :

   (a) une œuvre textuelle est un particulier concret existant et actuel. Cette position peut être dite réductionniste puisqu'elle ramène les œuvres à n'être que des objets matériels (des textes) ou des événements (des exécutions) ;

   (b) une œuvre textuelle est un particulier concret inexistant et actuel ;

   (c) une œuvre textuelle est un particulier concret et simplement possible, une entité modale ;

   (d) une œuvre textuelle est un particulier abstrait actuel (essence individuelle, *haeccéité*, un trope, une forme intelligible *ante rem*) ;

   (e) une œuvre textuelle est un universel actuel *in res*, une forme hybride.

2. Oui, une œuvre textuelle est quelque chose de réel, mais seulement de manière hétéronome, c'est-à-dire que sa réalité est dépendante de l'expérience qui en est faite. Sa réalité n'est que relationnelle et peut être déclinée ainsi :

   (a) être un particulier existant abstrait et actuel créé (théorie des artefacts abstraits),

   (b) être un objet social (institutionnalisme pragmatique) visé par des pratiques qui ont cours dans le monde de l'art (George Dickie, John Searle, Jean-Pierre Cometti).

3. Une œuvre textuelle n'est pas une chose réelle (anti-réalisme).

Attardons-nous un instant sur la pertinence de cette classification. L'affaire des alternatives (1) et (2), qui sont toutes les deux des options réalistes, est de stipuler la *nature* des œuvres (objet physique, objet immatériel, entité abstraite) et de stipuler leurs modes d'existence (autonomie ou hétéronomie). Il reste ensuite à établir les conditions d'identité qui régissent de telles entités, qui peuvent être considérées comme particulières ou universelles. À l'inverse, pour les partisans de la dernière alternative, anti-réaliste, il ne s'agit pas de considérer que les œuvres aient une quelconque *nature*, puisque rien de tel n'existe, mais cela n'engage nullement à renoncer à des identifications.

Ce tableau n'a pas pour objectif de présenter une vue *cohérente* des options possibles, mais une vue d'ensemble des théories, indépendamment de la cohérence qu'elles peuvent présenter. Cette typologie résulte donc des strictes affirmations des auteurs concernant le choix des couples d'opposition réel/irréel, existant/inexistant, etc. auxquels ils souscrivent. Je n'ai donc pas tenu compte de ce que je pense moi-même de ces nuances, ni de la manière dont j'aurais pu ou voulu élaborer cette classification.

Roger Pouivet considère que c'est un tort d'assimiler la théorie des artefacts abstraits d'Amie Thomasson à une forme de réalisme des œuvres, et qu'elle devrait figurer plutôt dans la troisième partie de la typologie proposée, à savoir, dans les théories qui affirment que l'œuvre n'est pas réelle [9]. Outre que cette contestation paraît se désolidariser de l'usage général qui tend à affilier ladite théorie au réalisme des fictions (FRIEND, 2007) et à la volonté revendiquée par son auteur d'en défendre la cause (THOMASSON, 2011), il me paraît peu judicieux de considérer que des entités qui sont dites être générées de manière objective par des actes de création ne relèvent pas d'une forme particulièrement forte de réalisme des œuvres. Aussi, il me semble pour ma part que l'option revendiquée par Thomasson correspond bien à une forme de réalisme. J'aurais donc tendance à inscrire la théorie des artefacts abstraits parmi les formes de réalisme au sens plein du terme, mais encore une fois, ma propre façon de classer les théories n'est pas ce qui compte ici.

Certains diront qu'il s'agit là d'une affaire d'intuitions. Je parlerais plutôt de différences métaphysico-idéologiques. Je crois qu'il arrive que certains auteurs défendent le réalisme de manière superfétatoire à la théorie qu'ils proposent, théorie qui pourrait parfois tout aussi bien s'en passer. C'est peut-être le cas pour Thomasson, dont les thèses pourraient, sans trop de conséquences, se passer des artefacts abstraits et ne mobiliser que des contenus purement intentionnels. Mais choisir l'option réaliste, c'est déjà être réaliste : l'engagement est un acte performatif. Il est difficile de refuser à son auteur le statut de réaliste même quand cette revendication semble arriver comme un cheveu sur la soupe.

Si l'on veut maintenant situer l'œuvre textuelle par rapport au Texte et à ses reproductions, les théories précédentes nous convient à choisir entre les possibilités alternatives suivantes :

1. l'œuvre textuelle *est* le Texte ou la classe d'équivalence des textes ;
2. l'œuvre textuelle est une entité à laquelle *réfère* le Texte ou la classe d'équivalence du Texte.
3. l'œuvre textuelle est une entité qui *survient* sur le Texte ou sur la classe d'équivalence du Texte ;
4. l'œuvre textuelle n'est rien, au mieux une manière impropre de parler ;

[9] Communication personnelle de Roger Pouivet.

Dans *Qu'est-ce qu'une œuvre d'art ?*, Roger Pouivet en caractérise les occurrences par le fait d'instancier des propriétés extrinsèques relationnelles [10]. Une œuvre est une entité qui est déterminée non par des propriétés intrinsèques, mais exclusivement par des propriétés extrinsèques qui dépendent de ses émetteurs et de ses récepteurs. Il cite sur ce point Étienne Gilson qui écrit que :

> La cause de l'œuvre d'art lui est extérieure ; elle n'a pas en elle-même le principe interne de sa croissance et des ses opinions. Elle n'est pas une nature, elle ne vit pas. (GILSON, 1972, p. 141)

Le même point de vue est soutenu chez Pouivet par la thèse aristotélicienne que les entités sont déterminées par des causes matérielle (le texte particulier), efficiente (l'auteur du texte), formelle (la signification auctoriale) et finale (le but visé par l'auteur). Une œuvre d'art est, selon Pouivet, un objet esthétique, mais dont l'existence n'est pas seulement intentionnelle :

> Dans un monde où une chose peut être réelle même si elle est constitutivement relationnelle - c'est-à-dire elle n'existerait pas si d'autres choses n'existaient pas - alors, dans un monde de ce type, les œuvres d'art existent réellement. Même si elles existent en dépendant d'autres choses c'est-à-dire d'êtres comme nous sommes. (POUIVET, 2016, 35 :45)

Selon Pouivet, les œuvres d'art semblent exister en dehors de nous, mais pas indépendamment de nous. Elles ne sont pas intégralement matérielles, elles relèvent aussi de l'esprit. Ce qui les distingue des autres artefacts, c'est qu'elles exemplifient des propriétés esthétiques. C'est évidemment très discutable au sens où de nombreuses œuvres n'exemplifient aucune propriétés esthétiques, et inversement que de nombreux artefacts, pour ne pas dire des objets naturels, exemplifient, quant à eux, des propriétés esthétiques. En tout cas, l'identification de l'œuvre aux propriétés esthétiques ne va pas de soi.

En outre, si l'on considère que les œuvres sont des entités, il importe alors d'en assurer les conditions d'identité. Même si ces dernières devaient demeurer vagues ou inconnues, il serait néanmoins primordial qu'*une* œuvre puisse compter comme *une* chose plutôt que comme plusieurs. La remarque pourrait paraître anodine ou triviale, mais il me semble pourtant que croire que les œuvres sont des choses plutôt que des manières de décrire les choses engendre inévitablement toute une série de problèmes de doubles. Les œuvres semblent relever ni complètement du réel, ni complètement du virtuel, mais, en quelque sorte, *des deux à la fois*. Cette hybridité n'est pas ontologique mais descriptive. Nous utilisons les mots « texte » et « œuvre » pour décrire des aspects différents d'une seule et unique chose.

On peut aussi, et inversement, admettre que ce qu'on désigne sous le prédicat d'œuvre puisse compter de différentes manières, comme une chose, ou comme plusieurs choses, voire comme une infinité de choses distinctes, sans devoir multiplier pour autant l'entité qui est décrite. Dans ce cas, le qualificatif d'*œuvre* ne doit plus se rapporter à une substance, à une entité, qui par

[10] Rappelons qu'une propriété d'un objet x est dite *intrinsèque* si elle ne dépend que de x. Dans ce cas, elle peut être relationnelle, comme par exemple « être identique à soi-même » ou non relationnelle, comme « être sphérique ». Une propriété est dite *extrinsèque* si elle dépend d'au moins un autre objet que x. Elle est extrinsèque impure si elle dépend d'un objet particulier comme « être à 1 km de la Tour Eiffel » et extrinsèque pure s'il elle dépend d'autre chose que x sans que cette chose soit déterminée par un particulier donné comme « être à 1 km d'une tour ».

définition est toujours singulière, mais à une manière d'envisager, de décrire, certaines choses par certains aspects, qu'il s'agisse d'aspects esthétiques, culturels, institutionnels, didactiques, historiques ou sociologiques, voire souvent plusieurs d'entre eux, sinon tous à la fois. Il s'agit de descriptions élaborées à partir de concepts qui n'ont rien de statique. Elles font elles-mêmes l'enjeu de ré-interprétations successives. Chaque description a un effet « *feed-back* » qui peut modifier les descriptions ultérieures. Précisément, Je soutiens que ce qui est dénommé « œuvre de fiction » consiste en un processus social et dynamique dans lequel des objets matériels prévus à cet effet (les dupliques du Texte) font l'enjeu de projections modales élaborées à partir des inscriptions matérielles. Comme l'indique Kendall Walton, de telles inscriptions servent de support (*prop*) dans des jeux de croyance feinte (*make-believe*) (WALTON, 1990).

Il faut préciser que ce point de vue n'est en rien associé à une ontologie hiérarchisée. Je ne souscris donc à aucune des positions énumérées dans le tableau précédent, qui souscrivent toutes à ce que John Heil appelle « la théorie picturale du langage ». Toutes ces théories peuvent être résumées selon le schéma 4.12. Or nous pourrions remettre en question le caractère exhaustif de l'ensemble des théories en abandonnant l'idée que la réalité suit un ordre hiérarchique que lui imposerait le langage :

> Nous pouvons admettre des niveaux d'organisations, des niveaux de perplexités, des niveaux des descriptions, des niveaux d'explications sans opter pour des niveaux de réalité, comme le font aujourd'hui les anti-réductionnistes auto-proclamés. Il en résulte une conception du monde et de la représentation que nous en avons réductible ontologiquement mais pas analytiquement. (...) Je préconise donc d'abandonner l'idée que la réalité est hiérarchisée. (HEIL, 2011, p.)

4.2.2 Différentes versions du textualisme

Comme nous l'avons vu, une bonne façon d'identifier les copies d'un Texte à leur modèle est de souscrire à une option nominaliste au sujet des entités qui en ressortissent, autrement dit, à éliminer de notre discours les « types » abstraits au profit des tokens, des occurrences particulières. Et comme nous l'avons constaté, l'inscriptionnalisme fournit un critère d'identité basé sur la relation de C-équivalence. Mais cette conception permet-elle de rendre compte de manière satisfaisante de l'identification des œuvres ? Lorsqu'on reproduit un texte, la question est finalement de savoir au juste ce qui est reproduit. La reproduction du texte implique-t-elle *ipso facto* la reproduction de l'œuvre afférente ?

Le point de vue substantialiste admet de l'œuvre textuelle est une entité, qu'il s'agisse soit d'une entité matérielle, soit d'une entité abstraite. Je voudrais récuser ces deux options au profit de la suivante : l'œuvre est une construction sociale basée à partir de la notion de duplique, autrement dit, non par des doubles ontologiques comme les répliques. Mon point de vue est donc pragmatique : l'œuvre textuelle est la fonction d'assemblage qui corrèle les dupliques du Texte selon les ajustements adéquats. Ce sont des pratiques culturelles qui déterminent la tolérance d'ajustement entre les reproductions, les authentifiant ou non comme des dupliques du Texte.

Je souhaite donc, dans ce qui suit, critiquer le point de vue substantialiste qui rabat l'œuvre entièrement sur les textes, et le point de vue émergentiste qui soutient que l'œuvre survient sur le

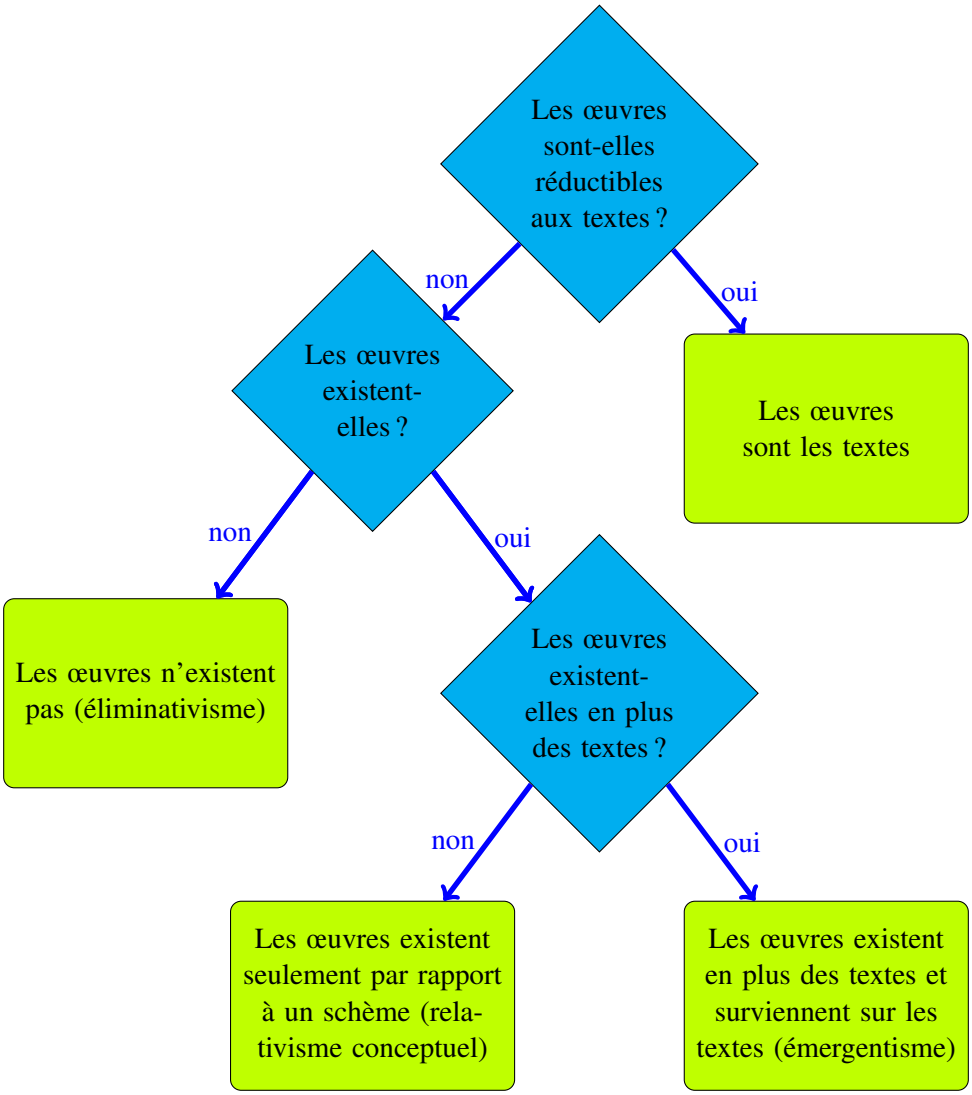


FIGURE 4.12 : Les différentes conceptions ontologiques liées à la théorie picturale du langage

texte.

L'affirmation que l'œuvre textuelle est identifiable, d'une manière plus ou moins complète, au texte est qualifiée de thèse textualiste. En ce sens, le textualisme relève du cas particulier de l'inscriptionnalisme appliqué aux œuvres textuelles. On peut discerner, me semble-t-il, plusieurs degrés de textualisme :

- Le textualisme radical affirme que l'œuvre est *un* objet physique particulier.
- Le textualisme fort affirme que l'œuvre est la classe d'équivalence du Texte.
- Le textualisme modéré consiste à admettre que l'équivalence des textes est une condition suffisante *mais non nécessaire* de l'identité des œuvres.

Les deux premières thèses sont relatives à la nature des œuvres (être un objet physique ou une collection d'objets physiques), la dernière est relative à l'identité des œuvres. Remarquons que le textualisme fort implique le textualisme modéré. Inversement, je qualifierai de thèse anti-textualiste l'affirmation que l'identité des textes n'est pas une condition suffisante de l'identité des œuvres, autrement dit, la thèse qu'un même texte peut produire deux œuvres différentes, et j'appellerai anti-textualisme radical la thèse que le l'œuvre textuelle ne correspond à aucun objet physique.

La crédibilité octroyée au textualisme radical peut être encouragée par certains usages langagiers. Elle peut par exemple être causée par des antonomases ou des synecdoques comme celle qui consiste à dire qu'en se procurant tel exemplaire particulier, on s'est procuré telle œuvre. De fait, on utilise parfois des expressions comme « elle m'a offert le dernier Bret Easton Ellis » ou comme « tu trouveras *La Divine Comédie* sur l'étagère du bas ». Le textualisme radical paraît cependant absurde et semble pouvoir être écarté immédiatement. Richard Wollheim souligne la faiblesse de l'argument langagier :

> Nous avons deux manières de décrire les faits : l'une en termes d'œuvre d'art, l'autre en termes de copies [...] d'œuvres d'art. Le seul fait qu'il y ait des contextes dans lesquels ces deux expressions sont interchangeables ne signifie pas qu'il n'y ait pas de contextes, encore moins de contextes décisifs, dans lesquels elles ne sont pas interchangeables. (WOLLHEIM, 2005, p. 19)

De tels contextes existent, notamment dans les situations où une reproduction du Texte est détruite. Si l'exemplaire d'*À Rebours* qui est actuellement sur l'une des étagères de mon logement venait à être brûlé, l'œuvre de Huysmans ne serait bien évidemment pas annihilée pour autant. L'objection de destruction semble donc être un argument permettant de contredire l'identification de l'œuvre à un texte particulier.

Cependant, avant d'écarter trop rapidement le textualisme radical, il convient de se demander si l'objection précédente est suffisamment efficace. Or, en utilisant l'hypothèse de la destruction d'une reproduction, nous avons peut-être mal choisi l'objet physique qui doit occuper le rôle d'œuvre dans l'argument. Parmi les choses matérielles, celle qui paraît être la plus adaptée pour occuper cette place est certainement le Texte plutôt qu'une simple reproduction, qu'il s'agisse du

manuscrit original ou du texte déclaré « bon à publier ». Renouvelons alors notre question : si, en raison de certaines circonstances, le Texte venait à disparaître, est-ce que l'œuvre disparaîtrait également ? Autrement dit, en perdant le Texte, perdrait-on aussi l'œuvre ? Cette fois, la réponse doit être nuancée. Elle serait bien sûr négative dans le cas d'un Texte dont on aurait préalablement tiré des reproductions sauves. Mais dans le cas contraire, la réponse est certainement positive. C'est bien parce que le manuscrit égaré du *Médecin Amoureux* de Molière n'a jamais été imprimé qu'on peut estimer aujourd'hui que la pièce de théâtre est perdue. Nous pourrions ainsi identifier l'œuvre au Texte seulement à la condition que ce Texte n'ait pas été reproduit. Mais est-ce véritablement un argument pour défendre le textualisme radical ? Cette affirmation contribue plus vraisemblablement à en affaiblir la thèse, puisque la destruction d'une œuvre serait entièrement dépendante de la condition que le Texte soit demeuré sans reproduction. Cela nous engage à penser que l'œuvre n'est pas identifiable au Texte, aussi important qu'il paraisse à son égard.

Si le textualisme radical échoue, ce n'est pas nécessairement parce qu'il confère une matérialité à l'œuvre, mais par sa prétention de localiser précisément l'œuvre dans tel ou tel particulier. Cela n'invalide donc pas, pour l'instant, l'hypothèse de la nature matérielle des œuvres. De surcroît, les objections précédentes au textualisme radical semblent fournir des arguments pour étayer la thèse du textualisme fort : l'œuvre n'est certes pas tel ou tel particulier, mais par contre, elle semble pouvoir être identifiée à l'ensemble complet des exemplaires. Si brûler un seul exemplaire d'une œuvre reste sans effet sur son existence, cela ne présume pas son « incombustibilité ». Car il est probable que la destruction de l'intégralité des reproduction et de leur modèle aurait bien pour effet de détruire l'œuvre. Ce genre de situation existe, citons par exemple l'épopée d'Arctinos de Milet intitulée *L'Éthiopide*, dont toutes les occurrences ont malheureusement été détruites. Bien sûr, l'argument de destruction généralisée des exemplaires ne préjuge pas d'erreurs éventuelles : nous pourrions découvrir à l'occasion d'une fouille archéologique un exemplaire subsistant. Mais cela ne serait pas un argument contradictoire à l'identité entre l'œuvre et l'ensemble de ses exemplaires. Car dans ce cas, nous aurions simplement cru à tort que l'œuvre était détruite alors qu'elle ne l'était pas.

La version forte de la thèse textualiste semble mieux résister aux objections que sa version radicale. Mais insistons encore. L'œuvre est-elle véritablement perdue lorsque le sont tous ses exemplaires textuels ? À la fin de son roman *Fahrenheit 451* (BRADBURY, 1985), Ray Bradbury évoque la situation de marginaux affairés à apprendre par cœur les plus grandes œuvres de la littérature mondiale. Leur activité est motivée par l'espoir de sauver, en les mémorisant, les œuvres littéraires menacées d'autodafés par un pouvoir politique totalitaire. La société futuriste décrite par Bradbury conduit à une forme de régression quant à la conservation et à la diffusion des œuvres littéraires. Cette régression ramène l'humanité à un stade technique antérieur à ce que Marshall McLuhan appelle la « galaxie Gutemberg », et même à un stade antérieur à l'écriture, octroyant une place prédominante à la communication orale. Le roman de Bradbury expose l'idée que l'œuvre peut persister dans la mémoire sans le secours des textes matériels. Si l'argument est plausible, conteste-t-il vraiment la nature matérielle des œuvres ? Peut-il servir de preuve à l'immatérialité des œuvres ? De la même manière que nous avons opté pour une caractérisation matérialiste des textes digitaux, nous pourrions aussi considérer ici que l'œuvre ne serait pas véritablement détruite, mais qu'elle aurait simplement changé de support, passant du papier aux neurones. Tout ce qu'on pourrait vraisemblablement inférer d'une telle situation, c'est que l'activation de l'œuvre

serait complètement dépendante de sa verbalisation plutôt que de sa lecture silencieuse.

Les besognes littéraires auxquelles s'adonnent les personnages de Fahrenheit 451 sont comparables à celles, pré-homériques, des aèdes récitant des épopées [11]. Que des inscriptions soient mémorisées plutôt que dessinées ou imprimées, elles n'en constituent pas moins une forme d'enregistrement inscrit dans le vivant. Les individus qui s'y livreraient pourraient être considérés non seulement comme des créatures vivantes, mais aussi comme des sortes d'artefacts qu'on pourrait activer pour exécuter l'œuvre, de la même manière qu'on ouvre un livre afin d'activer sa lecture. Nous aurions affaire à des hommes-livres comme Bradbury en convient d'ailleurs lui-même :

> - [...] Aimeriez-vous, un jour, Montag, lire *la République* de Platon ? - Certainement. - Je *suis* la République de Platon. Voudriez-vous lire Marc Aurèle ? Mr Simmons est Marc Aurèle. [...] Nous sommes tous constitués de morceaux d'histoire de littérature, de droit international, Byron, Tom Paine, Machiavel, le Christ, tout est inscrit. (BRADBURY, 1985, p. 175)

Comme on le voit, l'hypothèse que la mémoire des hommes puisse conserver le souvenir d'une œuvre dont tous les exemplaires matériels auraient été détruits, reviendrait à conférer la nature de texte au détenteur de sa connaissance, hypothèse de l'homme-livre. Cette éventualité n'invalide donc pas l'hypothèse que l'œuvre soit quelque chose de destructible, car, dans la situation exposée par Bradbury, si le cerveau enfermant l'enregistrement d'une œuvre venait à disparaître, l'œuvre en question serait bel et bien perdue.

Nous avons expliqué la relation de duplication des textes par la notion d'équivalence syntaxique. Nous pourrions alors peut-être identifier l'œuvre à l'ensemble complet des textes liés entre eux par des relations d'équivalence de caractères. Mais de quoi s'agirait-il au juste ? Si nous considérons qu'une classe est une entité abstraite, cela revient inévitablement à juger que l'œuvre n'est pas matérielle. Un ensemble est une chose abstraite, or on peut lire une œuvre, mais on ne peut pas lire un ensemble. Nous pourrions alors préférer dire que l'œuvre est l'extension de cet ensemble, un agrégat de textes, une somme méréologique. Notons que cette extension serait inévitablement ouverte, puisqu'elle pourrait toujours s'accroître d'un nouvel exemplaire, voire en perdre un. Richard Wollheim (WOLLHEIM, 2005, p. 21) considère qu'une telle manière de concevoir l'œuvre serait incompatible avec l'idée qu'elle soit achevée. En effet, l'œuvre serait ainsi identifiée à un processus de reproduction susceptible d'être indéfiniment prolongé. En identifiant l'œuvre à l'ensemble des copies, nous ne pourrions pas dire, au passé, que « James Joyce a écrit

[11] En outre, on peut rappeler ici que, dans l'Antiquité, les citoyens grecs ne pratiquaient pas la lecture silencieuse. La tâche de lecture, laborieuse en raison d'une absence de ponctuation dans les textes, était dévolue aux esclaves qui faisaient la lecture. On dit à ce sujet qu'Aristote faisait exception à cette règle, et que, pour cette raison, Platon le surnommait « le liseur », non sans une certaine condescendance. Retraçant l'histoire de la lecture, Orlando de Rudder souligne que « la lecture silencieuse, celle que nous pratiquons est, nous le savons, relativement récente. Dans l'Antiquité, au Moyen Âge, la lecture se faisait à haute voix. La preuve en est l'étonnement des contemporains d'Apollonius de Thyane, celui d'Augustin devant Ambroise, lorsqu'ils constatèrent que ces deux personnages hors pair ne prononcent pas ce qu'ils lisent. C'est parce que quelques lecteurs silencieux sont signalés comme des phénomènes que nous savons qu'on lisait à haute voix. » (DE RUDDER, 1983, p. 97). Concernant les traditions de diffusion orale des œuvres, on pourra se référer au livre d'Albert B. Lord, *The Singer of Tales* (LORD, 1971), inspiré lui-même des recherches ethnonarratologiques de Milman Parry au sujet de la poésie épique yougoslave. Ajoutons que considérer les corps vivants comme des supports de textes rejoint l'idée de Marshall McLuhan lorsqu'il définit, inversement, les artefacts comme des prothèses prolongeant les moyens d'action des êtres organiques. C'est aussi ce qu'indique le sous-titre de la *La Galaxie Gutemberg*, à savoir « la genèse de l'homme typographique ».

Ulysse », ou du moins, une telle expression devrait recevoir une interprétation non littérale. L'argument semble procéder selon un enthymème de la manière suivante : « l'œuvre est l'ensemble de ses exemplaires matériels. Or il est possible de produire un nouvel exemplaire. Donc l'œuvre ne peut être achevée ». Mais il semble que l'argument de Wollheim profite d'une confusion entre le temps de genèse de l'œuvre, qui doit effectivement être fini, avec la durée de sa diffusion, qui peut être indéfinie et posthume à l'auteur. Wollheim considère encore une seconde objection, plus sérieuse, à l'identification de l'œuvre à la totalité des reproductions, à savoir, de déterminer la relation adéquate qui relie les exemplaires entre eux. On pourrait commencer par répondre qu'il s'agit d'une ressemblance objective, que les reproductions matérielles sont des quasi-répliques, qu'elles présentent une similarité intrinsèque suffisamment significative pour que nous les associons les unes aux autres. Mais il reste à déterminer alors les critères de ressemblance qui, parmi une infinité de possibilités, nous permettent d'associer telle reproduction à telle autre. Le textualisme fort semble donc ici en difficulté.

Résumons-nous, nous avons considéré l'application de la thèse inscriptionnaliste aux cas particulier des œuvres textuelles, que nous avons appelé thèse textualiste. En admettant l'hypothèse intuitive que le texte noue des relations très fortes avec l'œuvre, nous avons décliné trois versions du textualisme : radicale, forte et modérée. Nous avons vu que les deux premières pouvaient être écartées, puisque les hypothèses de destructions développées plus haut indiquent ces deux thèses semblent vouées à l'échec. Reste la version modérée.

Ce que montrent vraisemblablement les arguments de destruction, c'est que les œuvres ne sont ni identifiables à un texte particulier, ni même à une collection de textes, car identifier l'œuvre à cette dernière consisterait à connaître d'emblée un critère d'identification des reproductions qui servirait aussi de critère d'identification des œuvres.

Les arguments de destruction montrent bien que les œuvres peuvent disparaître avec leurs supports matériels (textuels ou humains), mais cela n'implique pas que les œuvres soient nécessairement des objets physiques. Car les œuvres pourraient survenir sur les textes.

Ceci dit, ce n'est pas le point de vue que je souhaite défendre. Tenir l'intention de l'auteur et du récepteur à l'écart des manières d'identifier une œuvre allographique me semble de toute façon promis à de notables difficultés auxquelles ne peuvent répondre aucune théorie substantialiste, matérialiste ou abstraite, des œuvres textuelles. Je souhaite défendre l'idée que l'œuvre textuelle n'a aucun poids ontologique, qu'elle n'est rien. Pour montrer que les œuvres ne surviennent pas sur les textes, je vais donner des arguments visant à montrer que la même œuvre textuelle peut être déclinée par des textes différents, mais aussi, et réciproquement, qu'un même texte peut être associé à des œuvres textuelles distinctes.

Dans *Reconceptions en Philosophie*, Nelson Goodman et Catherine Elgin soulignent que la question de l'identité de l'œuvre requiert :

> que nous soyons au clair quant à ce qui constitue l'*identité d'un texte*. Cette identité est une affaire concernant seulement la syntaxe d'un langage - les agencements autorisés de lettres, d'espaces et de marques de ponctuation - tout à fait indépendamment de ce que le texte dit et de ce à quoi il peut référer. (GOODMAN & ELGIN, 1994, p. 58)

En défendant une individuation des œuvres textuelles, les deux auteurs soutiennent la thèse de

l'univocité entre texte et œuvre. Leur conception purement syntaxique implique que l'identité de l'œuvre ne dépende nullement de la signification littérale des textes, et encore moins de leur éventuelle interprétation métaphorique.

Chez Goodman, c'est la seule équivalence syntaxique, indépendamment de toute considération sémantique, qui permet d'identifier les textes entre eux. Nous ne devrions donc pas identifier complètement sa notion de réplique à nos dupliques, parce que celles-ci exigent un critère sémantique pour déterminer ce qui constitue la fonction d'assemblage du texte.

Le textualisme modéré de Goodman mène au critère d'identité des œuvres suivant :

x est la même œuvre que y si et seulement si le texte de x est le même que le texte de y.

Pour critiquer la condition de nécessité du critère, on montrera dans la section suivante que des textes différents peuvent être attribués à la même œuvre. Pour critiquer la condition de suffisance, on montrera dans la section 4 qu'à un même texte peuvent être attribuées des œuvres différentes.

4.3 Questions de filiations, de variations et de versions

4.3.1 Le textualisme comme essentialisme ontologique

La condition de nécessité d'identité des textes pour l'identité des œuvres relève, en quelque sorte, d'une forme d'essentialisme scriptural. Elle permet d'éviter les paradoxes sorites jouant sur la transitivité de l'identité à travers de petits changements, changements qui aboutissent en fin de course à une altération radicale. Si les conditions d'individuation d'un texte étaient telles qu'on puisse lui substituer ne serait-ce qu'un seul caractère, la succession de telles modifications engendrerait à la longue un texte n'ayant plus aucun caractère typographique en commun avec son modèle. Néanmoins, il peut sembler assez peu désirable de rejeter l'identification entre une copie et sa source sous le prétexte qu'entre les deux textes, un caractère aurait été substitué à un autre. Idéalement, le critère d'équivalence syntaxique est suffisant pour identifier les copies entre elles, puisqu'il repose sur la transitivité d'identité des classes d'équivalence syntaxique. Mais dans la réalité, les copies d'un texte sont fréquemment victimes de coquilles, autrement dit, de différences syntaxiques qui peuvent proliférer entre différentes éditions. Le textualisme fort semble accepter en tout état de cause cette conséquence, et conçoit que deux textes qui ne présenteraient qu'une légère différence syntaxique relèveraient néanmoins inévitablement d'œuvres différentes, comme l'indique d'ailleurs Goodman :

> Even remplacement of a character in a text by another synonymous character (if any can be found in a discursive language) yields a different work. (GOODMAN, 1976, p. 209)

On pourrait alors peut-être admettre que la copie d'un Texte puisse s'écarter dans une certaine mesure de sa syntaxe originale tout en continuant néanmoins à lui être volontairement identifiée. Bien sûr, il en découlerait un certain vague concernant le seuil de tolérance des modifications syntaxiques. Peut-on admettre 0,1% d'altérations, 1%, 10% ? En outre, par transitivité de l'identification des copies, on butera inévitablement sur les paradoxes de type sorite. D'un côté, admettre

la moindre erreur implique d'admettre qu'une copie puisse syntaxiquement s'avérer complètement différente de sa source tout en lui étant identifiée. D'un autre côté, refuser la moindre erreur a pour conséquence qu'il n'existe presque jamais de copies d'un Texte. Sortir de cette impasse nécessite d'introduire un critère pragmatique d'identification : tant que le nombre d'erreurs est tolérable, nous identifions la copie au Texte. Mais si le nombre de coquilles devient trop important, nous refusons. Le seuil dépend de nos intérêts et peut varier considérablement selon qu'il s'agit d'une reproduction destinée à une publication ou de la copie manuscrite faite par un collégien. Mais bien sûr, en admettant ceci, nous abandonnons l'idée que l'œuvre textuelle soit une substance [12].

En tout cas, il semble peu probable de parvenir à identifier systématiquement une œuvre littéraire à un texte particulier ou à un ensemble d'exemplaires reliés entre eux par une relation d'équivalence de caractères. Car la volonté de faire correspondre à toute œuvre une classe d'équivalence syntaxique du texte bute sur trois types de problèmes concrets. Premièrement, la question génétique de l'existence d'une source unique, deuxièmement la question temporelle des altérations et variantes du Texte et des copies, et troisièmement la question des traductions et des interprétations. Concernant la version modérée du textualisme, la question de l'interprétation montre inversement qu'il peut paraître difficile d'associer un texte à l'exclusivité d'une œuvre. Dans ce qui suit, nous examinerons successivement ces divers problèmes. L'idée est de montrer que toute tentative d'individuer de manière textualiste une œuvre littéraire par une équivalence de caractères est irrémédiablement vouée à l'échec.

4.3.2 Questions philologiques

Commençons par la question de la source. Les multiples filiations, les variations et les versions semblent indiquer qu'une œuvre puisse avoir pour source plusieurs Textes à la fois, et donc que l'équivalence de caractères n'est pas une condition nécessaire à l'identification des œuvres.

L'activité critique menée par la philologie a pour objet de rendre compte de l'authenticité des textes anciens par une remontée à leur source. L'attention est portée sur la question de la *répétition* du Texte en retraçant l'arbre généalogique de ses copies. Elle peut être motivée en cela par des raisons diverses, historiques, idéologiques, culturelles ou religieuses. Mais la remontée des copies vers la source est loin d'être un problème métaphysique qui se règle à coup de critères d'identité bien ficelés. Bien des difficultés de filiation peuvent apparaître, car l'enquête philologique indique invariablement que la production et la diffusion de copies étalées sur différentes générations engendrent inexorablement des variations, en particulier lorsque leur source s'avère ancienne. C'est ce que souligne notamment Daniel Ferrer dans son ouvrage consacré à la critique génétique :

> (...) l'origine (d'un texte) est toujours perdue, pervertie par les corruptions, il est donc impossible de la répéter simplement. Plus grave encore, il apparaît que la corruption est inhérente à la reproduction du texte. (FERRER, 2011, p. 30)

Les altérations que subit le texte peuvent parfois s'avérer exhaustives, par exemple quand un copiste croit bon de développer tout un passage qu'il juge obscur en l'illustrant abondamment. A cet égard, l'éventualité de faux ou de contrefaçons allographiques n'est sans doute pas illégitime,

[12] Sauf à admettre des substances vagues, ce qui représenterait de bien étranges entités.

notamment quand un copiste altère le texte original sans mentionner son intervention. Inversement, les dissemblances entre le texte original et ses copies peuvent également s'avérer lacunaires. Cela se produit par exemple en raison d'une censure exercée sur le texte de l'auteur, comme ce fut le cas des *Fleurs du Mal* de Baudelaire, dont le texte fut sévèrement amputé dans sa première édition. Une autre source de difficulté est relative à la citation, autrement dit, à la greffe de passages provenant d'autres textes à l'intérieur d'une œuvre. Un Texte peut comporter des parts importantes de citations, comme certains poèmes de Lautréamont qui comportent des paragraphes entiers issus d'articles encyclopédiques. Enfin, outre ces altérations intentionnelles, il peut aussi arriver des erreurs de composition ou d'impression, des mots, des pages, voire des cahiers entiers manquants.

Au long de différentes générations successives, les copies peuvent ainsi être altérées par diverses modifications volontaires et par erreur, dont la somme totale peut conduire à un texte éventuellement très différent de l'original. En outre, les altérations à des niveaux différents du processus de reproduction impliqueront des différences entre les copies elles-même. Or, comme une altération du texte à un certain niveau de la chaîne de reproduction est diffusée dans les copies qui en découlent, cela permet aux philologues de remonter la branche d'une certaine chaîne de copies par comparaison à d'autres branches dont les textes ne présentent pas les mêmes altérations. Cela signifie que, d'une manière plutôt paradoxale, le fait que des copies soient dissemblables sert ainsi à la quête de l'œuvre originale.

4.3.3 Questions génétiques

Nous venons de voir les altérations qui concernent les copies à leur modèle, ou les copies entre elles. Mais il existe une autre difficulté : les altérations du modèle lui-même, autrement dit, l'existence éventuelle de différentes versions du texte original.

À bien des égards, la carrière d'une œuvre ne commence pas avec le moment de sa reproduction. Une œuvre résulte d'un long processus d'élaboration qui nécessite de supposer une existence préalable à sa publication. Alors que la philologie focalise son attention sur les différentes phases des répétitions du Texte *après publication*, la critique génétique s'intéresse symétriquement aux différentes phases qui précèdent l'édition. Là où la première cherche à garantir l'authenticité du texte par devers les variations successives liées à sa reproduction, l'autre cherche à confronter la putative authenticité du texte par la comparaison à ses variantes génitrices. Car ce n'est pas seulement depuis un niveau post-éditorial que se pose la question de l'authentification de l'œuvre. Elle a aussi un soubassement architectonique que révèle une analyse archéologique des palimpsestes que sont les brouillons des auteurs et qui constituent autant de variations du Texte. Les effets de distorsions produites sur la succession des copies d'un texte incitent aussi, en retour, à considérer ce que la critique génétique appelle « l'avant texte », c'est-à-dire l'ensemble des textes qui forment les pré-versions, les brouillons de l'œuvre, les ratés, les raturés et les recommencements des différentes étapes de création qui précèdent l'exemplaire dit « bon à publier ».

Dans « Logiques du Brouillon », Daniel Ferrer analyse les liens qui corrèlent les œuvres publiées à leurs versions pré-éditoriales. L'auteur objecte à l'idée que l'œuvre est le texte publié que, sous cette perspective, il devient difficile de prétendre qu'un mot était différent dans une version antérieure de l'œuvre, puisqu'une telle altération, aussi minime soit-elle, manifesterait la référence à une autre œuvre plutôt qu'à une version antérieure de la *même* œuvre. Cela paraît évidemment

très contre-intuitif, et nos explications théoriques devraient pourtant demeurer fidèles à nos intuitions et à nos pratiques concernant nos usages des fictions :

> Comment peut-on même dire qu'un seul mot d'une œuvre devait être ou était différent si on considère que l'identité de l'œuvre requiert une définition mot à mot, et même, lettre à lettre ? (FERRER, 2011, p. 30)

Si les processus d'authentification philologique et génétique d'une œuvre ont pour objet d'identifier un texte princeps, un original dont ressortissent les différentes reproductions, cette entreprise peut néanmoins poser de délicats problèmes dont la résolution semble manifestement dépendre de nos décisions plutôt que d'une putative individuation métaphysique. Bien sûr, dans les cas les plus simples, les copies peuvent être identifiées à un Texte qui bénéficie d'une notoriété princeps, comme la première édition d'une œuvre qui n'a pas été remaniée ni modifiée par la suite. Mais ce n'est pas toujours l'édition originale qui délivre l'authenticité, puisqu'il arrive que le texte qui serve finalement à authentifier les copies soit celui d'une édition ultérieure. En outre, il n'est pas certain qu'un unique texte serve de source à une même œuvre. Voire même, il arrive que le texte-source n'existe tout simplement pas. *Le Roi Lear* a été publié en deux versions, en 1608 et en 1623, qui sont des retranscriptions d'une pièce jouée en 1806. Laquelle d'entre elles est l'authentique ? La plus ancienne en raison de son originalité (son antécédence temporelle) ? La seconde en raison de sa réécriture ? Ou aucune des deux en raison de la préexistence de la pièce ? Lorsque l'auteur ne donne pas de prescriptions, il est alors bien difficile de trancher. C'est ce qui fait dire à certains critiques que l'œuvre est parfois un *inaccessible*.

Outre ces premières difficultés pour individuer un Texte authentique, il n'est pas rare que des œuvres de fiction soient proposées en différentes versions par leur auteur. C'est par exemple le cas du récit de Maupassant, *Le Horla*, paru en deux versions, la première publiée en 1886 et la seconde publiée en 1887. On peut également citer le cas de Dostoïevski qui publia une première version du *Double* en 1846, puis une seconde édition profondément remaniée, 20 ans plus tard, en 1866. Ce n'est d'ailleurs pas qu'en littérature qu'on trouve des versions différentes d'une même œuvre. Remarquons que la plupart des supports à usage familial des œuvres cinématographiques (VHS, DVD, Laser Disc, Blue Ray, etc.) contiennent des versions alternatives s'ajoutant à l'enregistrement de l'œuvre commercialisée. Ces versions diffèrent entre elles de manière plus ou moins importante. Certaines modifient considérablement le montage, en ajoutant ou en retirant des scènes, allant parfois même jusqu'à modifier une partie du récit et à proposer des fins alternatives.

Il me semble qu'on aurait tort de négliger ces doubles de l'œuvre en les caractérisant comme des sortes de remplissages lucratifs. La diffusion d'une œuvre est le résultat d'un processus social complexe mêlant différents intervenants. Aussi, de nombreuses contraintes obligent le réalisateur à modifier son projet au gré du processus de création. L'une de ces contraintes relève bien sûr de la rentabilité financière qui doit accompagner la marchandisation de l'œuvre.

Les altérations qui sont responsables des différences entre les versions sont souvent la conséquence du respect de conventions sociales, idéologiques, politiques ou religieuses. Dans les pays sensibles au puritanisme comme les États-Unis d'Amérique, différentes organisations surveillent la diffusion des œuvres et en accompagnent la publication par des prescriptions d'ordre moral plus ou moins contraignantes. C'est le cas du code Hays, qui sévit à Hollywood entre les années 1934 et 1968 lors de l'avènement des grands studios américains de cinéma. Ce code constituait

un ensemble de règles morales guidant la réalisation des œuvres cinématographiques, la plus notoire d'entre elles recommandant que le film s'achève sur une fin heureuse, une « happy-end ». En somme, ce code relevait d'une sorte d'auto-censure que s'imposaient les studios, censure visant à optimiser la réussite commerciale des productions cinématographiques en ajustant, pour ainsi dire, l'offre aux mœurs présumées des spectateurs. Dans le domaine de la diffusion des œuvres musicales, le paratexte « People Advisory Explicit Music », collé ou directement imprimé sur la pochette des enregistrements sert à avertir le futur acquéreur que le contenu des paroles n'est pas conforme aux mœurs établies. Ce règlement, en cours aux États-Unis d'Amérique depuis 1985, a été introduit par une association interprofessionnelle du disque, la Recording Industry Association of America, en raison des pressions exercées par le Parents Music Resource Center. Cette dernière entendait dénoncer l'évocation des drogues, de la violence, du sexe et des attitudes anti-chrétiennes dans l'univers musical. Pour certains styles musicaux privilégiant une attitude se proclamant non conformiste, l'obtention de cette mention est d'ailleurs devenue un gage d'authenticité incontournable, qui oriente la production des textes par l'intention d'être la cible de l'avertissement, ce qui constitue, bien sûr, une nouvelle sorte d'auto-censure par effet inverse, et par conséquence, une forme évidente de conformisme.

En conséquence de telles contraintes, et afin d'en atténuer les effets, les supports de diffusion des œuvres cinématographiques présentent des versions qualifiées de « director's Cut », c'est-à-dire les montages prévus par le directeur artistique avant qu'ils ne soient modifiés par les prescriptions des producteurs. Il arrive d'ailleurs que ces versions ne soient pas cantonnées au format privilégié de l'enregistrement pour diffusion familiale, puisque certains réalisateurs, aidés en cela par la notoriété acquise et les bénéfices engrangés par la version publiée, sont parvenus à faire diffuser dans les salles de cinéma une version plus proche de celle qu'ils avaient conçue lors de la réalisation du film [13].

Ce serait certainement un grand tort de rabattre une œuvre sur l'une de ses versions plutôt qu'une autre. Une version commerciale est parfois aussi riche qu'une version « privée ». En outre, une œuvre est non seulement enrichie par les interprétations multiples qu'elle autorise, mais également par la multiplicité des versions qui sont publiées.

Dans le domaine des enregistrements sonores, le mastering désigne la création d'une matrice audio, par traitement des différentes sources successives d'enregistrements, dont il s'agit d'accorder ensuite, pour les fusionner en un enregistrement unique, les intensités et dynamiques en leur ajoutant aussi divers effets (délais, distorsions, stéréophonie, etc.). Le mastering a pour vocation de produire l'exemplaire (le Texte) dont seront ensuite tirées les copies afin d'assurer la diffusion de l'œuvre, cela de manière directe (diffusion numérique) ou après pressage pour exécution sur du matériel Hi-Fi (LP, CD, etc.). Or les différentes remarques qui précèdent semblent indiquer qu'il n'y a pas de « Master Tape », pas plus en littérature qu'au cinéma ni même qu'en musique. Il

[13] On peut par exemple dénombrer au moins sept versions officielles de « Blade Runner » de Ridley Scott (SCOTT, 1982). Parmi elles, on peut citer le prototype *Workprint* (1982) utilisé dans les phases de montage du film pour tester les réactions de l'audience et l'avant première *San Diego Sneak Preview Cut* (1982). Font suite à ces deux premières quatre versions destinées aux salles de projection : l'*U.S. Theatrical Cut* (1982) qui est connue comme la version officielle du œuvre ou comme « *Domestic Cut* », l'*International Cut* (1982), le *Director's Cut* (1991) et enfin le *Final Cut* (2007). Il existe aussi une version pour diffusion à la télévision, l'*U.S. Broadcast Cut* (1986). Outre ces sept versions, dans un documentaire sur le making-of du film, Ridley Scott, dans *Dangerous Days : The Making of Blade Runner* (2007), en mentionne encore une huitième le *Early Cut*, d'une durée de près de quatre heures et réservée au personnel du studio.

n'y a pas d'enregistrement exclusif qui assurerait de manière univoque une unicité à l'œuvre. La croyance en l'unicité du Master relève d'un mythe, preuve en est le nombre considérable d'œuvres re-mastérisées [14]. Et dire qu'il s'agit d'œuvres différentes serait évidemment absurde. Le processus de création nécessite, comme pour un algorithme mathématique, un arrêt. Mais au contraire de l'algorithme qui vise une fin objective en répondant directement à sa raison initiale, l'achèvement de l'œuvre est toujours conventionnel et institutionnel, répondant à des dates butoirs, des nécessités de diffusion. Or cet achèvement est toujours susceptible d'être remis en question, il n'est pas définitif, même de manière posthume. L'œuvre est donc une sorte de palimpseste où se recouvrent une multitude de textes, de réécritures sur réécritures, c'est-à-dire ce que Genette appelle aussi, l'intertexte. C'est notamment ce qui pousse le critique Pierre-Marc de Biasi à écrire que :

> L'idée d'un texte établi et fixé dans l'immuabilité de sa lettre est un rêve démenti par l'histoire de la tradition écrite et par les réalités de l'édition (BIASI, 2011, p. 62)

Dans *Philosophie du Rock* (POUIVET, 2010), Roger Pouivet prétend pourtant, dans le cadre de la musique rock, que l'œuvre consiste ni plus ni moins en l'enregistrement. Il faut bien reconnaître que cette hypothèse témoigne de certaines intuitions qui font du support la principale, voire parfois même, l'unique source d'exécution et d'interprétation. Mais en revanche cette hypothèse présuppose qu'il n'y ait pas, d'*avant-texte* de l'œuvre, pour employer la terminologie génétique. Les œuvres non enregistrées seraient tout simplement des non-œuvres, ce qu'il est pourtant difficile d'admettre. Un groupe de rock qui répète un morceau peut heureusement y référer avant même que celui-ci n'ait été enregistré. Si le mastering désigne bien une étape de l'œuvre, il est difficile de considérer qu'il s'agisse de son commencement. En outre, nous pourrions même considérer qu'il s'agit plutôt de son achèvement, quoique, comme nous l'avons déjà souligné, la question de l'achèvement est toujours à remettre en question [15]. Il y a donc une œuvre avant un Master, ce qui rend manifestement fausse l'égalité entre l'œuvre et le Master.

De fait, certains cas montrent qu'un enregistrement peut également suivre le processus de diffusion qui est habituellement réservé aux arts autographiques pluriels. Citons l'exemple du musicien britannique Luke Haines qui enregistra en septembre 2010 cinquante versions différentes de son album *Outsider Music*. Chacun des enregistrements fut vendu comme une œuvre unique et numérotée, à l'instar des tirages d'une même gravure. On peut également citer le titre *Morning Bell* qui figure en versions différentes sur deux albums différents du groupe Radiohead.

Pour résumer, il me semble donc qu'en esthétique sévit parfois le mythe du *Master*, c'est-à-dire, le mythe d'un original (un antécédent temporel) qui soit, aussi, un authentique (au sens

[14] Il est inexact de prétendre que la musique enregistrée échappe à l'étape de l'interprétation. Tout dispositif électro-accoustique interprète un certain nombre de paramètres accoustiques (dynamique, balance, bande passante, etc.) ayant des effets notables sur la réception de l'œuvre. Il serait, en outre, faux de dire qu'il ne s'agit pas là d'une interprétation intentionnelle mais involontaire : ce serait se méprendre complètement sur le rôle des ingénieurs du son qui conçoivent les matériels d'exécution sonore en privilégiant la dynamique de certains sons plutôt que d'autres (par exemple, le matériel Hi-Fi contemporain accorde plus de place aux basses que celui des années 70). Les mêmes remarques s'appliquent aussi aux supports de diffusion : le MP3, le LP et le CD n'interprètent pas l'œuvre de la même manière.

[15] En outre, l'analyse des enregistrements de musique rock mériterait sans doute une attention plus fine que ne le fait Pouivet à l'égard de la distinction entre les deux phases que sont le mixage des enregistrements et le mastering proprement dit, étapes qui sont temporellement séparées, et qui concernent généralement des compétences bien distinctes relevant de professions différentes.

métaphysique). Or à mon sens, une œuvre est toujours la résultante d'une diffraction de versions, elle est sémantiquement plurivoque et toujours inachevée, en raison de la profusion sans fin de ses interprétations multiples. Le *Master* relève en fait d'une décision, issue d'une pression sociale de publication, aux dépends, parfois, de l'artiste lui-même [16] : le Master, c'est le problème de la muséification des œuvres et de leur diffusion commerciale, phénomènes qui peuvent certes s'avérer indispensables au fonctionnement des œuvres, mais qui montrent aussi qu'elles sont des objets sociaux dont l'identification fluctue selon les pressions exercées par leur réception.

4.3.4 Questions de traductions, d'adaptations et de reprises

Peut-on raisonnablement espérer individuer ainsi les œuvres par des textes entendus comme chaînes de caractères ? Y a-t-il une correspondance une à une entre l'œuvre et le texte ? En particulier, quelles relations y a-t-il entre les inscriptions suivantes ?

(T_1) « Als Gregor Samsa eines Morgens aus unruhigen Träumen erwachte, fand er sich in seinem Bett zu einem ungeheuren Ungeziefer verwandelt. » (KAFKA, 1980)

(T_2) « En se réveillant un matin après des rêves agités, Gregor Samsa se retrouva, dans son lit, métamorphosé en une ignoble vermine. » (KAFKA, 1980)

(T_3) Après une nuit agitée, Gregor Samsa se découvrit, à son réveil, transformé en quelque chose d'ignoble.

Si l'on souscrit à une équivalence culturelle entre les différents exemplaires d'un texte plutôt qu'à une équivalence purement formelle, il est très difficile de renoncer alors à une caractérisation intentionnelle. Entre l'expression originale (T_1) de Kafka et la traduction française (T_2) due à Alexandre Vialatte, on a affaire à la même histoire exprimée dans deux langues différentes. On a bien sûr ici affaire à la question de la traduction d'un texte dans une autre langue. Mais il y a encore un second problème, qui apparaît avec le passage de (T_2) à (T_3) qui relève, en quelque sorte, d'une traduction dans la même langue. Cette question est par exemple celle de l'adaptation. Du point de vue du textualisme, le problème que suscitent ces deux questions est celui de savoir si deux textes relèvent nécessairement de deux œuvres différentes ou s'ils peuvent se rapporter à deux versions différentes de la même œuvre. De fait, le textualiste qui souhaite associer de manière univoque l'œuvre au texte est confronté à l'impossibilité de la traduction et de la reformulation.

La question de la traduction

Une traduction est-elle encore la même œuvre que son modèle ? Pour Gérard Genette, c'est bien le cas. En effet, Genette fait remarquer qu'on dit indifféremment qu' « on a lu une traduction française de *Guerre et Paix* » ou qu'« on a lu *Guerre et Paix* en français », voire plus simplement qu'« on a lu *Guerre et Paix* ». Les traductions auctoriales, comme celle de Vladimir Nabokov

[16] En peinture, la production de Gustave Moreau est à ce titre exemplaire. Bon nombre de ses travaux demeurent, de son propre aveu, inachevés. Mais ses toiles n'en constituent pas moins d'authentiques œuvres, exposées dans des musées et bénéficiant des mêmes soins et des mêmes intérêts que ceux dispensés à des œuvres achevées.

traduisant *La Méprise* du russe en anglais ou comme celle de Samuel Beckett traduisant *Murphy et Watt* de l'anglais au français, semblent témoigner en faveur de l'identité de l'œuvre entre les traductions. En cela, écrit Genette,

> [...] il est bien artificiel de considérer la version française et la version anglaise de *Molloy* comme deux *œuvres* distinctes plutôt que comme deux *textes* [...] de la même œuvre. (GENETTE, 1994, p. 202)

En outre, les traductions auctoriales montrent que l'auteur dispose d'un certain degré de liberté dans la traduction qu'il donne, liberté qui permet d'importantes différences avec le texte original. En ce sens, les œuvres textuelles pourraient parfois être considérées, à l'exemple des gravures, comme des œuvres multiples, inscrites par des textes distincts. Deux textes pourraient ainsi correspondre à une même œuvre, ce qui, selon Genette, donnerait tort à la réduction des œuvres aux textes que pratiquent Elgin et Goodman.

Quel statut conférer à l'œuvre ? Plutôt qu'un objet matériel, Genette y voit un objet abstrait, dépendant de celui qui y pense. Dans l'article *What is a poem ?*, Charles Stevenson (STEVENSON, 1957) a introduit la notion de mégatype afin de témoigner de l'aspect pluriel des œuvres en utilisant la distinction de Peirce entre type et token. Il définit le terme ainsi : deux tokens appartiennent au même mégatype si et seulement s'ils possèdent *approximativement* la même signification. Dès lors, deux mots-token comme le terme anglais « cat » et le terme français « chat », qui n'appartiennent pas au même mot-type, appartiennent au même mot-mégatype. Genette en reprend l'idée sous le nom d'*architype* (GENETTE, 1994, p. 128). Cette notion lui permet d'enrôler la signification par delà la seule identité syntaxique, offrant ainsi la possibilité d'authentifier entre elles différentes traductions d'un même texte. Reste alors à préciser le statut ontologique qu'on veut bien conférer à la notion d'œuvre vue comme architype. S'agit-il d'une idéalité platonicienne, d'un idéel husserlien ou d'un concept mental ? Pour Genette, l'architype « n'*existe* nulle part hors de l'esprit ou du cerveau de qui le pense » (GENETTE, 1994, p. 115). C'est ce qui fait dire à Pouivet que l'ontologie de Genette « est modeste et empirique. C'est une ontologie de régime, allégée en matières grasses métaphysiques ou phénoménologiques. » (POUIVET, 2012), et qu'elle confère un mode d'existence exclusivement pragmatique aux œuvres allographiques. Loin d'y voir une objection, il me semble au contraire qu'une telle diététique est souhaitable. Mais faut-il encore qu'elle parvienne à expliquer nos intuitions à l'égard des traductions.

Nous avons eu l'occasion, dans la première partie, de souligner l'inscrutabilité de l'individuation. Cette thèse, qui appartient à Quine, est également adjointe chez lui d'un pendant sémantique, la thèse d'indétermination de la traduction, que suit ici Goodman. Selon Quine, pour accéder au statut de science, la sémantique doit être fondée sur une base exclusivement empiriste. La méthodologie que le philosophe doit employer pour rendre compte du fonctionnement du langage est donc nécessairement behavioriste. En outre, elle doit être fondée de manière génétique sur l'étude des réactions empathiques suscitées lors de l'apprentissage d'un langage. Imaginons, propose Quine, l'étude d'une langue alors inconnue dans laquelle les indigènes utiliseraient le mot « gavagai » à chaque apparition d'un lapin. Il serait tout simplement impossible de savoir si le mot français « lapin » a ou non la même signification que « gavagai ». En effet, une certaine indétermination demeurerait inévitable : les indigènes veulent-ils désigner par ce mot des tranches temporelles de lapin, une substance abstraite comme la lapinité, des événements ou des individus ?

Cette indétermination a pour conséquence que deux traducteurs pourraient conférer des significations très différentes en étant les témoins d'actes de langage qui seraient par ailleurs complètement indiscernables.

La conclusion qu'en tirent Elgin et Goodman, c'est que la transformation syntaxique que représente la traduction d'un texte implique inévitablement la création d'une œuvre distincte de l'original. Précisons qu'ils ne contestent pas que les traductions soient véritablement des œuvres. Ce qu'ils rejettent en revanche, c'est l'idée qu'il s'agisse de la *même* œuvre. Il n'y a donc pas identification entre le texte original et sa traduction, mais une éventuelle relation de rectitude qui dépend à la fois de l'accès à l'œuvre que fournit la traduction et à la fois de la compréhension qu'elle en offre. Les traductions, bien que ressemblantes aux œuvres qui les inspire, sont des œuvres nouvelles.

Bien sûr, il reste que cette nouveauté est toute relative. Nous ne dirions pas que Baudelaire, son traducteur français, a créé *William Wilson*, même si le processus de traduction implique une certaine forme de créativité. La traduction *représente*, dans une autre langue, l'oeuvre qu'elle traduit. Le refus d'identifier ou non une traduction à son original est sans doute la conséquence de la recherche d'un critère absolu d'identité des œuvres. A l'inverse, admettre qu'une même œuvre puisse être identifiée à plusieurs textes écrits dans des langues différentes demande de déterminer alors ce qui fait alors qu'une traduction est *correcte*. Umberto Eco propose un critère poppérien de falsification consistant en la rétro-traduction, c'est-à-dire la traduction inversée, ramenant le texte traduit vers sa langue originale. Est dite correcte une traduction qui parvient à surmonter l'épreuve de la double traduction.

La question des adaptations

L'adaptation et son modèle relèvent-elles ou non de la même œuvre ? Certaines adaptations sont des réécritures transversales aux arts, qu'on peut qualifier de *transpositions* ou d'adaptations hétéro-médiatiques. Des romans ont été adaptés au théâtre, comme la pièce qu'Albert Camus a tirée des *Possédés* de Dostoïevski. Des nouvelles littéraires ont été adaptées en opéra, comme *Les contes d'Hoffmann* par Jacques Offenbach. Des partitions de ballets ont été remaniées en partitions pour orchestres, comme dans le cas de *L'Oiseau de Feu* de Stravinsky. Et bien des romans ont fait l'objet d'un film. Dans le cas d'adaptations hétéro-médiatiques, nous parlerions de traduction d'une manière sans doute plus métaphorique, tant les moyens employés peuvent s'avérer différents. Le changement de support médiatique fait surgir de nouvelles contraintes d'activation de l'œuvre, en provoquant chez le récepteur des stimulations esthétiques passant par des voies différentes. André Bazin décrit l'adaptation cinématographique d'un roman ni comme une œuvre complètement nouvelle qui viendrait s'ajouter à celle qu'elle adapte, autrement dit, ni comme une œuvre distincte de l'original, ni comme comme la même œuvre. Ni tout à fait distincte, ni tout à fait identique, Bazin parle de l'adaptation comme d'une œuvre à l'état second, une forme de déclinaison de l'original qui vient augmenter le rayonnement du roman par une forme de « produit », de fusion du roman et du film en un objet esthétique nouveau, un roman augmenté par le cinéma :

> Il ne s'agit pas ici de traduire, si fidèlement, si intelligemment que ce soit, moins encore de s'inspirer librement, avec un amoureux respect, en vue d'un film qui double l'œuvre, mais de construire sur le roman, par le cinéma, une œuvre à l'état second. Non point un film "compa-

> rable" au roman, ou "digne" de lui, mais un être esthétique nouveau qui est comme le roman multiplié par le cinéma. (BAZIN, 1958, p.126)

Les adaptations peuvent aussi consister en simple reprise d'une histoire, activité qui est bien décrite par l'anglicisme « remake » utilisé dans les arts cinématographiques, autrement dit, elles peuvent aussi consister en réécritures qui ne franchissent pas les démarcations de leur propre domaine artistique (je parlerai d'adaptations « homo-médiatiques ». Citons la pièce *Antigone*, réécrite par de nombreux auteurs, de Sophocle à Jean Anouilh en passant par Robert Garnier. Citons aussi des cas de réécritures d'œuvres par leurs propres auteurs, comme les adaptations que Paul Claudel a tirées de ses propres pièces, réputées injouables, ou comme l'adaptation pour la jeunesse *Vendredi ou la vie sauvage*, que Michel Tournier a tiré de son propre roman *Vendredi ou les Limbes du Pacifique*. Constatons que la modification du titre semble *a priori* inciter à y voir deux œuvres différentes.

On peut revenir ici sur le critère par double traduction que propose Umberto Eco, en ce qu'il débouche sur la question de la traduction d'un texte *dans sa propre langue*, ce qu'on peut aussi appeler une reformulation, qui est un cas particulier d'adaptation homo-médiatique. Si une bonne traduction de texte permet de l'identifier à la même œuvre que l'original, on peut se demander si la reformulation d'un texte dans sa propre langue peut prétendre à la même équivalence. Constatons que ce qui est préservé à travers les processus de traduction et d'adaptation serait vraisemblablement l'histoire, sans doute au détriment du style, et sans doute aussi du récit. Or s'il n'est sans doute pas raisonnable d'identifier complètement l'œuvre à la syntaxe, il n'est sans doute pas raisonnable non plus d'identifier entièrement l'œuvre à l'histoire.



Accepterions-nous qu'une même œuvre ait plusieurs inscriptions ? Si nous refusons, chaque occurrence textuelle doit alors tenue pour une œuvre différente. Or cette question ne peut être tranchée de manière métaphysique, à coup de critères d'individuation. Versions, traductions, adaptations semblent témoigner de la carrière dispersée d'œuvres fragmentées en textes multiples. Ainsi, la duplicité de l'œuvre n'est pas seulement inscrite dans son système de reproduction, mais plus profondément dans le mécanisme même des multiples phénomènes que nous regroupons sous l'appellation d'« œuvre textuelle ». La raison en est qu'une œuvre textuelle n'est rien en dehors d'un contenu intentionnel attaché à un ensemble de marques, et qu'elle est par nature inachevée, ou comme la qualifie Umberto Eco, « ouverte ». L'individuation d'une œuvre relève d'un processus social attaché à des jeux de croyance-feinte autour d'entités matérielles que sont les textes. L'œuvre n'est ni une substance individuelle matérielle ni une substance individuelle abstraite. Cela implique que la question de dénombrer des œuvres à partir de textes est vaine : une certaine interprétation peut relever ou non de la même œuvre selon la manière dont elle est décrite.

4.4 Le problème de Borges

4.4.1 Interprétations et identité

Les arguments précédents encouragent certainement à abandonner le textualisme fort. Il reste à savoir si l'on peut se contenter d'un textualisme modéré consistant à identifier l'œuvre par la condition suffisante d'identité des textes, tout en délaissant la condition de nécessité. Nous aurions ainsi la condition suivante :

Condition suffisante d'identification des œuvres : L'œuvre x est identique à l'œuvre y si le texte de x est identique au texte de y.

Si Goodman est tant soucieux d'authentifier l'œuvre à son inscription matérielle, c'est sans doute en raison de la thèse quinéenne d'indétermination des significations. Car à partir du moment où l'on fait jouer l'intentionnalité dans l'individuation de l'œuvre, celle-ci est inévitablement démultipliée par ses interprétations possibles. La variabilité des interprétations d'un texte pose des difficultés considérables au textualisme modéré. Car il n'est pas certain que l'exclusivité d'un ensemble d'inscriptions garantisse en retour l'exclusivité de l'œuvre. La question est de déterminer plus précisément les liens qui corrèlent interprétations, textes et œuvres. En particulier, dans ce qui suit, nous allons évoquer différentes situations dans lesquelles des signes se ressemblent en tout point, mais ne s'avèrent pas *au sujet* de la même chose.

Dans ce qui précède nous avons vu que plusieurs textes pouvaient être attachés à une même œuvre. La question de l'interprétation est de savoir si, inversement, deux textes indiscernables peuvent relever d'œuvres différentes. Donner une réponse affirmative à cette question remet en cause l'univocité entre l'inscription (le signe) et le signifiant. Constatons d'abord la différence entre l'interprétation du texte et l'interprétation de l'œuvre. Si le Texte et ses copies véhiculent un sens littéral, il arrive que celui-ci soit largement insuffisant pour déterminer le sens complet de l'œuvre, voire même, cela pourrait encourager le lecteur à une lecture à contre-sens par rapport aux intentions réelles de l'auteur. Car il arrive qu'un auteur écrive une chose pour en dire une autre. Que le texte ne signifie pas systématiquement ce que dit l'auteur, c'est par exemple ce dont témoignent l'usage des figures de style comme les tropes, métaphores, et autres métonymies et l'emploi de modes narratifs comme les allégories et le discours ironique. Le phénomène est généralement intentionnel, mais il peut aussi ne pas l'être.

En guise de premier exemple, mentionnons la célèbre illustration de Jastrov (figure 4.15) qui peut être lue indifféremment comme représentant un canard ou comme représentant un lapin. Un seul tracé physique occasionne ici plusieurs dessins intentionnels. Bien sûr, nous pourrions interroger son auteur : « qu'avez-vous voulu dessiner ? » Mais le dessin de Jastrov est justement réalisé avec l'intention d'évoquer plusieurs significations à la fois, offrant volontairement une double lecture. En outre, il existe des figures où l'intention de l'auteur est d'emblée écartée, comme par exemple dans le cas des figures qui servent aux tests de Rorschach, dont les dessins sont produits de manière aléatoire.

Par comparaison avec la figure du domino présentée à la figure 4.13 où ce qui est vu dépend des orientations successives de l'image, la figure du lapin-canard présente simultanément deux vues contradictoires.

Donnons encore un autre exemple, celui-ci tiré des *Recherches Philosophiques*, où Wittgen-

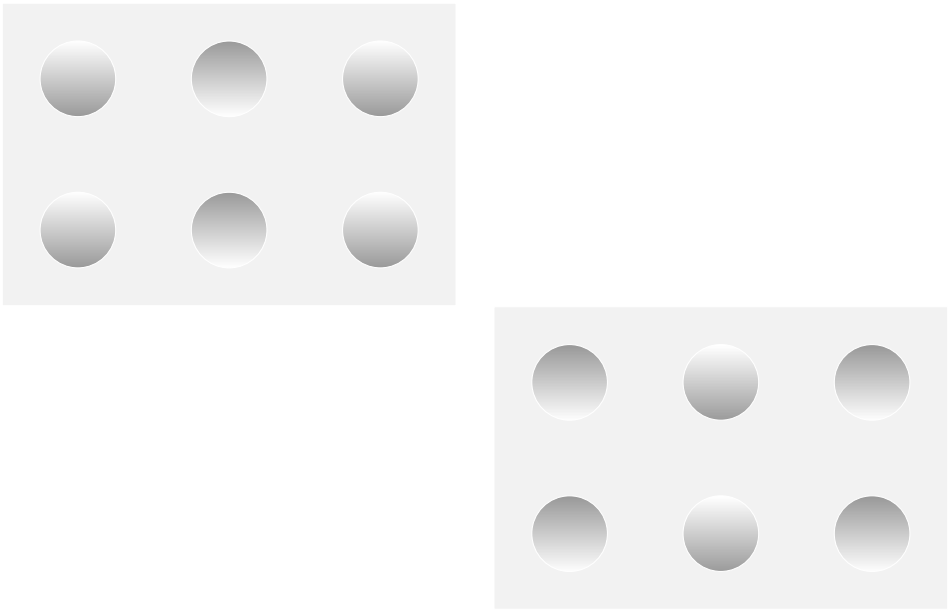


FIGURE 4.13 : illusion du domino. La même image pivotée de 180°

$$\int [f(x) + g(x)]dx = \int f(x)dx + \int g(x)dx$$

FIGURE 4.14 : inscription tribale ?

stein donne à imaginer la fiction d'une tribu qui utiliseraient les signes suivants en guise de décoration (figure 4.14) :

Il se pourrait, dans ce contexte, que le sens de l'inscription n'entretienne aucun rapport avec la propriété d'additivité des intégrales, quoique cette décoration serait graphiquement indiscernable de l'inscription mathématique correspondant à la-dite propriété.

Le problème est donc de déterminer si des textes indiscernables, au sens où ils seraient à la fois homographes, homophones et où ils fonctionneraient dans le même système linguistique pourraient ou non exemplifier des significations différentes, c'est-à-dire, des œuvres différentes.

4.4.2 Écrire une chose pour en dire une autre

Le problème de l'interprétation se rencontre à différents niveaux, et en particulier, il intervient déjà au niveau élémentaire de la syntaxe. Nous pouvons par exemple nous interroger sur les intentions d'un élève dont la copie mentionne des « gaz à effet de cerf ». Nous pourrions lui prêter, généreusement, l'intention d'une signification correcte au détriment d'une erreur de syntaxe, c'est-à-dire appliquer le principe de charité quinéen. Mais nous pourrions tout autant admettre l'absence de faute d'orthographe en faisant l'hypothèse que l'élève a précisément écrit ce qu'il entendait signifier. Cela pourrait par exemple arriver si celui-ci mêlait confusément l'écologie aux conséquences

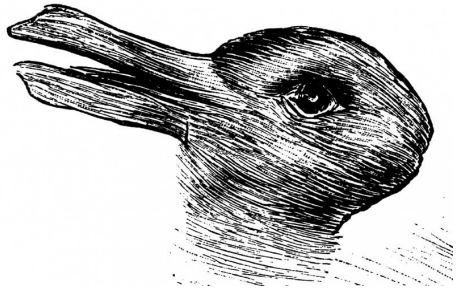


FIGURE 4.15 : canard-lapin de Jastrov

de la digestion des cervidés.

Comme le souligne Gregory Currie dans *Work and Text*, le textualisme modéré (et donc aussi sa version forte), peuvent être mis à défaut par la présence d'erreurs orthographiques et autres coquilles. Celles-ci peuvent non seulement apparaître dans les copies d'un texte, mais aussi, et pour commencer, dans la source de l'auteur. Doit-on alors, sous prétexte de fidélité syntaxique, reproduire, dans les éditions successives, les erreurs présentes dans leur source ? À ce qui semble, la plupart du temps, ceux qui sont chargés de dupliquer le Texte, par exemple les copistes et les éditeurs, prennent le soin et la responsabilité de corriger ce qu'ils estiment être des erreurs. Mais il reste à savoir si de telles modifications sont toujours pertinentes et plus particulièrement ce qui définit la notion d'erreur dans de telles situations.

Considérons un exemple pour clarifier ce problème. Admettons qu'un auteur ait rédigé, puis qu'il ait donné à publier le douteux haïku suivant :

> De retour de la Côte-d'Azur
> Sa peau devenue
> Aussi prune que des grains de café

L'éditeur en question pourrait se demander si l'intention de l'auteur n'était pas de dire tout à fait autre chose. Par exemple, l'éditeur pourrait se rappeler que l'auteur a longtemps baigné dans un milieu culturel pratiquant de manière courante l'idiolecte alsacien où le phonème correspondant à la lettre « b » est indifférencié du phonème correspondant à la lettre « p ». Il pourrait alors en déduire que ce que l'auteur avait *réellement* l'intention d'énoncer, c'est que la peau avait plutôt pris une couleur *brune*. Dans ce cas, les copies du texte devraient corriger l'erreur en mentionnant « brune » plutôt que « prune ». Mais d'un autre côté, la phrase pourrait être tout à fait appropriée, elle pourrait être fidèle à l'intention sémantique de son auteur. L'objectif serait par exemple d'amuser le lecteur en moquant les travers orthographiques de quelque touriste alsacien, ou même de produire de la prose surréaliste d'influence éluardienne connotant, au passage, la célèbre Terre « bleue comme une orange ». Et il pourrait aussi s'agir d'une courte poésie de science-fiction décrivant un monde où les rayons ultra-violets auraient un effet manifestement différent sur la pigmentation de la peau. Dans tous ces cas, le Texte ne devrait évidemment pas être modifié lors de sa duplication.

Ce que manifestent de telles situations, c'est qu'une erreur, à l'intérieur d'un texte, peut pro-

duire une œuvre très différente de l'intention originale de l'auteur. Il faut alors en déduire que le critère de fidélité syntaxique n'est pas un critère suffisant pour l'individuation d'une œuvre. Dans l'exemple précédent, le texte donné peut relever d'une œuvre réaliste ou d'une œuvre de science-fiction selon qu'on admet la présence d'une erreur typographique. La condition d'équivalence syntaxique des copies n'est donc pas suffisante, et doit être complétée par un critère sémantique tenant compte de ce que l'auteur voulait *réellement* épeler plutôt que ce qu'il a effectivement épelé lorsqu'il existe une différence. Ce qu'on appelle ainsi « erreur » résulte de deux intentions conflictuelles de l'auteur, la première étant d'épeler de manière appropriée un terme, la deuxième étant de l'épeler de manière erronée avec l'attitude de croire qu'il s'agisse de la bonne manière d'épeler le terme. Dans ce cas, l'intention d'épeler doit être considérée comme dominante, et l'éditeur devrait modifier le texte pour rester fidèle aux intentions sémantiques plutôt que syntaxiques de l'auteur. C'est notamment ce que constate Currie :

> The author intends to spell correctly, and intends to spell this way because she believes this way is correct. So the intention to spell correctly is dominant, and we respect her intention by correcting. Whether or not intentions conflict, it is intention that determines the text. (CURRIE, 1991, p. 326)

C'est donc l'intention de l'auteur qui détermine la signification du texte, et non pas la seule inscription matérielle.

Soulignons que l'appréciation de la présence d'une erreur peut demeurer hésitante. Dans les cas les plus simples, on peut toujours faire appel à l'auteur pour qu'il précise son intention véritable, ce qui a pour effet de trancher définitivement la question. Mais bien sûr, ce n'est possible que du vivant de l'auteur. L'éventuelle impossibilité de connaître directement l'intention de l'auteur justifie alors le travail de généalogie qui est l'affaire de la philologie. Il peut arriver cependant que, même s'il est possible de les connaître, les intentions de l'auteur ne permettent pas vraiment de résoudre ce conflit, autrement dit, qu'elles ne permettent pas de prescrire le texte qui convient à l'œuvre. Soit que, quoique suffisamment précises, ces intentions ne permettent néanmoins pas de déterminer quelle séquence syntaxique est la plus pertinente, soit qu'elles s'avèrent vagues et imprécises, ce qui suffirait déjà pour désolidariser les textes des œuvres. Dans ces cas, Currie souligne que :

> Undominated conflict among, or indeterminacy of, authorial intention creates a multiplicity of textual candidates for the title of work, with no thing to choose between them. (CURRIE, 1989, p. 326)

On voit que l'œuvre ne peut plus être identifiée à une classe de textes syntaxiquement équivalents, mais à toute une classe de textes pouvant manifester des différences syntaxiques.

Comme on le voit, dans certaines circonstances, le Texte comme manifestation d'inscriptions doit encore être dédoublé d'un autre Texte, celui-ci sauf d'erreurs, un Texte qui témoignerait de la véritable volonté de l'auteur par delà les erreurs éventuelles. Les textes ne doivent donc pas être passivement recopiés tels quels, mais parfois corrigés lorsque se présente une erreur. Or le jugement d'erreur fait appel à l'intention véritable de l'auteur. Si un auteur voulait signifier « brune » en écrivant « prune », il faut alors que figurent dans les copies du Texte la première mention plutôt

que la deuxième, parce que la signification qui est véritablement visée n'est pas celle du second terme. Comme on le constate, on passe ainsi à une conception qui réinjecte l'intentionnalité dans l'identification des textes. C'est ce qui conduit Currie à proposer une condition suffisante d'identification du Texte qui n'est plus seulement syntaxique mais sémantico-syntaxique :

> I offer as a sufficient condition that the two tokens have the same semantic and syntactic properties (Roughly, they should consist of words, in the same order that mean the same thing and are spelled the sameway). (CURRIE, 1991, p. 325)

Ainsi, Currie distingue-t-il le *vrai* texte du texte effectif qui ne lui correspond pas toujours. Mais plutôt que d'affirmer qu'un texte est plus vrai qu'un autre, on peut aussi concevoir deux modes différents de textes : le texte matériel consistant en l'ensemble des inscriptions, et le texte intentionnel comme ensemble des significations qui lui sont attachées.

Goodman considère que l'identification d'un texte ne nécessite de le voir que comme une unité syntaxique. Or nous avons vu que des textes différents peuvent être associés à la même œuvre, ce qui constitue une véritable difficulté pour le textualisme.

4.4.3 Le cas Ménard

Le constat d'indétermination de la traduction mène Quine à affirmer que non seulement celle-ci vaut pour deux langues différentes, mais qu'elle vaut aussi pour une seule et même langue. Comme le dit Quine, quand on y réfléchit bien, la traduction radicale commence « *at home* » (W. V. O. QUINE, 1977, p. 58). Pour s'en convaincre, il suffit d'imaginer qu'un terme soit traduit selon un dictionnaire dans le langage indigène, puis retraduit dans la langue d'origine avec l'aide d'un second dictionnaire attribuant un sens différent. La conséquence de ceci est que, quand deux personnes emploient les mêmes termes, il est bien possible qu'elles aient exactement les mêmes comportements de communication, sans nécessairement y attacher la même signification. D'ailleurs, souligne Quine, lorsque nous sommes en désaccord avec l'usage fait d'une certaine suite homophonique utilisée par un individu à qui nous concédons cependant le principe de charité sémantique, nous échangeons cette suite de sons au profit d'une autre qui nous paraît plus légitime pour véhiculer la signification imputée au message, « nous reconstruisons ses propos hétérophoniquement » (W. V. O. QUINE, 1977, p. 58).

La question de l'interprétation semble particulièrement articulée à la question des indiscernables. Dans la nouvelle intitulée *Pierre Ménard, auteur du Don Quichotte* (BORGES, 1957), Jorge Luis Borges conçoit le cas d'un écrivain symboliste du XXe siècle, Pierre Ménard, qui se propose de produire une œuvre littéraire à partir de la reproduction du *Quichotte* de Cervantes. Soulignons d'emblée que l'objectif de Ménard n'est pas de plagier Cervantes. Il n'en est pas question ici, et ne peut l'être parce que Ménard cite explicitement la source du texte dont l'évidente notoriété entrave inexorablement toute velléité de plagiat. En outre, Borges précise que l'entreprise que se fixe Ménard ne consiste ni à produire une imitation du roman de Cervantes, qui ne serait au mieux qu'un ersatz du *Quichotte*, ni à recopier le texte de Cervantes, ce qui reviendrait seulement à produire une nouvelle occurrence du *même* texte. Pour parvenir à ses fins, Ménard vise plutôt le projet, non d'une *reproduction*, mais d'une *production* fragmentaire et absolument indiscernable du texte du *Quichotte* :

> Il ne voulait pas composer un autre Quichotte - ce qui est facile - mais *le Quichotte*. Inutile d'ajouter qu'il n'envisagea jamais une transcription mécanique de l'original ; il ne se proposait pas de le copier. Son admirable ambition était de reproduire quelques pages qui coïncideraient - mot à mot et ligne à ligne - avec celles de Miguel de Cervantes. (BORGES, 1983, p. 45)

Ainsi, toute l'ingéniosité de la fiction de Borges provient du développement qu'il tire de l'hypothèse qu'un auteur puisse produire une œuvre nouvelle à partir du texte d'une œuvre plus ancienne. Le texte de Cervantes et celui de Ménard, quoique indiscernables, n'en semblent pas moins produire des œuvres numériquement différentes.

Considérons les détails de l'histoire de Borges. La première méthode que tente d'employer Ménard pour élaborer sa création littéraire consiste à reproduire les intentions de Cervantes, à prendre la place de l'auteur. Cela demande évidemment d'oublier tous les faits culturels et historiques survenus depuis la création du Quichotte par Cervantes. Cela contraint également Ménard à acquérir toutes les connaissances de Cervantes, en particulier sa maîtrise de l'espagnol du début du XVIIe siècle. Mais prenant acte de l'impossibilité physique d'être Cervantes à la place de Ménard, Ménard conçoit alors une autre méthode, encore plus improbable et donc, dit-il, encore plus intéressante, consistant à être Cervantes tout en demeurant Ménard.

Remarquons que ce que partagent les œuvres respectives de Cervantes et de Ménard, ce ne sont pas des caractéristiques relativement contingentes comme la typographie ou la mise en page, mais ce qui apparaît comme une caractéristique essentielle d'une œuvre : son identité scripturale. Ainsi, Borges nous convie à nous interroger : un même texte peut-il vraisemblablement servir à identifier deux œuvres différentes ? Ou pour le dire encore d'une autre façon : peut-on éviter de dupliquer l'œuvre quand on duplique son texte ? Si le textualisme permet de mettre de côté le rôle des intentions, autant celle de l'auteur que celle du lecteur, avec le problème de Ménard, on voit qu'il nous faut réinjecter la question du double au niveau de l'œuvre, c'est-à-dire qu'il nous faut réinscrire le rôle de l'intentionnalité au cœur du problème qui nous occupe.

Certains auteurs assimilent le problème de Ménard à celui de la traduction, que nous avons déjà évoqué précédemment. C'est par exemple ce qu'affirme George Steiner qui reconnaît dans le récit de Borges le commentaire « le plus dense et le plus percutant portant sur la traduction » (STEINER, 1998, p. 76) constatant qu'au fond, tous les discours sur la traduction ne sont que des commentaires de la fiction de Borges. Maurice Blanchot rapproche également le problème de Ménard de celui de la traduction, mais contrairement à Steiner, il en souligne le caractère inverse :

> Lorsque Borges nous propose d'imaginer un écrivain français contemporain écrivant, à partir des pensées qui lui sont propres, quelques pages qui reproduiront textuellement deux chapitres de *Don Quichotte*, cette absurdité mémorable n'est rien d'autre que ce qui s'accomplit dans toute traduction. Dans une traduction, nous avons la même œuvre en un double langage ; dans la fiction de Borges, nous avons deux œuvres dans l'identité du même langage et, dans cette identité qui n'en est pas une, le fascinant mirage de la duplicité des possibles. (BLANCHOT, 1959, p. 133)

D'une manière un peu obscure mais néanmoins saillante, Blanchot fait donc mention du caractère modal des œuvres. Le double est ainsi ancré au texte, et il annule l'effet d'originalité.

Quelles conséquences le problème de Ménard a t-il sur l'ontologie des œuvres ? S'il constitue une évidente contradiction au textualisme fort et modéré, il sert aussi d'argument contre les ontologies esthétiques stratifiées. Car ce qu'indique le paradoxe de Borges, c'est non seulement que l'œuvre n'est pas réductible au texte mais aussi qu'*elle ne survient pas sur le texte*. La thèse de survenance affirme en effet que si deux objets x et y sont esthétiquement différents, il doit alors exister une autre différence, celle-ci non-esthétique, permettant d'expliquer leur différence esthétique. Autrement dit, toute différence esthétique s'explique par une différence non-esthétique. Et de manière converse, si x est non-esthétiquement indiscernable de y alors x est esthétiquement indiscernable de y. Soulignons toutefois que, selon les partisans de l'émergentisme, il n'est pas systématiquement exigé qu'un caractère explicatif soit conféré à la différence sous-jacente.

La fiction de Borges mérite sans doute qu'on s'y attarde encore, en l'examinant plus précisément afin de savoir si l'on peut tirer une quelconque validité logique de sa concevabilité. La fiction de Borges décrit-elle une expérience de pensée du double ? Autrement dit, fournit-elle, dans le domaine des objets littéraires, un contre-exemple au principe d'identité des indiscernables par l'exemplification de deux choses indiscernables sans incohérence logique ? On pourrait y voir ici un exemple de ce que nous avons appelé, du point de vue métaphysique, des coïncidents purs. En effet, un seul et même texte ferait coïncider (au moins) deux œuvres différentes.

Considérons le problème d'un peu plus près encore. L'auteur insiste sur l'idée que le texte de Ménard, bien qu'indiscernable de celui de Cervantes, s'en distingue cependant en tant qu'œuvre. Quels sont ses arguments ? Borges en donne essentiellement deux. Le premier, c'est que la réception du *Quichotte* de Ménard par le lecteur peut se faire à travers des grilles d'interprétations qui l'enrichissent et qui sont inaccessibles à l'œuvre de Cervantes. Les composantes thématiques du texte de Ménard ne sont pas identifiables à celles du texte de Cervantes. Ils diffèrent en genre. Le récit de Cervantes peut être interprété comme le récit d'une opposition entre l'idéal chevaleresque médiéval et la société où vit son auteur. Le texte de Ménard relève quant à lui du genre de la fiction historique, décrivant de manière rétrospective l'Espagne de la bataille de Lépante. La seconde différence entre les deux œuvres relève des composantes herméneutiques qui divergent. Le fait que ces auteurs aient des carrières temporellement éloignées implique un changement de contexte d'énonciation qui enrichit la lecture du texte de Ménard par toutes les circonstances historiques, toutes les créations littéraires et tous les concepts philosophiques accumulés entre les époques respectives des deux auteurs. Borges constate ainsi que le récit de Ménard peut être lu par le filtre critique de l'influence probable qu'ont exercée sur lui les récits de Lope de Vega, les essais de Friedrich Nietzsche, la psychologie de William James et la logique de Bertrand Russell. Ce qui amène Borges à conclure que « le texte de Cervantes et celui de Ménard sont verbalement identiques, mais le second est presque infiniment plus riche. » (BORGES, 1983, p. 49) Troisièmement, la distinction numérique des œuvres n'est pas seulement due à des différences thématiques et herméneutiques mais aussi stylistiques. En effet, les expressions employées par Ménard avec une certaine lourdeur relève d'un style archaïsant et donc audacieux, alors que le langage que manie avec souplesse Cervantes s'avère au contraire la langue courante de son époque. Enfin, Jean-Marie Schaeffer (SCHAEFFER, 1989, p. 133) complète la liste d'arguments de Borges par une troisième différence, générique : alors que le livre de Cervantes est un anti-roman ou une parodie de roman de chevalerie, le livre écrit par Ménard relève autant du roman psychologique, historique, métaphysique que parodique.

Peut-être faut-il conclure avec Schaeffer que le problème de Borges ne

> démontre en fait une chose toute simple : dans la mesure où un texte n'est pas seulement une chaîne syntaxique mais aussi et surtout un acte communicationnel, l'identité à travers le temps de la chaîne syntaxique ne garantit pas son identité comme message. L'identité syntaxique est sans doute nécessaire, mais elle n'est pas suffisante : pour qu'il y ait identité de message, il doit y avoir aussi identité de contexte. (SCHAEFFER, 1989, p. 134)

Menant son analyse des fictions à partir d'un point de vue cognitiviste, Schaeffer souligne l'importance de la notion de message qui est au cœur de la théorie de la communication. Le texte est un signal, et le contexte historique, changeant, occasionne sur celui-ci des variations d'interprétations. Avec le temps, le message s'engrossit de différentes nouvelles significations que permettent les connaissances acquises. Et inversement, il dégrossit aussi de tous les savoirs qui ne sont plus accessibles.

Simplifions la situation exposée par Borges en considérant des textes composés d'un unique mot. Admettons par exemple deux textes consistant en deux séries de caractères orthographiquement indiscernables comme « chat » et « chat ». Les deux textes sont syntaxiquement identiques. Devons-nous pourtant leur associer la même signification ? Certes non, car considérés en fonction des systèmes linguistiques français et anglais, les deux textes n'ont pas la même signification. Quoique les deux inscriptions apparaissent indiscernables, les deux mots désignent en effet, dans ces langues, des choses différentes, à savoir, respectivement un félidé et un bavardage. On constate en outre que l'homonymie n'est d'ailleurs pas exclusivement graphique. Deux textes peuvent s'avérer holorymes sans être homographes, comme par exemple « Un vieil armagnac » et « un vieillard maniaque » qui sont des expressions homophones, et comme le sont aussi, « sain » et « saint ». L'indiscernabilité sonore n'implique pas toujours une identité sémantique. En outre, dans certains cas, l'homophonie peut s'avérer relever de deux idiomes différents, comme dans : « Lucullus apportavit legato aurum » et « Lucullus apporta vite le gâteau au rhum ».

Afin de défendre leur position strictement inscriptionnaliste, Goodman et Elgin font usage de la technique d'épellation pour manipuler les inscriptions. Cela permet de distinguer les homophones n'ayant pas la même graphie, comme l'illustrent les exemples précédents. Bien sûr, l'usage de l'épellation n'est cependant pas systématiquement pertinente, comme l'indiquait déjà le cas des morphèmes français « chat » et anglais « chat ». Mais Elgin et Goodman contestent que la différence entre ces deux morphèmes soit exclusivement d'ordre sémantique. Autrement dit, ils réfutent l'indiscernabilité présumée des deux morphèmes. En effet, Elgin et Goodman commencent par constater la différence phonologique qui les distingue. Les deux mots, bien que homographes, ne sont néanmoins pas homophones : « chat » ne se prononce pas, en français, de la même manière que « chat » en anglais. Or cette différence n'est pas négligeable car les répliques d'un mot ne sont pas seulement des inscriptions écrites et imprimées mais aussi des instances sonores.

Goodman et Elgin indiquent également un deuxième argument contre les indiscernables sémiotiques, qui relève encore du rejet de la prétendue indiscernabilité entre les deux occurrences de « chat ». Selon les auteurs, bien que les deux mots soient constitués exactement des mêmes lettres ordonnées exactement dans le même ordre, leur fonctionnement syntaxique diffère néanmoins selon le système linguistique dans lequel ils sont employés. En effet, en les comparant, on peut noter une variation dans leur aptitude à être combinés à d'autres morphèmes du même

système linguistique notamment quand on élabore des textes plus longs. Chez Goodman et Elgin, ces différences modales résultent donc d'une conception holistique de la syntaxe d'un langage. La mêmeté des termes ne peut être attribuée par la seule épellation « c-h-a-t » indépendamment du système linguistique dans lequel on trouve cette épellation.

Aussi, les usages linguistiques différents de l'inscription « chat » révèlent que les deux mots ne sont qu'apparemment, mais pas *véritablement*, syntaxiquement indiscernables. Il ne s'agit donc pas, ici, exclusivement d'une différence sémantique, mais véritablement, selon les auteurs, d'une discernabilité syntaxique :

> C'est *un mot différent, un texte différent*, dans les deux langages. Tout cela est pure affaire de syntaxe ; la distinction est indépendante de ce que les mots signifient ou de ce à quoi ils réfèrent. (GOODMAN & ELGIN, 1994, p. 59)

On pourrait objecter que certains mots comme « bar » existent à la fois en anglais et en français et semblent avoir la même signification dans les deux langues.

Dans l'article *On Likeness of Meaning* (GOODMAN, 1949) Goodman affirme qu'il faut distinguer pour chaque terme deux types d'extensions. L'extension primaire d'un terme est son extension habituelle. L'extension secondaire d'un terme est l'extension de n'importe quelle expression composée contenant ce terme. Selon Goodman, deux expressions ont la même signification si et seulement si elles ont les mêmes extensions primaire *et* secondaire (GOODMAN, 1949, p. 5). Une telle perspective a pour conséquence immédiate un rejet des synonymes, car leur existence présupposerait qu'ils aient à la fois la même extension primaire et secondaire. Or si les mots « bar » (anglais) et « bar » (français) ont la même extension primaire, à savoir un bar, ils n'ont sans doute pas la même extension secondaire. Le caractère holistique de la langue implique que la signification d'un terme est enrichi de l'ensemble des phrases dans lequel il peut intervenir, et ces ensembles peuvent s'avérer différents dans deux langues véhiculées par deux cultures différentes. Mettant à part les considérations d'homophonie, les deux auteurs appliquent d'ailleurs leur critère de distinction syntaxique aux cas de mots à la fois homographes *et* synonymes, comme l'expression « permission » qui a la même référence en français et en anglais. Elgin et Goodman sont amenés à rejeter l'idée que l'indiscernabilité des chaînes de caractères suffisent à elle seule à déterminer l'identité des textes. D'après eux, l'hypothèse d'indiscernabilité ne pourrait être validée que si les textes *fonctionnaient* dans le même système linguistique. Puisqu'un texte est toujours inscrit dans un langage, son identité syntaxique dépend entièrement de ce langage. Aussi, Elgin et Goodman concluent-ils que :

> « Chat » en anglais et en français ne consiste donc pas exactement dans le même groupe de répliques ou d'instances, même si les classes d'inscriptions coïncident. (GOODMAN & ELGIN, 1994, p. 59)

Même si les arguments d'indiscernabilité sont assez répandus dans la littérature consacrée au problème de Borges, ils ne sont sans doute pas très efficaces. Une œuvre n'est de fait jamais restreinte à un mot, et sur un texte long, on retrouve généralement la même signification littérale entre deux expressions identiques. L'argument d'homonymie entre deux langages différents peut de toute façon être écarté dans la mesure où, pour un texte long, la signification littérale est exclusive à un système linguistique, autrement dit, le cas de traduction dans des langues différentes

n'apparaît pas. Comment faire intervenir un tel critère d'individuation autrement qu'en souscrivant à un caractère intentionnel des textes ? Dire que les extensions secondaires sous-tendent la différence linguistiques, c'est admettre que ces différences sont le fruit d'habitudes et de convenances culturelles nécessairement ancrées sur l'intention.

4.4.4 Des textes homographes ?

Toute la question de Borges n'est peut-être qu'un problème d'homonymie appliqué à un texte entier plutôt qu'à un unique mot. Goodman et Elgin proposent inversement de considérer un texte minimal constitué de l'exclusif mot « serviette » [17]. De fait, on peut prêter au mot deux significations différentes, puisqu'il signifie à la fois un morceau de tissu destiné à la toilette et une sorte de sac destiné à véhiculer des documents. Les termes qui signifient ces choses sont donc simultanément parfaitement homographes et parfaitement homonymes, autrement dit, ils partagent une parfaite indiscernabilité sémiotique. Toute la question est alors de savoir si l'on doit considérer qu'on a affaire à un mot (texte) unique ayant deux interprétations possibles ou au contraire si on a affaire à deux mots (textes) ayant chacun une unique interprétation. Les auteurs répondent que le mot est certes ambigu par le choix de ses interprétations possibles, mais néanmoins unique. Autrement dit, Goodman et Elgin appliquent, sans autre justification, une sorte de version sémiotique du principe d'identité des indiscernables : si les signifiants sont indiscernables, c'est donc qu'on a affaire à un seul signe.

Un texte peut exprimer diverses significations, comme un nuage, d'après Hamlet, différentes dénotations :

> Hamlet : Voyez-vous ce nuage là-bas qui a presque la forme d'un chameau ?
> Polonius : Par la messe ! on dirait que c'est un chameau, vraiment.
> Hamlet : Je le prendrais pour une belette.
> Polonius : Oui, il est tourné comme une belette.
> Hamlet : Ou comme une baleine.
> Polonius : Tout à fait comme une baleine.

Si le même texte écrit par deux auteurs différents peut recevoir deux sens différents, on peut aussi se poser la question de savoir si le même texte d'un même auteur peut être reçu de deux manières différentes par un même lecteur. Le problème de l'indiscernabilité des œuvres est aussi celui de la diversité des expériences esthétiques qui sont tirées du même objet physique, et qui semble indiquer le caractère profondément subjectif de l'expérience esthétique. Sur une durée de vingt années, le texte d'une œuvre de fiction sera resté le même. En revanche, un homme qui relit, à 60 ans, une œuvre qu'il avait déjà lue à vingt ans, aura bien changé dans l'intervalle. Sa sensibilité ne sera certainement plus la même. Il ne réagira plus de la même façon au texte. L'expérience vécue modifiera inévitablement son appréhension esthétique de l'œuvre et la compréhension qu'il en tirera. Comme l'observe Étienne Gilson dans *Peinture et Réalité*,

> Bien peu restent fidèles aux admirations de leur jeunesse. L'homme mûr ne cherche plus

[17] Il s'agit de la traduction française. Dans le texte anglais, c'est le mot « cape » qui est utilisé.

> dans les (œuvres) ce qu'il leur demandait enfant ou jeune homme et s'il continue d'aimer les mêmes œuvres, ce n'est plus pour les mêmes raisons. (GILSON, 1972, p. 28)

Considérons encore un autre argument contre l'unicité de l'œuvre. Certains philosophes ont déduit de la multiplicité des interprétations possibles l'hypothèse que la signification n'est pas inhérente au texte. La radicalisation de cette thèse a été prônée par les théories déconstructionnistes pour lesquels l'œuvre peut prendre n'importe quelle signification, la plus absurde soit-elle. Cette thèse rend compte de la possibilité de significations incompatibles et de la nécessité pour une œuvre de passer par le stade de l'interprétation pour acquérir un statut. Mais en revanche, elle explique mal pourquoi nous refuserions certaines interprétations plutôt que d'autres. Il est vrai que n'importe qui peut considérer qu'une inscription donnée signifie n'importe quoi. Je pourrais admettre que l'inscription « Longtemps je me suis couché de bonne heure » signifie quelque chose comme « Attention ! Arrivée imminente d'extra-terrestres belliqueux ! ». Et lorsque Tolkien utilise le mot « dragon », je pourrais y voir la référence à telle personnalité de mauvais tempérament. Il s'agirait manifestement d'une interprétation, mais ce ne serait évidemment ni une interprétation *correcte* d'*A la recherche du temps perdu* ni une interprétation *correcte* du *Seigneur des Anneaux*.

La défense du caractère extrinsèque de la signification n'exige pas de rejeter le critère de pertinence interprétative. De fait, ce critère ne peut être attribué ni à l'auteur, ni au lecteur. Il doit donc être affranchi de la seule subjectivité, ce qui n'induit pas d'en faire pour autant une objectivité au sens réaliste du terme. Il revient simplement de constater que la construction de la signification d'une œuvre n'est pas un jeu solipsiste que le lecteur joue avec lui-même : il ne s'agit pas d'extrapoler des formes à partir de masses nuageuses au hasard d'une rêverie contemplative. Il s'agit de considérer que la validité d'une interprétation est le fruit d'une élaboration sociale impliquant différents acteurs, de l'auteur aux lecteurs en passant par les critiques, qui chacun participent à l'élaboration de la signification. Cela explique ainsi que si toute interprétation relève en une part de la subjectivité, le processus social que représentent les jeux de fiction implique aussi des composantes inter-subjectives. Les raisons de croire qu'une interprétation est bonne ou non sont relatives à des systèmes en dehors desquels de tels jugements s'avèrent injustifiables. Il n'y a donc pas de critères de pertinence trans-historiques, ils sont toujours dépendants de causes psycho-sociologiques. Est-ce à dire que les interprétations ne sont jamais guidées par ce que les critiques savent ou croient être vrai ? Si, à condition d'admettre un réalisme intentionnel, autrement dit au sens où les intentions doxastiques jouent un rôle causal dans le comportement des individus. C'est notamment ce que remarque Charles Stevenson :

> Dans quelle mesure au juste une telle connaissance peut-elle « guider » une décision ? Nous ne devons pas supposer que cela implique une relation logique. Car une décision est *processus* au cours duquel certaines façons de répondre aux œuvres d'art sont acceptées et d'autres rejetées. Il ne s'agit donc pas de l'assertion d'une proposition [...]. Et c'est seulement si une proposition avait été impliquée que l'on aurait pu dire de façon sensée que la décision entretenait une relation logique avec la connaissance. La relation en question est bien plutôt *causale* [...]. Comme n'importe quel processus psychologique, la décision critique a de nombreuses causes ; et parmi ces causes, nous devons inclure les croyances du critique qui [...] l'influencent, à la fois conativement et affectivement. (STEVENSON, 1957, p. 359)

Ouvrir l'œuvre au subjectivisme des multiples interprétations des récepteurs n'implique donc

pas nécessairement qu'il faille légitimer une pratique déconstructviste ni un relativisme postmoderne en acceptant libéralement que le texte puisse référer à n'importe quoi. Dans *Les limites de l'interprétation*, Umberto Eco relativise la théorie de l'œuvre ouverte à laquelle il a lui-même fortement contribué, écrivant que :

> Avant qu'un texte ne soit produit, on pourrait inventer toutes sortes de textes. Après qu'un texte a été produit, il est possible de lui faire dire beaucoup de choses – parfois, un nombre potentiellement infini de choses –, mais il est impossible, ou du moins illégitime d'un point de vue critique – de lui faire dire ce qu'il ne dit pas. Souvent les textes disent plus que ce que leur auteurs entendaient dire, mais moins que ce que beaucoup de lecteurs incontinents voudraient qu'ils disent. (ECO, 1992, p. 130)

Le sens d'un texte n'est pas seulement défini par une *intentio auctoris* et par une *intentio lectoris*, mais également par une *intentio operis*, par ce que le texte dit *littéralement* indépendamment des phantasmes des uns et des autres. Un texte est un objet social dont la signification ne peut être entièrement restreinte à un solipsisme psychologique.

Afin de répondre plus amplement au problème de Borges dans le cas d'un texte qui ne serait pas restreint à un unique mot, Goodman et Elgin proposent leur propre expérience de pensée, imaginant que deux jumeaux produisent chacun un texte décrivant des faits différents, mais avec des termes indiscernables :

> Deux jumeaux identiques, ayant vécu étroitement ensemble depuis plusieurs années, furent envoyés pour l'été dans des camps similaires, mais distincts, en Nouvelle-Angleterre. Après coup, on demanda à chacun de rédiger un bref rapport sur ses expériences. Alors même que les jumeaux parlaient de personnes, d'endroits et d'événements distincts, leurs rapports furent composés d'exactement les mêmes séries de mots. Ne serait-ce pas un cas manifeste de deux œuvres pour un unique texte ? (GOODMAN & ELGIN, 1994, p. 60-61)

Selon Goodman et Elgin, et contrairement aux apparences, on a en fait affaire à un texte unique mais recevant différentes lectures possibles. Une interprétation ne définit donc pas une œuvre, mais relève d'une manière de lire l'œuvre, c'est-à-dire le texte. En inscrivant l'œuvre dans et par le texte, Goodman et Elgin admettent qu'il y ait différentes interprétations correctes d'une même œuvre, mais maintiennent fermement l'identité entre le texte et l'œuvre. Mais cela exige alors de réfuter certaines des différences que Borges attribue aux deux textes du *Quichotte*. Elgin et Goodman défendent ainsi le principe d'identité des indiscernables en attaquant l'hypothèse des différences censées diversifier numériquement les textes de Cervantes et de Ménard. Mais ni les différences herméneutiques ni les différences historiques ne semblent menacer la thèse inscriptionnaliste, car de telles différences peuvent s'expliquer de manière interprétative en maintenant l'identité numérique du texte. C'est essentiellement les différences stylistiques qui posent un évident problème à Elgin et Goodman : les deux textes doivent avoir les mêmes composantes syntaxiques, donc stylistiques. Car le style ne peut survenir que sur les inscriptions. Deux styles différents, c'est donc, nécessairement, deux textes différents. Plus précisément encore, la différence d'interprétation ne peut expliquer la différence de style.

Ainsi, ce sont seulement les différences de style que Borges prétend découvrir entre les deux textes qui sont récusées par Elgin et Goodman. Mais il n'est, selon ces deux auteurs, qu'un seul

texte, avec diverses interprétations possibles, certaines correctes, et d'autres non. Les textes ne sont pas des coïncidents purs, autrement dit, des choses indiscernables, y compris de manière spatio-temporelle. Elgin et Goodman font donc appel au principe d'identité des indiscernables en réfutant les différences syntaxiques entre les textes de Ménard et de Cervantes.

> S'il est incorrect, pour un lecteur contemporain, d'interpréter Cervantes comme archaïque, il est également incorrect d'interpréter ainsi celui de Ménard. Car les « deux » textes n'en font qu'un. Tout comme le mot « serviette », malgré ses deux applications, est un seul mot et pas deux, *Don Quichotte*, malgré ses multiples interprétations admissibles est une seule œuvre et pas plusieurs. (GOODMAN & ELGIN, 1994, p. 63)

Parce que les auteurs identifient œuvre et texte, ils placent l'interprétation en marge de la question de l'identité, comme un phénomène qui ne la concerne qu'indirectement. Car comme l'écrit Jacques Morizot dans son essai *Sur le problème de Borges* :

> Si la diversité des interprétations ne remet pas en cause l'unicité du texte, inversement, l'identité scripturale ne supprime pas l'hétérogénéité des œuvres. [...] Que pourrait signifier de parler d'« œuvres » si l'on n'accorde pas le moindre intérêt à ce qui démarque des situations aussi dissemblables ? (MORIZOT, 1999, p. 83)

Notons en particulier qu'en identifiant texte et œuvre, Goodman et Elgin dissocient la question de l'identité de l'œuvre de celle de son authenticité. L'attribution d'un texte à un auteur est une autre question que celle de l'identité des textes, c'est une question historique et juridique. L'ayant droit d'un texte est le premier producteur reconnu de son inscription. En cas de producteurs synchrones, il s'agit soit d'une collaboration (pluralité volontaire des producteurs) soit d'une création à auteurs multiples (pluralité involontaire des producteurs). Mais, encore une fois, Elgin et Goodman insistent pour dire que l'authentification n'a rien à voir avec l'identité. Pour eux, l'identité de l'œuvre est exclusivement une question d'identité syntaxique du texte. Leur stratégie leur permet de défendre un textualisme modéré en utilisant l'identité syntaxique des textes comme condition suffisante de l'identité des œuvres.

Résumons-nous. Nous avons vu que la volonté d'exorciser l'intentionnalisme hors de conditions d'identification des textes est confrontée au problème des textes synonymes (traductions, adaptations, reprises) et au problème des homonymes (interprétations divergentes, métaphores, contenu implicite, etc.). Nous pouvons encore soulever un dernier problème qui est celui du statut de textes obtenus de manière non intentionnelle. En effet, la faiblesse du textualisme provient principalement du fait qu'il écarte complètement le rôle de l'auteur et du lecteur dans la création de l'œuvre. L'œuvre vue comme un enregistrement de marques est complètement désolidarisée de son mode de production historique, de sa cause efficiente et de son activation. Si l'on pousse alors un peu plus loin le textualisme dans ses retranchements, on peut se demander s'il est vraiment nécessaire, de ce point de vue, qu'un texte soit issu d'un acte humain de création pour mériter le statut d'œuvre.

En effet, il peut être intéressant de compléter le problème mis en exergue par Borges en indiquant une variante pour laquelle le statut d'auteur ne peut, en aucune manière, intervenir. Imaginons, comme nous invite Émile Borel dans le compte rendu de l'Académie des Sciences intitulé

« Sur les probabilités universellement négligeables » (BOREL, 1972), des singes tapant au hasard sur des machines à écrire, ou, de manière plus contemporaine, imaginons qu'un programme informatique produise de manière aléatoire des chaînes de caractères alphabétiques [18]. Si l'activité s'exerce sur une durée infinie, il est certain que le singe et que le programme produiront une copie de n'importe quel texte existant, par exemple le texte du *Quichotte*. Mais devrait-on dès lors qualifier d'œuvre le texte qui résulterait d'un tel processus purement mécanique ? Si on suit les prescriptions inscriptionnalistes de Goodman, il faudrait répondre *oui*, car le fait de partager les mêmes caractéristiques syntaxiques devrait impliquer l'identité des œuvres, c'est-à-dire conférer *ipso facto* le statut d'œuvre au texte obtenu de manière automatisée.

Il est vrai que rendre plausible la production mécanique du texte de Cervantes exige de disposer d'une durée de frappe plus longue que celle qu'on attribue à l'existence de l'univers. Mais nous pourrions aussi considérer la possibilité qu'un unique mot, mettons « eau », soit obtenu par un programme tirant trois fois de suite une lettre au hasard parmi les vingt-six lettres que compte l'alphabet. La probabilité d'obtenir de mot en question est alors de $\frac{1}{26^3}$, ce qui, pour un microprocesseur, serait très plausible et même quasi-certain au bout d'une durée très raisonnable. Nous pouvons donc insister sur la question : l'inscription obtenue serait-elle ou non le même texte que celui qui apparaît dans la phrase « L'eau est incolore, inodore et insipide » ? L'inscription ne prend le sens de mot que si nous lui accordons un tel rang. Mais en dehors de cette intention qui soutient son statut, elle n'aurait pas plus de sens que « uae » ou que « eee ».

A propos du problème de Borges, Elgin et Goodman concluaient que

> [...] les deux œuvres supposées n'en font qu'une [...] Ménard produit simplement une autre inscription du texte. N'importe lequel d'entre nous peut en faire autant, des presses d'imprimerie ou des photocopieurs aussi. On affirme que, si des singes en nombre infini tapaient sur des machines à écrire pendant une durée infinie, l'un d'entre eux, probablement, produirait une réplique du texte. Nous maintenons que cette réplique serait tout autant une instance du *Don Quichotte* que le manuscrit de Cervantes, que celui de Ménard et que chaque exemplaire du livre qui a été ou sera imprimé. Que le singe soit supposé avoir produit l'exemplaire par hasard ne change rien au problème. C'est le même texte, on peut en faire les mêmes interprétations que celles des exemplaires rédigés intentionnellement par Cervantes, Ménard et les divers copistes et typographes qui ont fait des exemplaires de l'œuvre. Ce qui relève de l'intention ou de l'intelligence du producteur d'une inscription donnée n'a aucune importance pour l'identité de l'œuvre... (GOODMAN & ELGIN, 1994, p. 272)

Quelles différences y a-t-il entre le problème de Ménard et celui des singes dactylographes et

[18] Le scénario de la production aléatoire d'un texte a différentes sources. Il apparaît notamment dans *Les voyages de Gulliver* (1721) où Jonathan Swift dépeint un professeur qui tente de générer de manière aléatoire la liste complète des connaissances scientifiques. Des étudiants sont employés à tourner des manivelles d'une machine destinée à concaténer des mots inscrits sur les différentes faces de petits cubes de bois reliés au mécanisme (SWIFT, 1965, p. 193). On peut aussi citer la nouvelle de Jorge Luis Borges, *La Bibliothèque de Babel*, qui mentionne un nombre illimité de livres remplis de chaînes aléatoires de caractères, dont la réunion contiendrait inévitablement toutes les grandes œuvres de la littérature mondiale, passées, présentes, et futures, en plus de textes médiocres, absurdes et inintelligibles. Hilary Putnam propose, quant à lui, l'hypothèse de lignes tracées par une fourmi qui marche sur le sable (PUTNAM, 1984, p. 61), dont la composition ressemblerait accidentellement à une caricature. Enfin, on peut citer aussi le jeu de société *Boggle* où des lettres alphabétiques sont inscrites sur les faces de dés qui sont lancés et que les joueurs doivent ordonner différemment dans le but d'obtenir le plus grand nombre possible de mots.

le détournement qu'opère Marcel Duchamp modifiant le statut d'un objet banal pour l'introduire dans le monde de l'art ? Le cas de Ménard pourrait exemplifier une sorte de *ready made* littéraire. En créant « Fontaine », Duchamp confère à l'artefact ordinaire qu'est un urinoir la nature d'œuvre d'art. Ce n'est pas le cas de Ménard, qui, en produisant un texte en tout point similaire, ne remet pas en question la nature de l'« entité » qu'il considère, puisqu'il s'agit d'emblée d'une œuvre littéraire éminemment reconnue comme telle. Le texte de Cervantes et celui de Ménard, quoique indiscernables, pourraient néanmoins être lus comme des œuvres très distinctes : un roman médiéval d'un coté et une espèce de performance, un « gag conceptuel », comme l'écrit Genette (GENETTE, 1994, p. 280), de l'autre.

Tout ceci montre qu'on ne peut faire l'économie des intentions dans le fait de caractériser des mots à partir d'inscriptions. Notre point de départ était de considérer qu'un texte peut être décrit comme un ensemble d'inscriptions physiques, mais nous voyons que nous ne pouvons nous en contenter, et qu'il faut aussi concéder qu'il puisse être décrit de manière intentionnelle.

La richesse de l'hypothèse de Borges tient justement à ce qu'elle tente de faire coïncider texte écrit et texte lu, c'est-à-dire de corréler l'œuvre à son exécution. Aussi, selon Genette, le problème de Borges peut se ramener à celui de la part créatrice de l'exécution :

> À chaque lecture, chaque livre est mentalement « réécrit » par son lecteur comme Ménard ré-écrivit le Quichotte. Ainsi, l'infatigable fable borgésienne est peut-être moins une parabole sophistiquée de la littérature, qu'une description fidèle et somme toute évidente de l'acte de lire. (GENETTE, 2016)

En rejetant l'intentionnalité, Goodman fait une pétition de principe en faveur de l'univocité reliant texte et œuvre. Mais comme ne cesse de le souligner Genette à la fin de *L'œuvre de l'art*, Elgin et Goodman ne démontrent à aucun moment qu'il n'y a qu'un seul texte plutôt que deux. Ils appliquent le principe d'identité des indiscernables sans discuter le comment et le pourquoi de cette application, se contentant de réfuter la possibilité logique de la fiction de Borges. Mais comme le souligne Genette (GENETTE, 1994, p. 273), puisque Elgin et Goodman n'ont pas démontré qu'il est faux d'associer un même texte à plusieurs œuvres, cette hypothèse reste probable.

La réécriture de l'œuvre de Cervantes par Ménard encourage aussi, de manière plus générale, le détournement d'auteur.

> Ménard (peut-être sans le vouloir) a enrichi l'art figé et rudimentaire de la lecture par une technique nouvelle : la technique de l'anachronisme délibéré et des attributions erronées. Cette technique aux applications infinies, nous invite à parcourir *L'Odyssée* comme si elle était postérieure à l'Énéide [...] attribuer l'*Imitation de Jésus Christ* à Louis-Ferdinand Céline, ou à James Joyce, n'est-ce pas renouveler suffisamment les minces conseils spirituels de cet ouvrage ? (BORGES, 1983, p. 52)

Eddy Zemach proteste que si l'on pouvait réattribuer ainsi n'importe quelle œuvre ancienne, on devrait alors créer des chaires universitaire de littérature anachronique (ZEMACH, 2005, p. 124). Certains auteurs prennent cependant cette invitation tout à fait au sérieux, comme en témoigne par exemple l'essai *Et si les œuvres changeaient d'auteurs ?* (BAYARD, 2010) de Pierre Bayard, où la notion d'auteur devient une fonction tributaire de celle de lecteur. C'est d'ailleurs ce que reconnaît

Barthes, quand il dit que « la naissance du lecteur doit se payer de la mort de l'auteur » (BARTHES, 1984). La place du lecteur prend une importance capitale dans la production de l'œuvre, qui est, *in fine*, création du lecteur. Les lectures d'un texte s'engrossissent nécessairement des contextes historiques et esthétiques qu'elles traversent. S'appuyant sur un court vers de Jean Racine situé dans l'acte I scène IV de *Bérénice*, « Dans l'Orient désert, quel devient mon ennui ! », Valéry commente :

> [...] l'accord magnifique de ces trois mots, quand le temps le transporte et le fait traverser le XIXe siècle, trouve un renforcement inattendu et une résonance extraordinaire dans la poésie romantique ; dans une âme de notre époque, il se mélange merveilleusement à quelques-uns des plus beaux vers de Baudelaire. Il se détache d'Antiochus, il prend une généralité pure et nostalgique. Son élégance finie se transforme en beauté infinie. (VALÉRY, 1957, p. 495)

L'œuvre de Racine bénéficie donc de celle de Baudelaire, comme si, au fond, Racine avait été influencé par Baudelaire, ou mieux, comme si Baudelaire était le véritable auteur du vers en question. Mais n'allons pas plus loin creuser ces paradoxes temporels. Il nous suffit de constater que l'application du principe d'identité des indiscernables aux inscriptions est certainement abusif.

4.4.5 Deux manières d'identifier l'œuvre textuelle

Ce que montrent les commentaires qui précèdent, c'est l'importance de la prise en compte de l'intentionnalité dans l'identification de l'œuvre. Les tentatives consistant à individuer les œuvres seulement par leurs inscriptions matérielles ne sont pas satisfaisantes, parce qu'elles ne donnent pas la place qui revient à l'intentionnalité dans le processus d'identification. L'œuvre n'est pas le texte et elle ne survient pas sur le texte. Les significations débordent le texte. Vouloir réduire les œuvres aux enregistrements, comme le voudraient Elgin et Goodman, est donc irrémédiablement voué à l'échec. Etant donné que Goodman et Elgin privilégient la question de l'individuation de l'œuvre par ses reproductions à celle de l'authentification des reproductions, le statut qu'ils confèrent aux reproductions d'un Texte est plus proche de celui des répliques (reproductions non intentionnelles) que de celui des dupliques (reproductions intentionnelles). Or, si le nominalisme à tendance métaphysique qu'ils défendent peut se passer d'entités idéales en considérant les œuvres allographiques comme des collections de répliques associées entre elles par des équivalences purement syntaxiques, l'inconvénient d'une telle association réside dans son caractère par trop rigide : des textes légèrement différents comme des versions corrigées sont inévitablement identifiés à des œuvres différentes.

Les difficultés auxquelles se heurte le textualisme proviennent sans doute de la volonté d'individuer les œuvres par des propriétés seulement intrinsèques des textes. Le cas de Borges décrit un unique texte qui est écrit deux fois. Le fait que Ménard ne se contente pas de réimprimer le texte, mais qu'il le « raconte » une nouvelle fois est tout à fait important. Car comme le souligne David Lewis, deux actes de raconter la même histoire impliqueront inévitablement deux fictions différentes : « différents sont les actes de raconter une histoire, différentes sont les fictions ». Un même texte peut donc relever de deux œuvres différentes. En outre, Lewis constate qu'il y a autant de fictions que d'interprétations : « un acte de raconter une histoire peut, cependant, être le récit de deux fictions différentes : l'une est une fantaisie inoffensive racontée aux enfants et aux censeurs,

l'autre une allégorie subversive racontée simultanément aux *cognoscenti* » (LEWIS, 2012, p. 42).

L'importance de l'intervention de l'auteur pourrait inciter à faire reposer l'individuation des œuvres en y introduisant des critères extrinsèques. Si l'on ne rejette pas d'emblée l'hypothèse que des textes strictement indiscernables puissent relever d'œuvres différentes, il peut alors être judicieux d'introduire, entre le texte et l'œuvre, un critère d'identité génétique, qu'on pourrait qualifier de composition. A cette fin, considérons un triplet $\langle \mathcal{T}, \mathrm{A}, t_0 \rangle$ constitué d'un certain exemplaire de texte $\mathcal{T}$, d'un agent A, et d'un moment temporel t_0. Le triplet est interprété ainsi : l'occurrence $\mathcal{T}$ provient d'un certain auteur A qui l'a rédigée à un certain moment t_0. La composition permet alors une authentification des textes par un moyen externe, en individualisant les textes et en les reliant à des artefacts concrets comme des manuscrits, des copies, des éditions. C'est notamment le genre d'idée que suit Amie Thomasson dans *Fiction and Metaphysics* pour garantir les conditions d'authenticité d'un texte (THOMASSON, 2011, p. 92–93). Focalisant au niveau du personnage plutôt qu'à celui de l'œuvre, elle donne en effet raison à Borges concernant l'hypothèse Ménard :

> Je maintiens que si un certain Fred Jones écrivait, par pure coïncidence (c'est-à-dire sans connaître l'œuvre de Richardson ou de Fielding), une pièce sur une jeune servante nommé « Pamela », sa Pamela serait un personnage différent de celui de Richardson - dont il serait, au mieux, extrêmement semblable. Et il n'en irait pas autrement, même si Jones en venait, par le plus grand des hasards, à écrire exactement les mêmes mots et dans le même ordre que ceux du récit de Richardson. (THOMASSON, 2011, p. 117)

À la différence du Ménard de Borges, le Jones de Thomasson ne doit pas connaître les auteurs qui le précèdent parce que la théorie de Thomasson associe œuvre et composition, c'est-à-dire qu'une œuvre est une création qui apparaît à un certain moment. Le fait, pour Jones, de connaître la composition de Richardson impliquerait alors l'identité des textes, ce qui n'arrive pas lorsque Jones créé le même texte que Richardson sans le connaître. Dans ce cas, poursuit Thomasson, un contenu de pensée au sujet de Pamela pourrait aussi bien concerner celui de Richardson que celui de Jones. La théorie du pur contenu intentionnel n'est pas en mesure de distinguer deux actes intentionnels dont l'un vise la Pamela de Jones et l'autre la Pamela de Richardson. Les deux ayant le même contenu, ils seraient identiques. Aussi Thomasson en conclut que

> Les solutions sur la base de l'individuation par la conscience sont [...] incapables d'affronter (des) cas où tout se passe comme s'il y avait deux ou plusieurs expériences du même objet : le cas de deux individus différents qui pensent à Hamlet, et celui d'un individu qui pense à Hamlet à deux reprises, sans se souvenir de son expérience précédente. (THOMASSON, 2011, p. 116–117)

La solution de Thomasson passe, nous l'avons déjà vu dans le chapitre précédent, par une métaphysique créationniste considérant les œuvres textuelles comme des artefacts abstraits. Sans céder à cet élan métaphysique, Gérard Genette introduit de son côté deux « modes » d'existence des œuvres. Le premier mode est désigné par le terme d'*immanence*, quand il s'agit d'ancrer les œuvres à des composantes physiques, à des textes matériels. L'autre mode d'existence est la transcendance, quand il s'agit de considérer que l'œuvre comme un *irréel*, un pur contenu mental. Il recouvre, précise Genette, « toutes les manières fort diverses et nullement exclusives les unes des

autres, dont une œuvre peut brouiller ou déborder la relation qu'elle entretient avec l'objet matériel ou idéal en lequel, fondamentalement, elle "consiste", tous les cas où s'introduit une sorte de "jeu" entre l'œuvre et son objet d'immanence ». Mais les cas de transcendance peuvent eux-mêmes être regroupés en trois modes, que Genette décrit selon trois formules-type. Le premier mode est la transcendance par *pluralité d'immanence*, dont la formule est n *objets d'immanence pour 1 œuvre*. Il englobe les phénomènes de traductions, de variantes, etc. que nous avons relevés dans les sections précédentes, lorsqu'une œuvre est immanente de textes non identiques et concurrents. Le deuxième mode de transcendance, appelé *transcendance par partialité* et dont la formule est $\frac{1}{n}$ *objet d'immanence pour 1 œuvre* correspond au cas où l'œuvre se manifeste de manière fragmentaire, comme dans le cas du *Château*, le récit inachevé de Kafka. Le troisième et dernier mode, qualifié de *transcendance par pluralité opérale*, est résumé par la formule *1 objet d'immanence pour* n *œuvres*. Ici, un seul objet d'immanence détermine plusieurs œuvres en fonction des contextes de réception. C'est le cas qui recouvre le problème de Borges, remettant en cause l'application du principe d'identité des indiscernables pour l'individuation des textes.

Le mérite de la conception de l'identification bifide de Genette est bien sûr de recouvrir tous les cas où l'œuvre textuelle n'est pas réductible à ses inscriptions, mêmes les cas les plus limites comme celui du Ménard de Borges. Mais cela exige d'abandonner l'idée d'individuer les œuvres par une seule et même source métaphysique (l'immanence), au profit de différents modes d'identification dépendant des intentions. La notion de duplique exige d'identifier un texte de deux façons différentes, d'une part d'une manière physique référant à certains objets matériels couverts d'inscriptions physiques, mais également selon une manière toute intentionnelle. Cette dissociation entre deux voies de la reproduction, physique et intentionnelle, peut être rapprochée de ce que Hintikka a introduit dans la logique de la perception [19], adaptation de la logique modale aux cas des verbes désignant des attitudes intentionnelles comme voir, etc.

Contrairement aux répliques pures, les dupliques nécessitent un double schéma d'identification, l'un dont les marqueurs sont des ressemblances intrinsèques de formes (les caractères du texte) et l'autre des critères pragmatiques extrinsèques et culturels. Deux reproductions d'une même inscription matérielle ne sont donc pas nécessairement deux dupliques, au sens où elles peuvent, selon le contexte intentionnel considéré, fonctionner de manières différentes, autrement dit, ne pas articuler les mêmes significations. La duplication d'une œuvre est un processus relevant donc, en réalité, de deux sources distinctes. La première consiste en l'opération de reproduction d'un artefact physique, par imitation des ressemblances syntaxiques (à savoir, les chaînes de caractères). La seconde consiste à véhiculer certaines significations et représentations associées à cet objet. Ainsi, le cas du *Don Quichotte* de Pierre Ménard met à jour la situation paroxystique visant à reproduire l'artefact physique du texte sans toutefois reproduire l'objet intentionnel associé.

4.5 Conclusion

Dans ce chapitre, nous avons montré les difficultés afférentes à l'individuation métaphysique des textes. Plusieurs textes peuvent être associés à la même œuvre textuelle et plusieurs œuvres textuelles au même texte. Nous avons vu qu'une notion fonctionnelle de dupliques pouvait servir de

[19] Concernant les verbes transitifs intentionnels, on pourra consulter (REBUSCHI, 2018).

base à l'identification des œuvres. L'œuvre textuelle peut ainsi être identifiée par un double processus. D'une part physiquement sur la base d'une ressemblance scripturale, et de l'autre sur la base de relations extrinsèques intentionnelles.

La conclusion que j'en tire est que les œuvres textuelles ne sont ni des inscriptions matérielles (qu'il s'agisse de textes particuliers ou d'ensemble de textes), ni des substances immatérielles émergeant sur des particuliers concrets (les textes). L'une ou l'autre de ces thèses conduisent inévitablement à rabattre les occurrences textuelles sur des répliques pures, autrement dit, sur des entités intrinsèquement indiscernables. En affirmant qu'elles ne sont pas des substances, cela ne présume cependant pas de soutenir que les œuvres ne sont rien (nihilisme). Les œuvres devraient plutôt être considérées comme des performances. En ce sens, elles peuvent être assimilées à des fonctions d'assemblage qui régissent l'authentification des dupliques d'un texte dans un contexte de pratiques culturelles.

A mon sens, nous devrions conclure qu'il n'y a pas de réponse absolue à la question de savoir si une œuvre est la même qu'une autre. La réponse dépend de nos attentes concernant « l'assemblage de significations » que nous accordons à un texte donné.

En introduisant les dupliques, on n'envisage plus les ressemblances comme étant d'ordre naturel, mais comme des constructions culturelles qui dépendent de sélections proprement humaines. Ces sélections sont motivées de manière fonctionnelle, par des critères contextuels de pertinence.

Après avoir traité la question de l'identification des textes, il nous reste maintenant à aborder de front l'identification des personnages de fiction, ce que nous allons nous employer à faire dans le chapitre suivant.

Chapitre 5

Le personnage et ses doubles : concevoir le double en fiction

> M. Goliadkine avait tout à fait reconnu son nocturne compagnon. Son nocturne compagnon n'était que lui-même… M. Goliadkine lui-même, un autre M. Goliadkine, mais tout à fait identique à lui-même… en un mot ce qui s'appelle son double sous tous les rapports…
>
> Fodor Dostoïevski, *Le Double* (DOSTOÏEVSKI, 1969)

L'objet de ce chapitre sera de mettre en exergue les paradoxes qui émergent de l'individuation métaphysique des personnages. Son contenu peut être résumé par la question suivante : quelle dose de ressemblance la représentation fictionnelle peut-elle supporter ? Nous voilà donc ici arrivés au cœur de notre problématique, concernant la tension occasionnée entre l'imitation et la représentation. L'enjeu est de montrer que la conception du double ne nécessite pas d'engagement ontologique envers des entités fictionnelles, et plus encore que des engagements ontologiques engendrent diverses difficultés qui seront discutées.

Dans le chapitre précédent, nous avons vu ce que nous pouvions entendre par copie d'une œuvre. En particulier, nous avons soutenu que l'identification d'une œuvre consiste en un double procès, l'un physique, l'autre, intentionnel. Nous avons discuté de l'identité d'une fiction, considérons maintenant l'identité *dans* la fiction. Demandons-nous en quoi consiste au juste l'identification d'un personnage. Ici, le procès physique semble d'emblée en défaut, puisqu'il n'y a pas d'individu matériel sur lequel directement faire reposer l'identification. Le problème est alors d'expliquer ce qui rend possible à différents actes intentionnels de viser *un seul et même* personnage, aptitude dont témoignent nos discours fictionnels ordinaires. La question est celle de l'identification d'un même personnage par différentes personnes, mais aussi celle de son identification par une seule et même personne à des moments différents.

Les théories meinongiennes procèdent, nous l'avons dit, par une réponse hyper-réaliste : les personnages sont, littéralement, des objets inexistants, des *ficta*. Il en découle la possibilité de copier de tels objets, soit comme des objets récurrents entre des œuvres différentes, soit comme des personnages indiscernables au sein d'une même œuvre. Apparaît-il, alors, dans les théories

réalistes qui soutiennent l'individuation métaphysique des personnages, des difficultés qui leur sont propres ?

D'emblée, nous pouvons constater qu'il y a deux problèmes inverses pour le partisan des *ficta*. Le premier est de parvenir à caractériser deux personnages par deux objets fictionnels quand un auteur nous invite à les concevoir indiscernables en dépit des descriptions nécessairement différentes qui se dégagent du texte - par exemple par le schéma narratif focalisant sur l'un des personnages plutôt que sur son *doppelgänger*. Le deuxième problème qui se pose pour l'ami des *ficta* est de réussir l'opération inverse, à savoir de parvenir à distinguer des personnages par le biais de *ficta* lorsque les descriptions sont insuffisantes pour les discerner alors que le lecteur reçoit l'invitation implicite d'imaginer qu'ils ne sont pas indiscernables. Par ailleurs, ces problèmes peuvent se retrouver à plusieurs niveaux. Au niveau intrafictionnel quand on se préoccupe d'une seule et unique fiction, au niveau interfictionnel quand il s'agit d'identifier ou non des personnages à travers des fictions différentes. Pour résumer, voici les différents types de questions qui émanent de l'identification des personnages de fiction :

(i) Problème des personnages indiscernables (William Wilson et son double, les membres anonymes de foules, etc.).

(ii) Problème des personnages récurrents (Sherlock Holmes d'*Une étude en rouge* (DOYLE, 1979b) et celui du *Signe des quatre* (DOYLE, 1979a), Hercule Poirot d'Agatha Christie et celui de Sophie Hannah (HANNAH, 2014)).

Je souhaite défendre la thèse que les personnages de fiction ne sont pas des objets fictionnels, ni au sens d'entités extra-mentales, ni au sens d'entités métaphysiquement objectives dépendant d'actes mentaux. Les personnages ont rapport à des processus sociaux-culturels, c'est-à-dire à des actes impliquant des intentions collectives et institutionnalisées. Comme l'écrit John Searle, les personnages sont des « faits institutionnels », ils « n'existent qu'à l'intérieur de systèmes de règles constitutives » (J. SEARLE, 1998, p. 46). Leur identification est le fruit de négociations consensuelles entre les récepteurs des fictions plutôt que la résultante d'une réification extra-mentale produite par un acte quasi-magique de stipulation auctoriale. La théorie des artefacts développée par Amie Thomasson à partir des travaux de Roman Ingarden autour des relations de dépendance ontologique pourrait paraître assez proche de cette conception si elle n'inversait pas le rapport entre l'intentionnalité et l'objectivité : là où Thomasson affirme que ce sont les contenus fictionnels qui individuent les actes intentionnels, je défends au contraire l'idée que ce sont les actes intentionnels qui prescrivent l'identification des contenus.

La question de l'individuation des personnages n'est pas métaphysique, elle n'est pas de savoir si, réellement, le personnage d'Hercule Poirot dans la fiction de Sophie Hannah est le même que celui des fictions d'Agatha Christie. La question est de savoir si nous sommes d'accord pour jouer un certain type de jeu de croyance feinte dans lequel nous admettons, intersubjectivement, l'identification de l'un à l'autre. Cela ne veut pas dire que tout vaut pour n'importe quoi, que nous pouvons identifier Averell Dalton à Hercule Poirot, puisque nos identifications sont régies par des consensus culturels, par l'élaboration collective des règles des jeux de croyance-feinte auxquels nous participons à travers l'immersion fictionnelle. La question de l'identification n'est pas de

savoir ce qui individue tel personnage, mais de déterminer ce qui particularise tel jeu de croyance-feinte. Or c'est bien à la critique littéraire, dans le sens de communauté de réception esthétique, que cette tâche est dévolue.

En contextualisant les vérités par rapport aux mondes, nous échappons à une théorie mimétique de la vérité, permettant ainsi de rendre leur caractère informatif aux énoncés fictionnels tout en admettant que les noms qu'ils contiennent ne réfèrent à rien de réel. Mais il reste encore à considérer comment concevoir le domaine de tels mondes. Pour le meinongianisme modal défendu par Francesco Berto et pour le nonéisme de Graham Priest, le domaine est uniformément le même, sinon que, d'un monde à l'autre, le prédicat d'existence est attribué ou non à l'entité visée. Mais les mêmes objets subsistent dans tous les mondes possibles. Cela nécessite donc à nouveau de souscrire à un réalisme fort des objets inexistants, et c'est en cela que ces stratégies peuvent encore être rattachées à la théorie de l'objet.

Ce n'est pas la voie que je suivrai ici, mais une autre conception, également attachée à la théorie des modèles, mais celle-ci d'obédience anti-réaliste, considérant que les domaines des mondes possibles sont variés, qu'ils sont différents de monde à monde. Cela permet un désengagement du vrai en rapport au réel. Ce qui est vrai n'est plus nécessairement ce qui est réel. En outre, je fais l'hypothèse que nos usages d'identifications en fiction ne relèvent pas de critères métaphysiques d'individuation, mais de pratiques discursives et critiques. La distinction numérique entre personnages n'est pas quelque chose que l'on *constate* passivement (une fois qu'on en aurait acquis le concept), mais quelque chose que l'on *attribue* activement.

Dans ce chapitre, nous accorderons une place à la question des rapport du double, dans l'imagination, à l'impossible, à la contradiction. Je propose de montrer que le double peut être considéré de manière rationnelle sans devoir souscrire nécessairement à la notion de mondes impossibles. Il me semble que, loin de revendiquer d'emblée le domaine de l'impossible, du paradoxe, de la contradiction, la littérature du double élabore ses productions en poussant la logique au plus profond de ses retranchements. Je me propose de montrer, dans la dernière section de ce chapitre, que penser le double n'est pas nécessairement penser l'impossible à travers une proposition de lecture de la nouvelle *Lui !* de Maupassant.

Le plan du chapitre est le suivant. Dans la première section, on propose une typologie des figures du double différente de celles avancées par les catégorisations allégorique et symbolique, que nous critiquerons au passage. Les problèmes afférents à l'identification des personnages doubles sont exposés dans la seconde section. La troisième section propose des éléments pour une théorie contextuelle de l'identification des personnages. Enfin, la quatrième section traite de la manière de gérer, à l'aide des mondes possibles, les identités impossibles sur un cas particulièrement pathologique d'identification imaginaire.

5.1 Figures du double en fiction

Dans cette section, nous allons présenter les analyses allégorique et symbolique du double en fiction, avant de proposer notre propre typologie, en récoltant différentes exemples de doubles fictionnels qu'il s'agira ensuite, dans la section suivante, de confronter aux théories des *ficta*.

5.1.1 Les analyses du double en fiction

Selon Paul Coates, le double relève de l'essence-même de la littérature. Il décrit l'existence des œuvres de fiction comme relevant d'un domaine propre entre « le double et l'autre », de la duplication qu'impose au récepteur l'acte de lecture :

> Works of fiction exist in a space between the Double and the Other. To enter into a work of fiction is in a sense to transform the Other into a Double: to discover in the apparent foreignness of another person the lineaments of one's own aspirations and hopes. The manner in which this process of identification and transformation expands the sympathies can be seen as a concomitant of the centralisation and integration of society carried out during the modern era, when more fiction will be written than ever before. (COATES, 1988, p. 1)

Ce point de vue hégémonique du double parmi les œuvres de fiction ne fait pourtant pas consensus au sein de la critique littéraire. Le plus souvent, la critique stipule une particularité de la littérature du double, comme il y a aussi une littérature de la guerre ou une littérature sentimentale. Pourtant, il n'est guère certain que le double relève d'un concept unifié. En témoigne le fait que les critiques ne s'accordent pas sur la question de la nature du double en littérature : s'agit-il d'une figure, d'un thème ou d'un genre ? Les auteurs ne s'entendent pas non plus sur l'extension que recouvre le concept, c'est-à-dire sur l'ensemble des œuvres que recouvre la littérature du double. Il semble que le thème du double ait des démarcations assez vagues et, pour ainsi dire, assez lâches. Nous allons cependant tenter de resserrer « le tissu » fictionnel du double pour en faire saillir les caractéristiques les plus discriminantes.

Le double comme motif ou comme thème fictionnel

De fait, la littérature du double, souvent considérée comme une variété du genre fantastique, pose d'emblée la question de savoir s'il s'agit d'un motif ou d'un thème. Un motif, ou figure, est un concept nominatif qui se fixe exclusivement dans un personnage, comme par exemple les personnages de Sosie et de Goliadkine. A l'inverse, le thème est défini par Tzetan Todorov et Oswald Ducrot (DUCROT & TODOROV, 1972) comme une catégorie sémantique éventuellement présente dans l'intégralité d'un texte, voire dans la totalité de la littérature, comme la réalité, la vie, etc. Serait-ce le cas du double ? Relève-t-il seulement du motif, ou occupe-t-il la place plus importante d'un thème ?

L'aspect le plus manifeste du double est certainement l'évocation de personnages indiscernables. À partir de la publication de son roman *Siebenkäs* (1796-1797), l'écrivain allemand Jean-Paul met en circulation le terme « *doppelgänger* » pour désigner le pendant d'un individu qui lui est similaire en tout point quoique numériquement distinct. À partir du romantisme allemand, les indiscernables deviennent effectivement le motif de prédilection d'un genre, le fantastique. C'est pour cette raison que, dans son approche comparatiste de la littérature du double, Cécile Kovacshazy écrit que « la définition la plus élémentaire du double est l'apparition par deux fois ou plus d'un même personnage. » (KOVACSHAZY, 2012, p. 13). Cette définition a l'élégance de la simplicité, mais s'avère néanmoins malheureuse, parce qu'elle s'applique uniformément à toutes les occurrences d'un personnage, même lorsqu'il s'agit d'un texte où l'on n'a aucunement affaire au

double. Quand le nom d'un personnage est mentionné à la page 5 d'un roman et qu'il réapparaît à la page 7, rien ne laisse présumer que nous avons affaire à la figure du double. Aussi, Kovacshazy précise-t-elle encore que cette figure ne surgit que lorsque « un personnage en rencontre un autre qui lui est en tout point identique, ou bien une partie du personnage, ombre ou reflet, se détache de lui, ou bien encore le récit comprend deux ou plusieurs personnages similaires mais seul le lecteur constate leur identité. » (KOVACSHAZY, 2012, p. 13). Mais quelle identité le lecteur est-il invité à constater ? S'agit-il d'une identité des qualités ou de l'identité numérique des personnages ?

Les approches thématiques tendent généralement à identifier les récits de double à l'allégorique ou au symbolique.

L'allégorie consiste en une forme narrative. Il s'agit d'une histoire généralement courte, basée sur des éléments concrets, mais dont la compréhension ne doit pas être exclusivement littérale. La véritable signification d'une allégorie, comme par exemple celle de la caverne de Platon, relève en effet d'une idée abstraite dont l'expression nécessite le recours à une comparaison avec des éléments concrets. Elle ressortit donc de l'ineffable, de ce qui ne peut être exprimé directement, même si le concept est directement accessible. Elle traite d'une idée qui est certes compréhensible, mais difficile à évoquer par le langage.

Quant au registre du symbole, Gilbert Durand écrit de lui qu'il est, « comme l'allégorie, reconduction du sensible, du figuré au signifié, mais en plus il est par la nature même du signifié inaccessible, *épiphanie*, c'est-à-dire apparition, par et dans le signifiant, de l'indicible. » (DURAND, 1989, pp. 11–12) [1]. Contrairement à l'allégorie, le symbole se rapporte donc à une notion qui n'est pas seulement ineffable, mais dont la connaissance est, de manière directe, tout à fait impossible, comme par exemple le mythe symbolique du Phédon dont l'objet, l'au-delà de la mort, est par nature interdit à toute expérience. Le symbolique ressortit donc au domaine des choses intangibles, sinon absentes, dont le concept ne peut être atteint directement.

Le double comme allégorie

La conception allégorique permet de conférer aux figures du double un aspect plus général en les haussant au niveau du thématique. Ainsi, Cécile Kovacshazy (KOVACSHAZY, 2012) propose une typologie du double selon une classification en quatre grandes figures allégoriques. La première figure est exemplifiée par le couple gémellaire Castor-Pollux qui représente la fidélité indestructible. La deuxième figure est incarnée par Sosie qui représente l'idée d'usurpation par similarité. La troisième relève du couple Tirésias/Hermaphrodite qui représente la notion de dualité. Enfin, la quatrième catégorie ressortit à la figure de Narcisse qui représente ce que Kovacshazy désigne comme l'« identité impossible ».

Les explications allégoriques tentent de réduire la figure du double à un principe qui n'est pas exprimé sous la forme d'un concept général, mais exemplifié par une histoire concrète à travers un concept nominatif comme Narcisse ou Sosie. Ceci étant, les explications allégoriques focalisent

[1] Il convient donc de distinguer la caractérisation « symbolique » du symbole de sa caractérisation sémiotique, inaugurée par Charles Sanders Peirce, qui place le symbole à côté des autres formes de signes que sont les icônes et les indices. Comme signes, les indices entretiennent une relation causale avec ce qu'ils signifient (la fumée pour le feu), les icônes entretiennent une relation de similarité (le dessin de flammes pour le feu) et les symboles sont totalement arbitraires, et culturels (comme la marque « x » indiquant la nocivité d'un produit chimique).

généralement leur attention sur des aspects duaux, comme ceux qui sont présents dans le mythe d'Hermaphrodite ou dans l'allégorie platonicienne de l'Androgyne [2]. Ces analyses ne traitent jamais la question de la ressemblance parfaite en tant que telle. Bien que les fictions qui mettent en scène des duaux, entendus au sens de couples d'opposés contradictoires (masculin/féminin, lumière/nuit, etc.) soient, du point de vue de la critique littéraire, souvent associées au thème du double, je les exclurai pourtant hors de mon analyse, parce qu'elles ne relèvent pas directement d'une forme d'hésitation, chez un agent, entre un constat d'identité et un constat de différence numérique. En outre, l'une des caractéristiques du double tel que je le conçois est la multiplicité expansive, et non la polarité binaire : le double est toujours aussi triple, quadruple, il ne cesse de se reproduire dans des subdivisions et des démultiplications infinies, ce qui n'est jamais le cas des opposés duaux qui doivent toujours être pensés comme un binôme solidaire et irréductible. De ce point de vue, les duaux ne relèvent donc pas du double, mais du complémentaire. En quelque sorte, l'analyse allégorique du double lui confère une place sans doute trop générale, en laissant échapper la question de la ressemblance parfaite et de l'identité dans leur rapport à la représentation.

Le double comme symbole

Passons à l'examen de l'interprétation symbolique. C'est sans doute Otto Rank qui en a donné l'expression la plus aboutie. Comme nous l'avons indiqué, elle consiste à exposer par une figure quelque chose d'inconcevable. Dans son essai sur le double, Rank insiste particulièrement sur le symbole de l'ombre, qu'il juge être une sorte de témoignage de l'âme. De fait, dans la plupart des analyses qui en sont faites, l'immatérialité de l'ombre fait pendant à une putative part immatérielle de l'individu. Mais l'ombre n'est pas seulement « l'indice » d'une individuation, d'une identité personnelle, elle entretient, dans certaines conceptions, des liens plus forts. Dans la section intitulée « L'âme comme ombre et comme réflexion » du *Rameau d'Or*, l'anthropologue James George Frazer rappelle qu'à travers leurs rituels, les sociétés traditionnelles associent l'ombre, sinon à l'âme, du moins à une propriété essentielle de l'individu qui la projette. Certaines précautions s'avèrent donc indispensables pour assurer la sécurité de l'individu à travers celle de son ombre. Ainsi, la tradition folklorique véhicule par exemple la croyance que si l'ombre d'un individu est « foulée aux pieds, si elle est frappée ou percée, l'homme sentira la blessure comme si son corps à lui la recevait ; et si on la détachait de lui, comme il le croit possible, il mourra » (FRAZER, 1981, p. 529). Symbolisant le caractère volatile de l'âme, l'ombre prolonge et double les corps matériels. Elle fait parfois l'enjeu d'enchères, comme peut l'être un rein, ainsi que le rappelle *L'étrange histoire de Peter Schlemihl ou l'homme qui a vendu son ombre* du poète et botaniste franco-allemand Adalbert von Chamisso (CHAMISSO, 1973). Dans ce récit, l'ombre, immatérielle, ne représente presque rien pour la personne, mais sa disparition a pour conséquence la perte de presque tout. Jugée à l'aune d'une double évaluation économique et métaphysique, l'ombre est considérée, comme l'écrit le critique Wladimir Troubetzkoy, à la fois « sans aucune valeur et (...) infiniment précieuse » (TROUBETZKOY, 1996, p. 111).

Or, si dans l'analyse symbolique le double spéculaire représente la part immatérielle de la personne, comme le souligne Troubetzkoy :

[2] Fait exception à ce constat l'ouvrage de Pierre Jourde et de Paolo Tortonese (JOURDE & TORTONESE, 1996, p. 179).

> Qu'on se le dise bien, un nez est un nez, et l'ombre est ombre, s'ils étaient autre chose, cela se saurait, est s'ils étaient pour autre chose, le récit se serait plus qu'une allégorie à clés s'épuisant dans ses interprétations. (TROUBETZKOY, 1996, p. 19)

Et de fait, l'association étroite que Rank tisse entre l'ombre et une signification symbolique renvoyant à l'âme est loin d'être également partagée, et cela même par Chamisso.

Dans la préface que Chamisso donne à l'édition française de 1838, l'auteur répond aux sollicitations de lecteurs le pressant de fournir une explication à son récit. En effet, dès la première publication du roman en 1814, le public de l'œuvre s'est empressé de rechercher la signification symbolique de l'ombre perdue de Schlemihl. Ainsi, Troubetzkoy relève qu'elle fut successivement interprétée comme le symbole « de l'honneur (perdu) de Schlemihl, de la patrie (elle aussi perdue de Chamisso), de l'identité nationale, de la *persona* sociale selon Carl Gustav Jung, de la face sombre ou nocturne de moi, de l'intégralité de la personnalité, de la solidarité avec la communauté humaine, de la participation à la société bourgeoise, des rapports humains dans la société capitaliste naissante » (TROUBETZKOY, 1995, pp. 21-22). Or, dans sa préface, Chamisso répond ainsi à ses lecteurs :

> Cette histoire est tombée dans les mains de gens réfléchis qui, accoutumés à ne lire que pour leur instruction, se sont inquiétés de savoir ce qu'était l'ombre. Plusieurs ont fait à ce sujet des hypothèses fort curieuses ; d'autres, me faisant l'honneur de me supposer plus instruit que je ne l'étais, se sont adressés à moi pour obtenir la solution de leurs doutes. Les questions dont j'ai été assiégé m'ont fait rougir de mon ignorance. Elles m'ont déterminé à comprendre dans le cercle de mes études un objet qui jusque-là leur était resté étranger, et je me suis livré à de savantes recherches dont je consignerai ici le résultat. (CHAMISSO, 1874, p. III)

La suite de l'explication est une analyse physico-géométrique des ombres tirée du *Traité élémentaire de physique* de Haüy, définissant l'ombre comme zone protégée de la projection d'une lumière sur un solide. Le message de Chamisso est donc sans ambiguïté : l'ombre ne doit pas être interprétée pour autre chose qu'elle-même. Un récit ou un personnage perd son ombre doit être compris tel quel. L'ombre n'y a d'autre signification que physique. La compréhension de son récit ne doit pas recourir à une interprétation symbolique. Il s'agit de concevoir véritablement qu'un individu perd son ombre, rien de plus, rien de moins, ce qui constitue une sorte de rejet de l'interprétation symbolique qui déporte son attention hors de la question des liens entre ressemblance et identité.

L'analyse structurale du double

Je m'inscris en faux contre ce qu'écrit Troubetzkoy quand, dans son approche symbolique, il écrit que « [...] on peut *imaginer*, on peut rêver, fantasmer, (croire) voir le double, on ne peut le *concevoir* » (TROUBETZKOY, 1996, p. 2). Pour qu'une fiction du double « fonctionne », le double doit y être concevable.

La position que je souhaite défendre consiste à éviter de réduire les fictions du double à des interprétations allégoriques et symboliques. De mon point de vue, le double constitue une expérience fictionnelle qui n'est réductible à aucun de ces types de lecture. Elle se rapporte, à mon

sens, à la possibilité logique que deux choses soient parfaitement similaires, ou de carrières alternatives vécues par un même individu. En rejetant l'explication herméneutique du double, on peut recentrer l'attention sur le sentiment qu'il favorise, sentiment trouble qui résulte d'une hésitation entre un constat d'identité ou de distinction numérique. Les fictions du double sont ainsi définies comme celles où il arrive que le lecteur, voire le personnage, ne puisse savoir exactement s'il y a une ou deux choses en présence.

De fait, vouloir enrégimenter le double dans le registre du symbolique ou de l'allégorique conduit inévitablement à en perdre la richesse. Dans leur remarquable ouvrage synthétique sur la littérature du double, Pierre Jourde et Paolo Tortonese remarquent que « les formes de doubles sont si multiples qu'il apparaît vite illusoire de chercher à assigner un sens déterminé à chacune d'entre elles » (JOURDE & TORTONESE, 1996, p. 179). Cherchant à dégager le thème littéraire du double, Jourde et Tortonese constatent que pour qu'il y ait double, il faut « que l'on sente avant tout une perturbation de la loi de différence » (JOURDE & TORTONESE, 1996, p. 5). Autrement dit, il faut soit qu'une discernabilité *stipulée* en vienne à être perturbée par une identité *présumée*, soit inversement qu'une identité stipulée soit perturbée par une discernabilité présumée. Les deux auteurs remarquent aussi que les objets doubles inanimés sont bien moins scandaleux que les consciences doubles. C'est que la tradition occidentale ancrée sur la théologie chrétienne refuse la perspective de deux âmes indiscernables. Selon les auteurs, le thème littéraire du double est justement caractérisé par cet affranchissement, il doit donc être restreint aux seuls cas des récits où une excessive ressemblance corrèle deux consciences :

> Seul, le maintien, dans le récit, d'une tension créée par l'excès de ressemblance, en général, corporellement enracinée, permet de parler de doubles. (JOURDE & TORTONESE, 1996, p. 16)

La redondance physique chez un être intentionnel est alors le facteur discriminant des récits de doubles. Le double est caractérisé par l'effet psychologique qu'occasionne le rejet de l'indiscernabilité des identiques.

En dehors des approches symboliques et allégoriques, une analyse qui me semble plus intéressante est celle d'ordre structurel. De ce point de vue, plusieurs sortes de questions peuvent être posées. D'abord celles concernant la nature des doubles : dans telle fiction, à quel type de doubles a-t-on affaire ? De fait, les fictions du double peuvent se rapporter à deux aspects inverses l'un de l'autre : celles qui annoncent des ressemblances frappantes qui défient une distinction numérique présumée (cas des indiscernables), mais aussi celles qui, inversement, stipulent des différences qualitatives qui défient l'identité attendue (doubles-mêmes, contreparties). En considérant l'aspect narratif du double, on peut encore les différencier, comme le font Jourde et Tortonese, selon leur incidence purement psychologique (double interne) ou physique (double externe), selon que le double n'est qu'une hallucination subjective ou qu'il existe objectivement dans le monde du personnage. On peut prendre en compte l'aspect narratif en examinant comment évoluent les dénombrements au fil de l'intrigue. Y a-t-il une constante différence numérique ? L'identité est-elle au contraire « retrouvée », comme dans les fictions qui annoncent au départ une distribution numérique qui s'effondre dans une unité au court du récit ? Inversement, l'indiscriminabilité peut aussi être défaite quand une identité annoncée s'avère finalement être fausse. L'indiscriminabilité indifférente n'annonce aucune identité et n'en rejette aucune. Il y a également la question de savoir pour qui l'identification est problématique. Serait-ce pour le récepteur de la fiction ? Pour tous

les personnages, ou pour quelques-uns, voire pour un seul ? Les doubles peuvent être subjectifs (seuls les protagonistes se sentent dupliqués), ou objectifs (il y a des témoins du double). Enfin, nous pouvons poser la question du mode de production des dupliques. Par quel moyen est rendue concevable la duplication ? Par un processus surnaturel (par exemple un dieu prenant la forme d'un avatar), ou contraire naturel (les sosies et les jumeaux), voire artificiel (clone, androïde) ?

Les tentatives d'individuation des mondes par les personnages ont suscité chez Lubomír Doležel un intérêt particulier pour le cas des récits de doubles. Dans un article de 1985, intitulé « le triangle du double », l'auteur soutient qu'en fiction, la thématique du double est entièrement circonscrite par trois catégories modales qui en structurent les variantes possibles. La première catégorie relève des cas où un même personnage réside dans différents mondes séparés dans l'espace ou dans le temps, comme dans l'Orlando de Virginia Woolf. La deuxième catégorie relève du cas des récits qui parlent d'individus indiscernables habitant un monde commun. Enfin, la troisième catégorie concerne les cas où deux incarnations alternatives d'un même individu habitent le même monde, ce qui correspond proprement, affirme Doležel, aux récits de doubles. Doležel illustre ces trois possibilités selon le schéma de la figure 5.1. Les lettres A, B, X et X' représentent des personnages, les lettres w, w', etc. représentent des mondes fictionnels, et les doubles flèches représentent des relations d'identité.

Le premier type, identifié par la figure d'Orlando, relève en quelque sorte des *mêmes*, autrement dit, de la question d'entités qui seraient entièrement présentes dans différents mondes à la fois, de la même manière qu'on considère, dans une perspective endurantiste, qu'un individu est entièrement présent à différents moments du temps. Le second cas, relevant de la figure d'Amphitryon, relève des indiscernables numériquement distincts. Enfin, contrairement aux deux précédents, le troisième cas soulève un véritable problème d'identité contradictoire, à savoir, le cas d'une chose qui serait entièrement présente à deux emplacements différents, non dans deux mondes distincts, mais cette fois dans un même monde. Je m'attarderai plus particulièrement sur ce problème dans la troisième section de cette partie avec l'examen d'un récit fantastique qui relève de cette figure, à savoir, la nouvelle *Lui !* de Maupassant.

Le diagramme de Doležel permet essentiellement d'établir les variantes que l'on peut générer à partir des contraintes structurales qui lient les trois thématiques principales. Doležel définit ainsi les modes de construction du double (par fusion, par fission, par métamorphose), les relations paradigmatiques (par similarité vs. contraste, par antagonisme vs. coopération), les relations syntagmatiques (simultanéité et exclusivité) et enfin l'authenticité (garantie, récusée, ambiguë). Leurs corrélations sont ensuite déterminées par une dispersion triangulaire comme représenté sur la figure 5.2 :

Je ne reprendrai pas telle qu'elle cette catégorisation du double, et j'en proposerai une autre, alternative. L'une des raisons à cela est que la typologie de Doležel est limitée aux catégories intrafictionnelles, et d'autres raisons apparaîtront dans ce qui suit.

5.1.2 Les doubles natifs

Je distinguerai pour ma part deux premières catégories de doubles, les doubles natifs, qui sont des doubles intrafictionnels, et les doubles importés, qui sont des doubles interfictionnels, comme le sont les personnages récurrents.

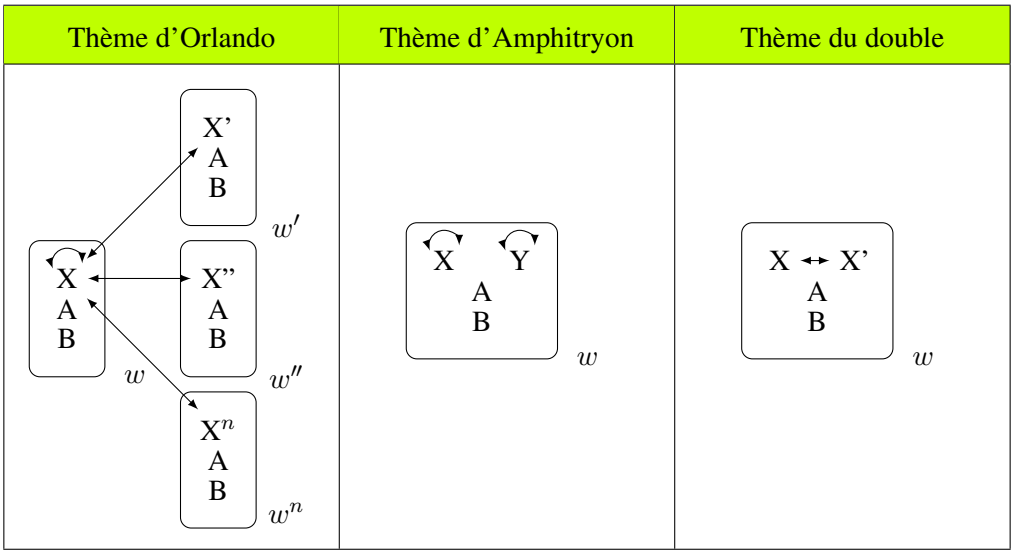


FIGURE 5.1 : Distinction des thématiques du double selon Lubomír Doležel (DOLEŽEL, 1985)

Si la fiction enrichit nos connaissances à l'égard du monde, ce n'est pas lui faire offense de dire qu'elle s'alimente des thèses métaphysiques, des connaissances et des hypothèses scientifiques, des croyances religieuses. D'un point de vue proprement structurel, ces trois sources permettent de diversifier les types de doubles et la manière dont ils sont engendrés. Dans ce qui suit, je cheminerai sur le terrain des fictions du double pour y analyser nos manières de comprendre l'indiscernabilité qu'elles évoquent, en particulier lorsqu'elles relatent des formes pathologiques de l'identité métaphysique. A cette fin, cette première section survole quelques exemples littéraires et cinématographiques de figures du double. L'idée est d'évaluer les invitations, lancées par la littérature, à concevoir des indiscernables et d'analyser la manière dont nos représentations s'emploient à en susciter l'idée.

À vouloir définir le thème du double dans la fiction, on se heurte inexorablement à des problèmes de classification des œuvres, à des mises à l'écart plus ou moins indues et à des agglomérations plus ou moins discutables. Ma contribution à la caractérisation du double en fiction propose de distinguer cependant trois formes principales de doubles natifs, les indiscriminables ostensifs, les indiscriminables occultifs et, enfin, les doubles-mêmes. Dans la catégories des indiscriminables, on retrouvera, sous une forme épistémique, les trois types d'indiscernables que nous avons déployés dans la première partie de ce mémoire, à savoir, les répliques pures, les doubles purs et les coïncidents purs, modulo le fait que de tels rapports à l'indiscernabilité sont relativisés à un agent et à un contexte donnés sous la notion épistémique de discriminabilité. Nous dirons ainsi que deux personnages sont indiscriminables si le lecteur, le narrateur ou l'un des protagonistes s'avère être dans l'incapacité de les différencier.

Précisons la terminologie qui sera utilisée. Nous parlerons d'indiscriminabilité *ostensive* quand l'auteur d'un texte attribue des caractéristiques *explicitement et exclusivement* identiques à des personnages qui, néanmoins, sont considérés comme numériquement différents. Il s'agit précisément

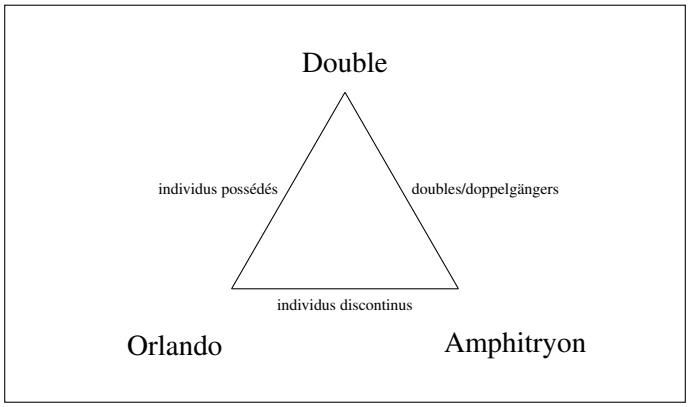


FIGURE 5.2 : Le triangle du double selon Lubomír Doležel (DOLEŽEL, 1985)

d'une indiscriminabilité comparative basée sur des informations de similarité qui peuvent sembler excédentaires - un trop plein de ressemblances entre des choses qui ne peuvent être finalement que difficilement considérées comme différentes. Au contraire, l'indiscriminabilité sera dite *occultive* lorsque l'auteur ne fournit *aucune* information permettant de mettre en comparaison les personnages (un manque complet d'informations). Dans les cas limites qui échoient à ces types, deux personnages peuvent relever de doubles purs ou de coïncidents purs, quand aucune propriété, ni intrinsèque ni même extrinsèque, ne permet de les distinguer l'un de l'autre.

Dans ce qui suit, je parcours quelques exemples de fictions du double, qui serviront à éclairer ces catégories. En ce sens, il s'agira moins d'illustrations que d'exemplifications. Précisons que je ne souhaite pas relever exhaustivement toutes les figures du double, ce qui serait une tâche quasi-impossible en raison de la variété du thème. Mon but est d'élargir la compréhension des fictions du double par delà les limites où elles sont trop souvent cantonnées, à savoir, selon des interprétations allégoriques, symboliques et structurelles. Je montrerai, à travers ces exemples, qu'on peut également offrir une compréhension conceptuelle des fictions du double, autrement dit, qu'elles ne se rapportent pas de manière nécessaire ni à de l'impensable, ni à de l'ineffable.

Les indiscriminables ostensifs

Le type de l'indiscriminable ostensif a fréquemment servi aux pièces comiques. S'il ne manifeste pas un contenu comique par lui-même, il peut engendrer cependant un comique de *répétition* et, en particulier, encourage ce que Bergson a appelé dans son opuscule *Le rire*, « l'interférence des séries » qui sert de ressort comique aux comédies classiques. Le principe d'interférence peut être défini de la manière suivante : il y a interférence entre deux séries d'événements indépendantes lorsqu'elles sont interprétées comme appartenant en totalité ou partiellement à une seule et même ligne d'événements. Or le motif du double favorise diverses occasions de quiproquos [3], d'identifications usurpées ou erronées. En particulier, il permet de jouer sur l'ambiguïté entre identité logique, identité personnelle, identité qualitative et capacité du langage à référer, produisant ainsi

[3]Rappelons au passage l'étymologie issue de *quid pro quo*, signifiant « une chose pour une autre ».

des confusions par la rencontre inattendue de lignes distinctes d'événements.

En témoigne le succès intemporel de la farce *Amphitryon*, créée en -197 par Plaute, et reprise, entre autres, par Jean Rotrou (1636), Molière (1668), John Dryden (1690), Heinrich Von Kleist (1807) et Jean Giraudoux (1929) [4]. Elle est tirée d'un mythe grec décrivant la rencontre de Zeus avec Alcmène, d'où résultera la naissance de deux jumeaux dont l'un est le héros Héraclès. L'argument de la pièce est le suivant : Zeus veut séduire Alcmène. Pour ce faire, il doit déjouer la fidélité inébranlable qu'elle voue à son mari Amphitryon. Pour satisfaire son désir et rejoindre la couche d'Alcmène, Zeus prend l'apparence d'Amphitryon, alors que ce dernier est parti guerroyer hors de Thèbes où réside sa maison. Plaute complexifie le thème en ajoutant, parallèlement au couple Zeus/Amphitryon, une seconde duplication, celle d'Hermès usurpant l'apparence de Sosie, le serviteur d'Amphitryon. Transformé, le messager des dieux barre la route à Sosie qui doit accéder à la maison d'Alcmène pour lui donner des nouvelles de son époux. C'est que, argumente Hermès en substance, s'il est lui-même Sosie, l'autre Sosie (le vrai) ne peut prétendre à l'être également. Sosie est donc privé du droit de franchir le seuil de la maison d'Alcmène, livrant celle-ci à la disposition de Zeus.

L'un des ressorts comiques de la pièce provient de la rencontre entre Sosie et Hermès transformé en Sosie, situation dont Hermès profite en s'attribuant une identité tour à tour légitime et usurpée, jurant sur sa propre figure « par Hermès, que Sosie soit rossé s'il (Hermès) n'est pas Sosie lui-même ». Sosie doit bien reconnaître que celui qui lui barre le passage manifeste sa propre apparence. Il revêt son vêtement, il a même allure, mais plus encore, il en a tous les traits. Sosie semble accepter de bon droit l'application du principe d'identité des indiscernables, admettant qu'il est tout à fait impossible que lui et celui qui lui barre le passage soient parfaitement ressemblants. Mais son double a pourtant toute son apparence. Si on lui emprunte son identité sociale, à savoir l'attache à son maître et son patronyme, que reste-t-il alors à Sosie ? « Qui suis-je alors, si je ne suis Sosie ? », demande-t-il. La raison du trouble où est placé Sosie n'est pas l'effraction du principe logique que A est A, que Sosie ne remet jamais en question. Le préjudice est donc imputable soit à une persistance temporelle défaillante de Sosie (« [...] Dieux immortels, je vous implore, où suis-je mort ? »), soit à la perte d'une identité qualitative (« Où me suis-je transformé ? Où ai-je quitté mon apparence ? [...] Car cet individu possède toute l'image de moi-même, celle que j'avais avant. »), voire à une perte de mémoire (« Me serais-je laissé là-bas, en m'y oubliant ? »). Finalement, conclut Sosie pour ce qui le concerne, il a dû perdre sa propre apparence lors de son voyage de retour vers Thèbes, et un autre la lui aura empruntée. L'inquiétude dont est sujet Sosie n'est donc pas la conséquence d'une mise en déroute logique, d'un principe d'identité qui s'avérerait défaillant, ou de l'appréhension d'une incongruité métaphysique. Cette appréhension provient plutôt d'une perte de reconnaissance sociale et d'une menace d'usurpation.

Ce qui décourage Sosie de mener à terme sa mission est finalement moins une hésitation d'identité suscitée par une ressemblance exagérée que les multiples coups qu'il reçoit d'Hermès. Notons d'ailleurs qu'Hermès reste discriminable en ce qu'il frappe « quatre fois plus fort ». Ce n'est donc pas Sosie qui se bat lui-même. Sosie n'est pas un *héautontimorouménos*, un bourreau de lui-même. Hermès n'est finalement pas assez ressemblant pour parvenir à faire douter Sosie de sa propre identité. Et d'ailleurs, la ressemblance est-elle suffisante pour tromper autrui ? Ce n'est pas certain, comme le souligne Giraudoux dans sa version de la pièce. Dans *Amphitryon 38*, Hermès

[4] Giraudoux affirmait avoir dénombré 38 versions connues de la pièce, la sienne comprise.

objecte à Zeus que même sous les parures d'un humain, il ne cesse jamais vraiment d'apparaître encore comme un Dieu. Le vêtement, constate Hermès, est trop parfait, trop résistant. Les yeux ? Trop brillants. Le corps ? Trop souple et robuste, il ne connaît pas la courbure des épaules portant le poids des années. L'être parfait qu'est Zeus sera toujours incapable de ressembler suffisamment à un homme, car il restera toujours diverses imperfections à combler entre lui et son modèle [5].

Ainsi, Sosie n'est jamais vraiment dupe de la tromperie dont il fait l'enjeu, et sans doute devine-t-il la manigance d'un dieu. Il ne croit pas à la possibilité d'une duplication merveilleuse de lui-même : son sentiment reste qu'il ne peut y avoir qu'un seul Sosie. La pièce de Plaute n'indique jamais l'idée d'un conflit du soi, d'une division entre le sujet pensant et son objet de pensée. Autrement dit, l'histoire de Sosie n'est pas encore celle d'un double réflexif comme elle le deviendra chez Dostoïevski. Dans la pièce de Plaute, l'intérêt du double ne réside pas dans l'évocation d'une situation fantastique et des troubles d'identité qui seraient occasionnés chez le personnage. Le double a rapport au sentiment comique qui provient des multiples quiproquos dus à l'incapacité qu'a Sosie à rapporter les faits. Les défaillances ne sont donc ni métaphysiques, ni logiques, ce sont des défaillances communicationnelles. En voulant décrire la mésaventure qu'il a endurée, Sosie a toutes les peines du monde à en rendre la réalité. Dans la version de Rotrou, les intrications de références s'enchaînent en défilant l'écheveau des identifications : « Qui t'a battu ? (demande Amphitryon) - Moi-même (répond Sosie) - Et pourquoi ? - Sans raison. - Toi ? - Moi, vous dis-je, moi qui suis à la maison » (acte II, scène 1). Et aussi chez Molière : « On t'a battu ? - Vraiment ! - Et qui ? - Moi. - Toi, te battre ? - Oui, moi, non pas le moi d'ici, mais le moi du logis, qui frappe comme quatre » (acte II, scène 1). Sosie fait un emploi tout aussi exclusif qu'excessif du pronom personnel « moi », utilisé de manière indexicale pour désigner à la fois le vrai et le faux Sosie. Alors que « moi » fait toujours référence au sujet de l'énonciation, dans l'emploi qu'en fait Sosie, ce sujet est encore projeté dans ce qui n'est plus lui, mais un autre. Cela conduit à confondre identité logique (moi, c'est moi) et identité qualitative (cet autre qui m'est indiscriminable). Sosie ne parvient ainsi à faire référence qu'à un seul individu là où il faudrait pourtant en indiquer deux. De même, dans la version de Molière, quand on demande les raisons pour lesquelles Sosie n'a pas délivré le message convenu, et qu'on sollicite qu'il précise ce qui l'en a empêché, Sosie répond « moi, que j'ai rencontré, moi qui suis sur la porte, moi qui me suis moi-même ajusté de la sorte, moi qui me suis chargé d'une grêle de coups, ce moi qui m'a parlé, ce moi qui suis chez vous. ». Sosie ne parvient pas à expliquer la situation à son maître parce qu'il ne fait usage que du seul indexical « moi » pour faire référence alternativement à lui-même et à son double illégitime. Le moi que ne cesse de répéter Sosie dans la pièce de Plaute est, somme toute, encore impersonnel, il est un « il », et non pas encore un « je ». Distinction qui reste hors de portée de Sosie, et qui mène Amphitryon à ne concevoir qu'un seul référent là où Sosie ne cesse de vouloir dire qu'il y en a deux, sans y parvenir. Le langage employé par Sosie n'est pas assez affiné pour parvenir à

[5] C'est d'ailleurs cette imperfection qui permet, dans une variante indienne rapportée par Voltaire, de confondre le dieu venu usurper la place du mari. Voltaire y voit même une forme de sagacité supérieure à celle du mythe grec. Dans cette variante, un dieu profite de l'absence d'un fugitif pour usurper son apparence et tromper sa femme. Mais le mari revient, et revendique son identité. L'affaire paraît devant le tribunal de Bénarès. Le président de la cour, qui a deviné la présence d'un dieu, propose d'identifier le mari ainsi : les deux prétendants s'accoupleront successivement avec la femme convoitée. Celui dont l'éjaculation sera la plus généreuse sera déclaré légitime, le mari étant réputé pour sa vigueur sexuelle. Le premier prétendant émet 12 jets, le second, 50 jets. Alors que les participants du tribunal concluent l'authenticité du second prétendant, le président du tribunal intervient en contestant que seul un dieu peut parvenir à une telle prouesse physique. Le dieu finit d'ailleurs par avouer sa tromperie, s'en retournant au ciel amusé de son aventure.

mentionner les choses dans leur diversité (moi qui suis là, cet autre qui me ressemble) et les réduit inévitablement à une identité grossière auquel ne souscrit manifestement pas Sosie lui-même. Ce sont donc les jeux de langage qui font émerger le sentiment comique. Cette situation contribue à l'amusement du spectateur, quant à lui avantageusement placé au point de vue omniscient, dans la confidence des faits.

Chez Plaute, la confusion entre le « moi et l'autre moi » est le symptôme d'un rapport à soi qui n'est jamais internalisé, mais toujours pensé de l'extérieur. Car comme l'affirme Jean-Pierre Vernant, s'il est indéniable que les Grecs anciens aient eu un rapport à eux-mêmes, ce rapport est sans doute organisé d'une façon très différente du nôtre. Le moi n'est ni délimité, ni unifié, mais ouvert à des forces multiples (VERNANT, 1989, p. 224). Il n'est pas encore pensé comme une relation d'intimité, toute intérieure, mais comme projection extérieure produite par les actions :

> L'individu [...] s'objective dans ce qu'il accomplit effectivement, dans ce qu'il réalise : activités ou œuvres qui lui permettent de se saisir, non en puissance, mais en acte, *enégeia*, et qui ne sont jamais dans sa conscience. Il n'y a pas d'introspection. Le sujet ne constitue pas un monde intérieur clos, dans lequel il doit pénétrer pour se retrouver ou plutôt se découvrir. Le sujet est *extraverti* [6]. (VERNANT, 1989, p. 225)

Chez le sujet grec, l'expérience de soi n'est pas encore existentielle, réflexive. La littérature semble ainsi indiquer que la conscience de soi chez les anciens Grecs est toujours secondaire, qu'elle est toujours donnée *après* l'existence concrète. L'expérience que fait Sosie d'une double apparition implique donc un désarroi qui n'est pas encore réflexif, mais seulement pratique, concurrentiel. Il n'est pas encore le double résultant d'une division dans les profondeurs de la psyché. Car le sujet n'est pas encore pensé sous une forme substantivée (*le Moi*), il reste toujours, dans sa description, un phénomène tout extérieur.

Entre la version de Plaute et celle de Molière, la notion d'identité personnelle a été profondément influencée par la philosophie cartésienne du sujet. Chez Molière, la confiance qu'a Sosie de son propre Moi semble fragilisée. C'est qu'il ne peut y avoir qu'une substance du Moi, et non deux. Ainsi, ce n'est plus seulement le jeu des apparences étranges, celui de deux apparitions similaires qui est responsable de l'effroi, mais l'indiscernabilité entre deux substances qui se donnent pourtant pour distinctes. Chez Molière, Sosie avoue à Mercure :

> Je ne saurais nier, aux preuves qu'on m'expose, que tu ne sois Sosie, et j'y donne ma voix. Mais si tu l'es, dis-moi qui tu veux que je sois. Car encor faut-il bien que je sois quelque chose. (MOLIÈRE, 1947, p. 231)

Si Sosie admet bien l'éventualité qu'il se méprend sur sa propre identité, encore lui faut-il être quelque chose. L'inflexion est d'ailleurs encore plus marquée lors de la rencontre entre Amphitryon et son double. Là, le comique fait carrément place à la terreur : « Mon âme demeure transie » écrit Molière. Le geste d'Amphytrion qui se jette sur son épée pour terrasser son double anticipe déjà la lutte finale que William Wilson livre à son double spéculaire, témoignant d'une menace de scission psychologique du Moi.

[6] C'est moi qui souligne.

Dans *Amphitryon* comme dans d'autres fictions, le moteur produisant la situation de duplication est d'ordre surnaturel, puisqu'il consiste en la visite d'un dieu revêtant toutes les apparences d'un individu donné. On désigne la présence d'une entité transcendante par le terme d'avatar. Le mot vient du sanskrit, où il signifie la descente sur terre du dieu Vishnu. L'entité incarnée devient le véhicule permettant au dieu d'agir dans le monde des hommes. Et parce que les incarnations de la divinité sont multiples, le mot exprime également la notion de transformation, de métamorphose, d'identités plurielles. La transformation indique toujours, par sa capacité qui échappe aux ressorts de l'être imité, une discernabilité, une appartenance spécifique différente, une nature autre.

Les raisons surnaturelles ne sont néanmoins pas les seules explications qu'un auteur peut fournir pour indiquer une ressemblance parfaite. Parmi les causes naturelles de ressemblances, la plus communément exploitée est sans doute la gémellité. Les jumeaux permettent en effet de rendre compte d'une explication simple de ressemblances, sans faire appel à des capacités surnaturelles de métamorphoses. Les ressemblances multiples entre deux personnages s'expliquent ainsi à peu de frais, en cela que la gémellité permet de maintenir un monde fictionnel dont la physique ne s'écarte pas trop de la réalité. Notons au passage qu'en faisant intervenir des jumeaux dans une histoire, on se débarrasse aussi des questions d'originalité et d'authenticité. En rapport à leur similarité réciproque, l'un des jumeaux n'est jamais plus original ni authentique que l'autre. Si Hermès copie Sosie, le membre d'un couple de jumeaux ne copie pas l'autre. Ce constat pourrait toutefois être nuancé, parce que nombre de traditions associent souvent les jumeaux à des figures divines et leur livrent un véritable culte [7]. Comme le remarque Françoise Frontisi-Ducroux dans son analyse exhaustive du double dans la mythologie grecque (FRONTISI-DUCROUX, 2007), rares sont les jumeaux issus de mortels. En effet, les mythes grecs attribuent généralement à l'un des jumeaux, mais pas à l'autre, une origine divine, comme dans le cas de Héraclès. Une différence ontologique est ainsi attribuée à chacun des individus formant un couple gémellaire [8]. L'un relève du monde humain, l'autre, d'un monde supra-humain. Bien sûr, de telles conceptions ressortissent à des croyances erronées et certainement désuètes.

Parmi les écrivains qui ont utilisé les jumeaux comme motif réaliste du double, c'est encore Plaute qu'on doit citer ici, dont la pièce *Les Ménèchmes* (-206) décrit les aventures de deux jumeaux indiscernables qui ont été séparés à leur naissance, et dont les retrouvailles occasionnent divers quiproquos. La séparation à la naissance permet de produire une situation dans laquelle les personnages de la pièce, qui ignorent la gémellité des frères, ne voient qu'un seul individu là où il faudrait qu'ils en constatent deux. À l'instar d'*Amphitryon*, la pièce a connu des reprises. Celles de Shakespeare, *La comédie des erreurs* (1592) et *La nuit des rois* (1602) sont sans doute les plus connues. Dans ces différentes versions de la pièce, l'effet comique surgit des mêmes maladresses de communication que celles que nous avons relevées dans *Amphitryon*, maladresses dues aux confusions faites entre des personnages distincts. Dans de telles situations, les personnages ne sont pas indiscriminables pour le spectateur de la pièce, mais seulement pour l'un ou plusieurs des protagonistes [9].

[7] Ce fait est notamment souligné par Otto Rank dans son essai (RANK, 1973).

[8] L'authenticité qui ressurgit par le biais de cette différence est également associée à une résurgence de la notion d'originalité, en ce que la figure des jumeaux est souvent attachée aux mythes des origines, pensons par exemple aux couples gémellaires Romus/Romulus (Rome), Hunahpu/Ixbalanque (Mayas), Tamendonare et Aricoute (Tupinemba), etc.

[9] Prenant l'indiscernabilité à contre-pied, la comédie intitulée *Jumeaux*, réalisée en 1988 par Ivan Reitman s'inspire

Lorsqu'elle ne trouve pas une explication par la fraternité gémellaire, la simple ressemblance, sans autre explication biologique, relève de ce qu'on appelle, par antonomase, le « sosie ». Avec l'explication naturelle de la ressemblance, nous sortons du récit merveilleux pour accoster aux rivages de l'étrange. La ressemblance devient alors parfois embarrassante, voire inquiétante quand la similarité est parfaite, ou qu'elle fait l'enjeu d'une auto-identification impossible. Comme nous l'avons vu, avec les indiscriminables surnaturels, un potentiel rapport de discrimination divin/humain peut toujours être avancé pour différencier la copie de l'individu authentique. Les répliques appartiennent à une catégorie sortale différente de ce qu'elle dupliquent. Dans le cas des indiscriminables naturels, cette différence n'est plus permise. Cependant, afin de restituer un rapport axiologique à l'authenticité, le couple sortal divin/humain est souvent remplacé par le couple de caractéristiques sociales comme seigneur/serf. Dans *Le Roi par Ressemblance* (1600), comédie attribuée à Lope de Vega, une reine fait assassiner son époux et le remplace par un paysan qui lui ressemble, ce dernier devenant finalement bien meilleur souverain que celui qu'il usurpe. La question est alors celle de savoir ce qui discrimine le vrai du faux souverain. Dans une comédie attribuée au même auteur, *Confusion au palais* (1630), un roi est remplacé durant son sommeil par un sosie. Le roi, qui se croit devenu fou parce que tous ses ordres sont démentis, s'efface progressivement. On découvre finalement que son remplaçant n'est autre que son frère jumeau. Citons encore *L'Aventure par le nom* (1630), de Tirso de Molina, où un berger remplace un roi dont il découvre ensuite qu'il est le demi-frère. C'est aussi la même intrigue dans le film *Kagemusha* (*L'ombre du gerrier*, 1980), où le réalisateur Akira Kurosawa décrit les tribulations, dans le Japon féodal, d'un voleur placé à la tête d'un clan. À la mort du seigneur Shingen Takeda, ses vassaux élaborent une ruse afin d'éviter de néfastes rivalités territoriales au sein du clan qui les réunit. Ils confient discrètement la place du shogun à un voleur qui ressemble suffisamment au défunt pour faire croire que ce dernier est toujours en vie. Le thème du double vise ici la mise en scène d'ambivalences entre le pouvoir politique et sa représentation. Se sentant d'abord illégitime et peu sûr des décisions qu'il doit prendre, le double prend progressivement confiance en lui, et finit par s'identifier complètement au seigneur du clan.

Constatons que toutes ces histoires présentent une même interrogation portant sur le rapport de l'identification aux conditions sociales. Le double sert ainsi une réflexion politique sur la légitimité du pouvoir. Rappelons ici le livre de l'historien Ernst Kantorowicz, *Les deux corps du Roi*. Le point de départ de l'ouvrage est un fait historique assez déroutant, à savoir, l'exécution, en 1649, du roi Charles I par les Puritains *au nom même de Charles I*. Kantorowicz explique le sens de

elle aussi de l'intrigue des *Ménèchmes*. Comme dans la pièce de Plaute, deux frères jumeaux ont été séparés à la naissance et privés de leur mère biologique. L'un des deux frères a été élevé dans un contexte idéal, il est devenu un intellectuel brillant. Le second a été abandonné dans un orphelinat et est devenu un petit escroc. Le film décrit leurs retrouvailles. Contrairement aux *Ménèchmes*, le récit ne tire pas son effet comique des méprises résultant d'une ressemblance parfaite. En effet, la discernabilité qui distingue les prétendus jumeaux respectivement joués par Arnold Schwarzenegger et Danny DeVito est bien plus qu'évidente, contrairement à ce que prétend le slogan français qui annonce que « seule leur mère peut les différencier ». En effet, les deux personnages diffèrent aussi bien du point de vue physique que moral. Le spectateur chercherait vainement une quelconque ressemblance et se demandera en quoi consiste au juste la gémellité des personnages. Cette fiction, assez maladroite et entachée de nombreux clichés, a cependant le mérite d'interroger la place de l'éducation dans le développement de la physionomie et du caractère. Deux individus génétiquement identiques à la naissance seront-ils encore indiscernables trente années plus tard s'ils sont élevés dans des conditions différentes ? Quelle est la part de l'éducation dans la ressemblance ? Les jumeaux permettent ainsi une interrogation sur le rapport de l'environnement au développement de l'individu. Cette question a d'ailleurs été exploitée par le psychologue René Zazzo, nous y reviendrons dans la dernière partie de ce mémoire.

cet événement en indiquant un double corps du roi : à côté de son corps matériel et donc mortel, le roi possède encore un corps moral, immortel, que célèbre d'ailleurs la fameuse locution « le roi est mort, vive le roi ». C'est au nom du corps moral que le corps physique de Charles I fut exécuté. Les divers corps matériels ne sont ainsi que des corps de substitution d'un unique corps politique persistant dans le temps, ce qui amène la problématique du double à la question du corps de substitution [10].

Les modes de duplication ne sont pas seulement surnaturels ou naturels, ils peuvent aussi s'avérer artificiels. Dans ces cas aussi, le mode de production du double réinstaure l'échelle axiologique de l'authenticité et de l'originalité. Font partie de ces doubles les clones, c'est-à-dire les duplications d'entités biologiques, qui font l'objet de nombreux arguments de films. Citons par exemple *La cité des enfants perdus* de Jean-Pierre Jeunet. L'intérêt de ce type en fiction est essentiellement une critique socio-politique des usages de la technologie et une réflexion sur les enjeux éthiques ressortissant à la bio-technologie. Dans *The Island* de Michael Bay, des clones humains sont stockés en attendant de servir de banque d'organes aux individus qu'ils répliquent. Soulignons qu'il est courant que, dans ces fictions, le point de vue soit inversé par rapport au double romantique : l'histoire focalise l'attention du récepteur non sur l'individu qui subit la duplication, mais sur la duplique qui se découvre être la copie d'un autre. C'est du point de vue de la réplique que surgit l'inquiétante étrangeté d'une ressemblance parfaite au modèle, comme le découvrent aussi, à leurs dépends, les répliquants du film *Blade Runner* de Ridley Scott tiré du roman *Do Androïds Dream of Electric Sheep ?* de Philip Kindred Dick, androïdes qui possèdent des souvenirs d'événements qu'ils n'ont pas réellement vécu, mais qu'ont vécu les modèles dont ils sont tirés.

Les androïdes servent d'ailleurs fréquemment à la justification du double. Dans le récit d'E.T.A Hoffmann intitulé *L'homme au sable*, le personnage principal, l'étudiant Nathanaël, s'amourache d'Olympia, la fille de son professeur qui s'avère en réalité être un automate, une poupée mécanique. L'indiscernabilité entre automates et humains n'est pas nécessairement une relation de ressemblance parfaite entre des individus particuliers, mais plutôt une indiscriminabilité dans l'espèce. Dans le *Frankenstein* de Marie Shelley, la créature partage avec son créateur une sorte d'indiscernabilité spécifique, la créature étant une imitation plus ou moins réussie d'homme, elle-même incarnée dans divers lambeaux de chairs humaines recousus entre eux. Qu'est-ce qui différencie la créature de Frankenstein d'un homme ? N'est-elle pas faite, d'ailleurs, d'un tissu d'humains ? La même question revient dans le film *Blade Runner*, où les répliquants, machines imitant le corps, le raisonnement et le comportement des humains, s'avèrent physiquement et psychologiquement, quasiment indiscriminables de leur modèle.

Concevoir que deux individus appartenant à des catégories ontologiques différentes puissent s'avérer indiscernables explique aussi, inversement, le rapprochement du thème du double à celui de la métamorphose. La créature tératologique, comme la chimère, incite à la conception qu'un même individu puisse appartenir à deux espèces différentes. C'est sans doute la raison pour laquelle les analyses thématiques littéraires incluent couramment *La Métamorphose* de Kafka dans les récits de double. Un même individu partage des composantes sortales diverses et contradictoires, comme c'est encore le cas dans *La Mouche Noire*, réalisé par Kurt Neumann en 1958 et

[10] On peut également mentionner ici l'hypothèse de Marcel Pagnol au sujet du Masque de fer, affirmant que le célèbre prisonnier aurait été le jumeau puîné de Louis XIV, frère caché afin de ne pas encourager de protestations sur la légitimité du roi.

repris en 1986 dans le remake *La Mouche* de David Cronenberg où le personnage est à la fois homme et mouche et dans *Alien resurection* de Jean-Pierre Jeunet où le personnage de Ripley a été cloné en fusionnant avec un Xénomorphe. Le monstre déjoue les catégories sortales qu'il mixe, il est une aberration taxonomique, une entité qui semble s'écarter de toutes les normes culturelles. Il ne donne pas de prise aux descriptions, qui échouent toujours à rendre compte de sa réalité. Il n'est pas compréhensible par concept, mais seulement « montrable » par un acte indexical, par une désignation directe, comme l'indique d'ailleurs l'étymologie même du mot.

L'indiscriminabilité artificielle peut aussi apparaître sans la moindre référence à un modèle original. Elle peut encore être le fait du déguisement ayant pour but, non d'imiter autrui, mais d'effacer sa propre personnalité, pour se fondre dans la conformité sociale. On peut illustrer ceci par l'exemple des deux inspecteurs Dupont et Dupond inventés par Hergé, qui, partageant tout, reproduisent non seulement la même physiologie (ressemblance naturelle), mais surtout les mêmes accoutrements, les mêmes allures et les mêmes occupations. Les Dupondt ne semblent pas particulièrement dérangés par leurs ressemblances frappantes, mais cherchent au contraire à les exacerber. Ils ne sont pas craintifs d'une perte éventuelle de leur singularité. Leur soucis est inverse, relevant d'un besoin intense de conformité sociale, de fusion dans l'indiscernabilité. Ils semblent constamment encouragés à ne faire qu'un. Poursuivant Tintin, les deux inspecteurs se heurtent comme pour se fondre l'un dans l'autre. Dans les lieux les plus déserts, ils suivent leurs propres traces en tournant en rond comme s'ils cherchaient à résoudre enfin leur différence dans une légitime coïncidence qu'ils ne parviennent jamais à réaliser vraiment. Quoi qu'ils fassent pour apaiser leur inextinguible soif d'unité, les Dupondt seront cependant toujours séparés d'une légère différence spatiale, et toujours distingués par les infimes différences que manifestent l'orthographe de leur patronyme et l'orientation de leur moustache, respectivement droite (d) ou troussée (t) [11]. C'est jusqu'aux paroles que les Dupondt s'apparentent, sinon que chez l'un, elles sont encore adjointes d'un « je dirai même plus » toujours démenti en ce qu'il précède la répétition stérile du propos de son compère. Comme l'analyse avec beaucoup de finesse Clément Rosset dans *Le démon de la tautologie*, là encore, les formes de cette répétition inutile encouragent diverses cocasseries langagières, selon qu'elles se rapportent à la pure tautologie (Page 5 de *L'île noire* : « Tintin est un petit gredin ! - Un gredin ? Je dirais même plus, Tintin, c'est... c'est un gredin »), selon qu'elles se rapportent au pléonasme (« Cette lettre est anonyme - Je dirais même plus : l'auteur de cette lettre ne s'est pas fait connaître »), selon que l'une d'elles soit remplacée par une onomatopée (en page 3 de *On a marché sur la Lune* : « Elle est bien bonne ! - Je dirais même plus : Ha ! Ha ! Ha ! ») ou selon qu'elles perdent tout sens par l'effet d'une perturbation dyslexique (*Cock en stock*, page 10 : « Notre métier à nous est de tout savoir - Je dirais même plus, notre métier à tous est de nous savoir »). Comme le souligne Rosset, éreintés par leurs tentatives aussi fréquentes que vaines d'en dire plus que ce qu'il y a, les deux inspecteurs perdent progressivement la capacité à exprimer quelque chose :

> Épuisés par leur tentative obsessionnelle d'éviter la tautologie en dénichant un terme introuvable qui en dirait plus tout en disant de même, les policiers, dont les circuits cérébraux

[11] Notons que cette orientation varie d'album en album. Cette maigre différence qualitative n'empêche pas Hergé de les confondre, volontairement ou non. Aussi, sauf à faire l'hypothèse d'une erreur de la part d'Hergé, on peut se demander si ce changement marque la volonté des Dupondt de s'imiter l'un l'autre en harmonisant leurs bacchantes mais sans jamais parvenir à un accord sur la manière de les arborer.

> disjonctent, en sont de toute façon réduits, sous couleur d'en dire plus, à ne plus rien dire. (ROSSET, 1997, p. 36)

Le travestissement auquel recourent à tout bout de champ les Dupondt peut parfois atteindre un degré paroxystique consistant à s'approprier la peau d'autrui, à être autrui en restant soi, à se vêtir de l'autre comme l'on rentre dans un vêtement. Dans *La Complainte de Fantomas* (1933), Robert Desnos écrit par exemple :

> Pour effacer sa trace
> Il se fit tailler des gants
> Dans la peau d'un trophée sanglant
> Dans d'la peau de mains d'cadavres
> Et c'était ce mort qu'accusaient
> Les empreintes qu'on trouvait

Mais les malversations d'identité ne nécessitent pas toutes cette pratique radicale consistant à revêtir la peau d'un autre. Il y a d'autres moyens, plus fins. Dans *Plein Soleil*, le film de René Clément tiré du roman de Patricia Highsmith, le personnage de Tom Ripley, joué par Alain Delon, tente d'usurper l'identité sociale d'un jeune et riche dandy, Philippe Greenleaf, qu'il vient d'assassiner et dont il a jeté le corps en pleine mer. Aucun témoin n'ayant constaté la disparition de Greenleaf, Ripley tente de reprendre le cours de sa vie là où elle s'est arrêtée, en usurpant son identité sociale pour profiter des largesses de son mode de vie. Afin de tromper l'entourage de la victime en faisant croire que Greenleaf vit toujours, Ripley est amené à agir à la fois sous sa propre identité, mais également sous celle du défunt. Pour cela, il use de deux stratagèmes. Le premier, se rapporte aux talents qu'évoque le titre du roman, précisément la capacité de Ripley à imiter voix et signature, ce qu'on pourrait appeler, en un sens plus général, des *signatures* d'un individu, c'est-à-dire des *indices* de sa présence. Le deuxième stratagème consiste à accréditer l'indiscernabilité par l'effacement. Dans les diverses situations où Ripley est amené à emprunter l'identité de Greenleaf, il ne le fait qu'en dissimulant son corps, de manière à ce que seules sa graphie et sa voix soient perceptibles. La stratégie menée par Ripley est basée sur l'idée qu'une personne n'est qu'un corps en mouvement, et peut être réduite à des signatures plus ou moins « naturelles » comme la voix ou la manière d'écrire.

Cependant, Ripley va découvrir progressivement qu'un individu possède de nombreuses signatures, et qu'il ne parvient pas à contrôler toutes celles de sa victime. Par exemple, son goût pour la décoration intérieure est fruste et prétentieux quand celui de Greenleaf était raffiné et élégant. Ripley est certes capable d'imiter certaines manières de Greenleaf, mais il ne parviendra jamais à substituer ses actes et gestes aux attitudes intentionnelles de sa victime, qui se manifestaient, de manière causale, dans de multiples aspects extérieurs, comme celui du choix d'un tissu de rideaux, d'un assemblage de coloris, ou de la manière de poser un objet sur une étagère. La conclusion vers laquelle le film semble s'acheminer est que pour usurper la carrière d'un homme, il ne suffit pas d'en imiter les qualités externes. Il faut aussi en connaître parfaitement le point de vue à la première personne, parce que celui-ci « émane » dans son environnement proche. Toute usurpation semble donc parfaitement vaine, en ce qu'elle sera inévitablement détrompée à une occasion.

Les indiscriminables occultifs

Dans les récits que nous avons évoqués jusqu'à présent, la duplique relève toujours d'une usurpation, et, d'un point de vue absolu, elle se distingue toujours, d'une manière ou d'une autre, de son modèle, soit par l'espèce (avatar de dieu/mortel), soit par une axiologie sociale (seigneur/serf), soit par une signature propre.

A côté de ces différents types d'indiscriminables, l'effet de duplication peut être également produit par un autre artifice que le déguisement, et en quelque sorte inverse, artifice qui relève du masque. Pensons aux masques de théâtre, aux loups, aux cagoules, aux cache-nez, aux voiles mais aussi aux éventails, qui cachent ou révèlent de manière dynamique des parties du visage. En revanche, les visages en caoutchouc que revêtent par exemple Fantomas et les scélérats des épisodes de la série de dessins animés Scooby-Doo ne sont pas, en ce sens, des masques, mais des travestissements car ils encouragent l'indiscernabilité par l'imitation d'un modèle dont il s'agit d'épuiser les ressemblances qualitatives. Le masque tend inversement à effacer les qualités afin de perdre tout moyen de comparaison. Tel est l'usage du masque dans la tragédie antique auquel recourt l'acteur pour effacer ses traits propres et se rendre, par ce moyen, indiscernable d'un héros qu'il incarne. Dans sa préface à *Histoire de Masques* (1900) (LORRAIN, 1900), Jean Lorrain constate d'ailleurs que le masque, fut-il le plus chamarré et le plus drôlatique, est toujours inquiétant, il suscite toujours, et à la fois, attirance et répulsion. Comme dans la gravure *L'Entrée du Christ à Bruxelles* de James Ensor, les foules masquées déambulant dans les avenues prennent l'apparence d'une sorte de masse informe où l'individu disparaît, le carnaval effaçant autant les différences physionomiques, sexuelles et sociales. Alors que le déguisement a pour tendance d'exacerber les similarités avec ce qu'il prétend dupliquer, le masque a au contraire pour principe de gommer tous les traits qui pourraient servir de comparaison, ceci afin de rendre les individus indiscernables par effacement de leurs caractéristiques propres. À l'inverse du travestissement qui recherche une indiscernabilité *ostensive*, comparative, le masque est un artifice d'indiscernabilité *occultive* : il rend impossible toute comparaison. L'indiscernabilité occultive est donc l'enjeu d'un rapport qui est essentiellement de perspective, et joue en ce sens un rôle pleinement épistémique. De manière plus générale, je désignerai par l'appellation « d'effet de masque » tout procédé permettant de dissimuler une partie, voire la totalité des traits d'un personnage.

Les effets de masque mettent aussi en question la vérité qu'ils sont censés dissimuler : comme prévient le titre d'un dessin de Max Ernst, *Attention : un masque peut en masquer (ou démasquer) un autre* (1949), on risque toujours de trouver, sous un masque, un autre masque. L'effet disruptif est aussi provoqué par la rencontre des apparences, par les visages fondus, ceux que le photographe Krzysztof Pruszkowski appelle *Photosynthèse*, comme son *De Gaulle, Pompidou, Giscard, Mitterrand : le Président de la Ve République française* (1984) où les visages des quatre présidents sont fondus les uns aux autres, où chacun surgit pour laisser aussitôt l'autre apparaître, à l'instar des récits de science-fiction de Philip Kindred Dick, où les protagonistes portent des masques digitaux qui révèlent une multitude de visages différents se succédant les uns aux autres.

Les effets de masque, en littérature, relèvent cependant moins de l'effacement et de la dissimulation de traits que d'une information auctoriale défaillante. Le double par effet occultif prend appui sur le caractère incomplet des personnages. A cet égard, et pour en illustrer l'effet, citons un passage d'un roman d'Arcadi Stougatski et de Boris Stougatski, intitulé *Le lundi commence un*

samedi, dans lequel un personnage visite des mondes fictionnels au moyen d'une machine de son invention :

> Je préférais regarder les gens. [...] Je rencontrais des personnages qui n'étaient vêtus qu'en partie ; par exemple un chapeau vert et une veste rouge sur un corps nu (rien d'autre) ; ou bien des souliers jaunes et une cravate à fleurs (sans pantalon, sans chemise ni même linge de corps), d'élégantes chaussures enfilées sur des pieds nus. Les passants ne réagissaient pas, moi j'étais très gêné, et puis je me souvins que certains auteurs ont l'habitude d'écrire des phrases de ce genre : "La porte s'ouvrit, un homme élancé et musclé, en casquette et lunettes noires se montra dans l'encadrement". (A. STROUGATSKI & STROUGATSKI, 1974, pp. 207-208)

L'indiscrimination occultive peut d'ailleurs fonctionner de différentes manières en littérature, en rapport avec la manière dont sont généralement identifiés des personnages de fiction. Leur existence peut être suscitée par la seule évocation d'un nom propre sans qu'aucune description ne leur soit directement attribuée, comme ce Godot qu'attendent vainement les personnages de la pièce de Samuel Beckett. Inversement, les personnages peuvent aussi faire l'objet d'une description sans avoir de nom, comme par exemple la jeune gouvernante dans *Le Tour d'Écrou* d'Henry James (M. R. JAMES, 1924).

Dans *Monsieur Klein*, le réalisateur Joseph Losey a réussi le tour de force de mettre en scène une indiscriminabilité au seul moyen d'une homonymie. Le double s'avère indiscriminable du personnage principal non en raison d'une part excessive de qualités partagées, mais au contraire, d'un défaut total de traits. Car on ne sait presque rien du double de Robert Klein sinon qu'il porte le même nom. La conséquence inattendue, mais néanmoins cohérente, est que rien ne permet de discriminer le personnage de son double, en dépit des vaines tentatives que celui-ci produit pour faire cesser l'usurpation d'identité dont il déplore être la victime. Le spectateur est alors lui-même en butte sur la question du bon dénombrement des personnages. Le sentiment du double n'est pas seulement celui que ressent le personnage principal, mais aussi celui du spectateur impliqué lui-même dans ce jeu de dénombrement.

Dans ce film, Alain Delon joue le rôle d'un homme qui tente de justifier son identité aux autorités durant l'occupation allemande. C'est que Robert Klein s'inquiète d'une grave méprise à son égard, puisqu'on le confond avec un homonyme juif traqué par la police. Par crainte que la police ne parvienne pas à démêler le vrai du faux, Klein, entame sa propre enquête afin de retrouver son double onomastique. Mais ce dernier manifeste la fâcheuse tendance à s'avérer aussi invisible qu'insaisissable. Au fil de l'intrigue, la distinction numérique stipulée au début du film entre les deux personnages tend, pour le spectateur, à s'effacer progressivement. Monsieur Klein, en poursuivant son double, en vient à fréquenter les lieux qu'occupe ce dernier, et se livre finalement aux mêmes activités. Tant et si bien qu'à la fin, le personnage incarné par Delon remplit complètement le rôle de son mystérieux double, qui sera demeuré, pour le personnage autant que pour le spectateur, imperceptible. Alors qu'il n'a jamais été aussi près de « serrer » son invisible adversaire, Monsieur Klein est finalement embarqué dans la rafle du Vélodrome d'Hiver, incarnant complètement le destin de celui qu'il poursuivait. Il finit donc par coïncider pleinement avec lui.

Contrairement à la plupart des fictions du double, l'originalité du scénario de Losey consiste en ce que son intrigue est construite sur le schéma d'une identité *octroyée*. Car le personnage principal n'y perd pas son identité sociale au profit d'un autre individu qui s'en serait accaparé,

comme c'est encore le cas dans nombre de récits de doubles du XIXe siècle. Ici, c'est bien plutôt le contraire : le personnage est doté d'une identité surnuméraire, une identité que non seulement il ne revendique pas être la sienne, mais plus encore qu'il rejette entièrement. Le personnage ne subit pas une usurpation d'identité, mais, si je puis dire, un gain indésirable. Notons aussi que Klein et son double sont indiscriminables tant par les propriétés intrinsèques que par les propriétés extrinsèques. Il s'agit d'une indiscriminabilité qui est proche de ce que nous avons appelé, en métaphysique, double pur.

Les doubles par fission

Dans la littérature du XIXe siècle, la figure du double étend sa généralité, et en vient à concerner les jeux d'identifications métaleptiques entre lecteur, auteur, personnage et narrateur. C'est sans doute les nouvelles de Nicholas Gogol qui initient une progression en ce sens. Dans *Le journal d'un fou*, le narrateur entretient des rapports serrés avec le personnage. L'attention du lecteur est centrée sur l'expérience intime du personnage, ce qu'autorise la forme même du journal intime relatant la décomposition mentale progressive du protagoniste par lui-même. Dans le *Le nez*, la place du narrateur omniscient est progressivement occupée par Kovaliov, le personnage en butte au double [12].

Le procédé diégétique des nouvelles de Gogol a été repris par Dostoïevski. Publié une première fois en 1846 puis en 1865 dans sa version définitive, le deuxième roman de Fédor Dostoïevski, intitulé *Le Double* (*Dvoïnik*) relate les aventures d'un fonctionnaire de l'administration Russe, Jacob Pétrovitch Goliadkine, confronté à la rencontre brusque avec un collègue de travail qui lui est en tout point indiscernable : mêmes traits, même nom, mêmes fonctions administratives, même bureau, même logement.

Le roman fut globalement mal accueilli par la critique. Mais, de l'aveu même de son auteur, *Le Double* occupe une place primordiale dans l'ensemble de son œuvre. Dostoïevski considère en effet qu'il est le premier auteur à défricher la question de la fragilité de l'identité sociale. Ainsi, alors qu'il doit défendre, en 1859, auprès de son éditeur la seconde publication du roman à l'intérieur de l'édition en deux volumes de ses œuvres, Dostoïevski écrit à son frère : « Pourquoi laisser perdre une magnifique idée, un type grandiose par son importance sociale, que j'ai découverte premier et dont j'ai été l'annonciateur ? ». En 1860, l'auteur écrit encore : « J'ai positivement raté cette nouvelle, mais l'idée en était assez lumineuse, et je n'ai rien lancé dans la littérature de plus sérieux que cette idée... ».

Les rapports troubles que Goliadkine entretient à son moi n'attendent d'ailleurs pas l'apparition du Doppelgänger. Au début du récit, Goliadkine croise incidemment son supérieur hiérarchique dans une situation qu'il estime gênante. Pris de panique, il cherche à se soustraire aux salutations d'usage par l'artifice consistant à faire comme s'il ne s'agissait pas de lui : « faire semblant que ce n'est pas moi, que c'est quelqu'un d'autre, quelqu'un qui me ressemble à s'y

[12] On peut également noter, dans le récit, un élégant jeu métaleptique. Kovaliov cherche à passer une petite annonce dans la presse, promettant une récompense à celui qui lui ramènerait son nez. L'employé du bureau d'annonce refuse de publier une telle absurdité, mais, témoin de l'absence de nez de Kovaliov, lui conseille de s'adresser aux compétences d'un écrivain pour en relater l'histoire. Rappelons à cet égard que la nouvelle, jugée inepte, fut d'abord refusée par la revue littéraire *L'Observateur moscovite*, avant d'être publiée en octobre 1836 dans la revue *Sovremennik*.

méprendre, et faire comme si de rien n'était ? Mais non, ce n'est pas moi, ce n'est pas moi et voilà tout ! » (DOSTOÏEVSKI, 1969, p. 8) Goliadkine veut ainsi apparaître exactement comme il est pour paraître numériquement différent. De fait, il y a quelque chose d'extrêmement trouble dans la volonté de faire paraître à quelqu'un à qui l'on fait face, qu'il ne s'agit en réalité pas de soi mais d'un autre, puisque cela contraint à faire comme si on était autre, tout en maintenant le droit de l'autre à se reconnaître comme étant le même. La stratégie de Goliadkine est comme enroulée sur elle-même, prise dans une auto-réflexivité paradoxale : « Oui, vous avez bien jugé, pourrait dire Goliadkine, la personne qui est là paraît être celle que vous connaissez. Mais justement, puisqu'elle lui ressemble, ce n'est pas elle. »

La tentative décrite par Goliadkine consiste en quelque sorte à « disparaître par la présence ». C'est sans doute le processus même de narration, pour lequel le narrateur prétend s'effacer devant ses personnages. Le mérite du roman de Dostoïevski est principalement de questionner leurs places respectives. Comme le remarque Mikhaïl Bakhtine, si, au début de l'histoire, le récit est raconté du point de vue omniscient, le flux des pensées de Goliadkine prend progressivement la place du narrateur :

> Avec un goût et un art stupéfiants Dostoïevski oblige la seconde voix de Goliadkine à passer, presque insensiblement et sans que le lecteur s'en aperçoive, de son dialogue intérieur dans le récit lui-même : elle commence à sonner alors comme voix du narrateur, voix d'un autre. (BAKHTINE, 1970, p. 249)

Ce procédé a pour effet de rendre indiscriminables les places respectives du personnage et du narrateur à l'intérieur même du processus diégétique.

Pour le lecteur, l'indiscernabilité n'est sans doute pas tant celle qui existe entre le personnage et son doppelgänger. Le lecteur du *Double* est lui-même amené à voir deux personnes là où ne s'en tient peut-être qu'une seule. Au fil du roman, Goliadkine et son doppelgänger sont respectivement distingués par des locutions comme « le vrai Goliadkine », « l'aîné » et « Goliadkine deux », « le cadet », « l'autre », « le sosie ». Les expressions accentuent l'effet de réification du double. Mais leur ressemblance ne semble pas tant frappante aux yeux de leur entourage. Quand Goliadkine attire l'attention sur la ressemblance qu'entretient son double avec lui-même, ses collègues s'avèrent tout au plus amusés, sinon indifférents, rétorquant que « ça arrive » et évoquant le cas des frères siamois. Seul Goliadkine semble affecté par cette duplication étrange, à l'exception, doit-on préciser, de son domestique, qui reproche à son maître que « les gens bien, ils vivent comme il faut, ils font pas d'histoires malhonnêtes, les gens bien, ils sont jamais deux à la fois [...] » (DOSTOÏEVSKI, 1969, p. 163) Notons toutefois que les doléances du domestique sont prononcées alors qu'il est sous l'emprise de l'alcool, ce qui favorise sans doute sa perception double sans qu'on puisse savoir si elle persisterait sans ébriété.

C'est que l'indiscriminabilité est surtout logée à un autre niveau, entre le narrateur omniscient et le personnage. C'est là que réside l'originalité du roman de Dostoïevski, et qui explique l'importance que son auteur lui attribue parmi ses autres romans. Dans ce discours polyphonique à une seule personne, le solipsisme de Goliadkine prend la forme d'une totalité fermée, les événements réels étant systématiquement filtrés et réinterprétés selon le délire de persécution du personnage. Le lecteur, qui n'a plus le recul du point de vue omniscient comme c'est encore le cas pour Sosie dans la pièce de Plaute, est donc associé à ce parti pris, à l'identification du personnage à son

double. Au fur et à mesure de l'avancée dans le récit, la distinction entre le narrateur et Goliadkine s'amenuise, cette ambiguïté ayant pour effet d'accroître le sentiment de confusion générale. Il n'y a qu'un pas à faire pour interroger alors le rapport entre l'auteur et le personnage, pas que franchit Dostoïevski en écrivant à son frère « Je suis désormais un vrai Goliadkine » (TROUBETZKOY, 1995). Le double n'est plus seulement une figure littéraire, celle des doppelgängers, mais un thème présent dans l'intégralité de la littérature, interrogeant les places respectives des personnages, du narrateur, de l'auteur, mais bien aussi, par extension, celle du lecteur.

Avec le récit de Dostoïevski apparaît une nouvelle manière d'interroger l'indiscriminabilité : la disparition progressive d'un repère absolu, à savoir le narrateur omniscient, pour dénombrer de manière objective les personnages. Ce processus incite à un perspectivisme qui ne permet plus de rendre compte d'une distribution relationnelle des propriétés : pour cette raison, le récit de Dostoïevski se rapproche d'un cas de double pur, en ce que le lecteur ne dispose plus de l'appui d'un espace absolu permettant de discriminer de manière spatiale les individus en présence.

Les doubles fissionnels et fusionnels

La démultiplication peut encore s'approcher, en littérature, du cas métaphysique des coïncidents purs, lorsque le personnage et son double semblent coïncider pleinement, et que rien ne semble les séparer sinon le sentiment vague, pour le personnage, d'une démultiplication des entités, de la présence d'un individu qui est comme ajouté à sa propre existence. C'est en particulier le cas du *Horla*, et sans doute encore plus de *Lui !*, deux contes écrits par Guy de Maupassant. Ici, l'identification du personnage à son double semble prendre une part paroxystique, et semble peut-être inconcevable. Pour cette raison, nous y consacrerons une analyse plus poussée dans la dernière section de ce chapitre.

Dans toutes les fictions examinées précédemment, la situation de double est produite par l'existence d'individus ayant pour effet la réfutation du principe d'identité des indiscernables. Précisément, ils tiennent de situations où le jugement plus ou moins avoué que des choses indiscernables devraient légitimement être des choses identiques se trouve réfuté. Mais la thématique littéraire du double prend aussi appui sur le jugement inverse, à savoir que des entités identiques devraient s'avérer, synchroniquement, exactement ressemblantes. Dans la fiction, le double ressortit non seulement à la démultiplication, c'est-à-dire de la ré-itération d'une entité reproduite, mais également à la division, au fractionnement d'une même entité éclatée dans le temps et l'espace. Le dédoublement procède alors en fission, en parties déconnectées.

La fiction permet d'éprouver ce dernier principe de différentes manières. Le roman de Dostoïevski apparaît ici comme un cas limite, le double n'étant plus seulement vu comme un double pur, une altérité exactement ressemblante. Il revêt aussi la possibilité de la mêmeté par une démultiplication d'une seule et même psyché occupant plusieurs corps dans l'espace, un double-même, un individu dont les parties spatio-temporelles sont dispersées d'une manière extraordinaire. En ce sens, le double questionne les limites, les frontières de l'individu : qu'est-ce qui n'en fait plus partie, qu'est-ce qui en fait encore partie, où tracer la ligne de démarcation ? La figure du miroir n'est pas seulement celle qui revoie l'image d'une ressemblance parfaite, mais encore celle d'une délimitation, renvoyant à celui qui s'y cherche les limites de l'entité qu'il est. Le miroir offre une possibilité de cerner les contours des corps. Mais paradoxalement, il réfléchit le corps, autrement

dit, il lui inscrit des frontières sur une image qui, justement, n'est déjà plus le corps physique. Cette image projette non seulement le corps ailleurs dans l'espace, c'est-à-dire, sur le plan opaque du miroir, mais le projette également ailleurs, dans un espace qui n'est plus physique, mais purement virtuel. Nous sommes, en quelque sorte, tous des Goliadkine parce que nous créons sans cesse un individu virtuel projeté au devant de nous afin de parvenir à démarquer nos propres frontières. Le double apparaît dès lors comme une extension corporelle.

Dans la fiction, l'extension virtuelle qu'apporte le double à l'individu peut prendre différentes formes : soit qu'il s'agisse d'une extension matérielle, d'un prolongement du corps sous une forme non connexe, séparée, mise à distance d'elle-même (comme dans *Le Nez* de Gogol), soit qu'il s'agisse d'un fractionnement psychologique lié à une subdivision de la personnalité (comme dans *Le Double* de Dostoievski), soit encore qu'il s'agisse d'un prolongement du Moi dont les occurrences s'étendent à travers le temps. La croyance en une substance immatérielle du moi qui persisterait par endurance dans le temps et la psychologie de l'identité personnelle sont des appuis qui permettent aux fictions de développer d'hypothétiques pathologies de l'identité, comme le cas d'une même personne habitant plusieurs corps, ou comme le cas de fusion de plusieurs personnes logées dans le même corps. La littérature a exploité l'idée de corps multiples habités par une seule et même personne, notamment le principe de la réincarnation d'une âme dans des corps différents, comme dans le roman *Orlando* (1928) de Virginia Woolf (WOOLF, 1928), en encore comme dans le film *Cloud Atlas* de Lanna Wachowski et Lilly (Léna WACHOWSKI & WACHOWSKI, 1999).

D'un point de vue perdurantiste, nous avons vu dans la première partie qu'un individu n'est pas une entité endurante, c'est-à-dire présente à tout moment de sa carrière temporelle, mais qu'elle est la somme méréologique de toutes ses parties spatio-temporelles. Cela permet de donner une explication logique aux paradoxes qui émergent dans les histoires de voyages dans le temps, lorsqu'un personnage se retrouve à proximité d'une de ses parties corporelles appartenant à une période passée. À la fin du film *Retour vers le Futur* de Robert Zemeckis, le personnage principal revient d'un voyage dans le passé quelques minutes avant son départ. Il a ainsi l'occasion d'apercevoir l'une de ses parties temporelles. Leurs différences chronologiques étant moindres, les deux occurrences possèdent synchroniquement la même identité qualitative sans néanmoins partager la même identité temporelle, puisque l'un a voyagé dans le temps, mais l'autre pas (encore). Mentionnons aussi la série télévisée *Dark* (ODAR & FRIESE, 2017), qui, s'inspirant du film de Zemeckis, en accroît encore les paradoxes. Dan cette série, les incessants voyages temporels des protagonistes les mènent à rencontrer certaines de leur occurrences temporelles plus âgées ou plus jeunes.

Les cas de fusions fictionnelles présentent à l'inverse l'idée de différents personnages logeant dans un même corps. Citons le roman *Body Snatchers* (1955) de Jack Finney dans lequel des envahisseurs extra-terrestres s'approprient les corps, la personnalité et les comportements d'individus humains [13].

[13]Il s'agit non seulement d'une expérience fictionnelle de pensée, mais également d'un syndrome délirant relevé par le psychiatre Joseph Capgras (CAPGRAS & REBOUL-LACHAUX, 1923) en 1923 sous le nom de « délire d'illusion des sosies ». C'est une psychose chronique non dissociative dans laquelle le patient soutient l'idée que ses proches ont été remplacées par des sosies parfaits, mais qu'il est certain de cette usurpation en vertu d'infimes différences dans la couleur des yeux, le comportement, etc. Contrairement aux cas de prosopagnosies (décelée en 1947 par le neurologue Joachim Bodame), dans lequel le patient, bien que possédant une acuité visuelle normale, est incapable d'associer un visage à un individu, les personnes qui souffrent de la maladie de Capgras sont tout à fait capables d'identifier les visages.

D'autres doubles par fission peuvent encore relever de figures synecdotiques par des cas de séparation des parties propres d'un individu amenées à produire deux occurrences autonomes et non connexes. La séparation chirurgicale des frères siamois sert aux récits de doubles comme dans *Dead Ringers* de David Cronenberg. Mais les fissions peuvent aussi s'avérer bien plus fantaisistes. Dans *Le Nez* de Gogol, la-dite partie, détachée du visage de Kovaliov mène grande vie de conseiller d'État au grand dam de son propriétaire. La transformation peut également relever de la seule psyché, comme dans *L'Étrange Cas du Docteur Jekyll et de Mister Hyde* de Robert Louis Stevenson, ou dans le film *Psychose* d'Alfred Hitchcock adapté du roman de Robert Bloch où un même personnage fissionne en deux personnalités différentes successives.

Certains récits centrent la thématique du double sur les relations qui se tiennent entre un individu et son ombre, son reflet ou son écho, autrement dit, aux phénomènes particuliers qui sont l'effet de leur existence, et dont la persistance temporelle dépend normalement de l'individu qui en est la cause. De fait, certaines fictions évoquent la rupture du lien qui unit la personne à ses *effets spéculaires*. Ceux-ci s'en désolidarisent : l'ombre que Peter Schlemihl cède au diable (CHAMISSO, 1973), celle que perd Peter Pan (BARIE, 1902), celle capricieuse du conte d'Andersen (ANDERSEN, 1876). Le cas du reflet est également intéressant. Le reflet peut se désolidariser de l'individu, comme celui, perdu, du vampire, voire gagner en autonomie, comme dans *Le Portrait de Dorian Gray* d'Oscar Wilde où le portrait du protagoniste vieillit à sa place.

De manière générale, les parties détachées, en usurpant la place entière de la personne dont elles émanent incitent à la pensée de ce qu'auraient pu être ces personnes. Elles constituent alors une sorte de halo de possibles confrontant la personne aux multiples carrières alternatives qu'elle aurait pu connaître. On retrouve alors l'idée de mondes possibles.

5.1.3 Les doubles importés

Les personnages récurrents transfictionnels

Il arrive que des personnages migrent d'une fiction à une autre. Ils mènent ainsi une carrière qui dépasse les limites des œuvres où ils ont été créés. Par exemple, Holmes est le protagoniste de différents romans de Conan Doyle, apparaissant pour la première fois en 1887 dans le roman *Une Étude en Rouge*, et pour la dernière fois en 1927 dans le recueil de nouvelles *Les Archives de Sherlock Holmes*. Il apparaît aussi dans des pièces de théâtre et des pastiches, dans des adaptations radiophoniques et cinématographiques, dans des séries télévisées, dans des dessins animés, dans des bandes dessinées, dans différents jeux vidéos, dans des jeux de société et dans des livres-jeux. En outre, il est mentionné aussi dans les critiques littéraires, les exégèses et les biographies imaginaires qui lui sont consacrées. Entre ces différentes fictions et ces différents supports, Sherlock Holmes demeure identifiable en dépit de modifications progressives (séries, romans) ou brusques (pastiche) qui peuvent altérer sa description lors de ses réapparitions.

Différents types d'histoires comportent des personnages récurrents. Certaines leur attribuent un rôle important et d'autres peuvent à peine l'évoquer, comme par exemple le personnage de Meursault de *L'Étranger* d'Albert Camus, dont le crime est relaté, dans *La Peste*, lors d'une conversation dans un bureau de tabac, sans que son nom ne soit mentionné. Lorsque le personnage récurrent occupe une place importante dans une nouvelle œuvre, nous avons alors affaire à

une sorte de *prolongement* de sa carrière initiale. Notons que le temps du récit qui constitue ce prolongement ne suit pas nécessairement la chronologie de l'histoire, car il peut concerner des événements ultérieurs au récit original - on parlera dans ce cas de séquelle, aussi bien que des événements antérieurs - on parlera alors de préquelle, voire d'événements intermédiaires (interquelle). En outre, il arrive aussi que le récit orignal soit tout bonnement *recommencé* (réquelle), en ce sens que certaines des affirmations du créateur de la fiction originale sont contredites dans la nouvelle fiction, comme c'est par exemple le cas dans le roman de Philippe Doumenc *Contre-enquête sur la mort d'Emma Bovary* (DOUMENC, 2007) où l'auteur conteste le suicide de Madame Bovary en défendant l'hypothèse d'un assassinat. Le Petit Chaperon Rouge, Antigone, Emma Bovary, Anna Karénine pourraient-elles ne pas mourir ? Dans l'histoire de Perrault, le Petit Chaperon meurt, mais pas dans celle des frères Grimm. L'histoire littéraire nous montre que les altérations de carrière d'un personnage sont possibles, voire souvent pratiquées. La fiction y est un refuge pour faire et refaire le monde. Lorsque Flaubert crée Madame Bovary, il lui attribue des propriétés importantes, d'autres qui le semblent moins et enfin il laisse tout un tas de propriétés indéterminées. Parmi ces propriétés, on pourra citer celle d'être une bourgeoise s'ennuyant de son mariage, femme qui, adultère et menant un train de vie qu'elle ne peut assumer, finit par se suicider. Un personnage en tout point ressemblant à la différence qu'il ne se suiciderait pas serait-il encore Madame Bovary ? Selon Umberto Eco, ce n'est pas le cas (ECO, 2013, p. 120), le suicide de Madame Bovary est nécessaire :

> Du fait qu'ils sont des entités fluctuantes, ces fidèles compagnons de nos vies (à la différence d'autres objets sémiotiques), qui sont culturellement soumis à la révision) ne changeront jamais et resteront pour toujours les agents de leurs actions. (ECO, 2013, p. 121)

Et pourtant, il y a bien un sens dans lequel nous pouvons admettre le fait que Bovary aurait pu ne pas se suicider. Les personnages ont des caractéristiques fictionnelles modales, ce que j'appellerai des propriétés contrefictionnelles. Le lecteur doit se poser ces questions : pourquoi le suicide, n'y avait-il pas d'autre solution ? Quitter son mari, disparaître, vendre sa maison pour payer les frais ? Nos réécritures intérieures du roman, nos pensées contrefictionnelles, aiguisées par les orientations données par l'auteur, ouvrent sur la dimension modale de l'œuvre. La compréhension du récit de Camus relatant le crime de Meursault dépend étroitement de l'idée qu'il aurait pu ne pas le commettre. La modalité est inscrite au cœur même du récit.

Qu'est-ce qui fait la récurrence d'un personnage ? La question est de savoir si l'on a affaire littéralement au même *fictum* ou à un autre, si la question est vraiment pourvue de sens. Constatons qu'affirmer que c'est l'identité qui corrèle les diverses occurrences d'un personnage récurrent implique que celui-ci soit nécessairement affecté par son contenu. Par le fait même de réapparaître dans un nouveau récit, un personnage se voit nécessairement attribuer de nouvelles propriétés. Autrement dit, sa description s'enrichit au long d'une carrière prolongée par des actes narratifs. Car il est inévitablement modifié, si ce n'est directement par des propriétés intrinsèques, du moins par des propriétés extrinsèques. Par exemple, le personnage de Meursault à peine évoqué dans *La Peste* est affecté d'une vie dans un monde où sévit la peste à Orange dans les années 40 et où vit un certain docteur Rieux.

Les personnages récurrents métafictionnels

On peut se demander si les histoires de doubles constituent des cas particuliers de récits à personnages *récurrents dans leur propre récit*. Remarquons qu'entre deux interruptions de lecture d'une même et unique histoire, nous retrouvons certes le même personnage, mais que nous ne parlerions pas de récurrence. Ainsi, le fait d'apparaître dans *deux histoires différentes* semble constituer une condition nécessaire de récurrence d'un personnage. C'est donc dans un sens métaphorique qu'on parlera de récurrence d'un personnage dans les histoires de doubles, par exemple au sens où William Wilson et où Goliadkine semblent réapparaître une seconde fois au cours du récit.

Il y a toutefois un cas problématique de récurrence qui survient quand un personnage réapparaît dans sa propre histoire, sous un statut ontologique différent, par exemple comme personnage plutôt que comme personne. Autrement dit, quand l'auteur fait deux usages différents du même nom, à la fois fictionnel et métafictionnel. Je qualifierai de métaleptique cette sorte de récurrence. De fait, le langage permet autant de parler des choses qui lui sont externes que, de manière purement réflexive, de lui-même. Un nom peut ne dénoter rien d'autre qu'un autre nom. C'est alors comme si le langage s'enroulait sur lui-même, avec possibilité de régression à l'infini. Lewis Carroll en donne un exemple bien connu à travers les embarras d'Alice s'affairant vainement à connaître le nom de la chanson du Cavalier. Mais ce dernier lui indique obstinément « le nom du nom du nom, etc. » de sa ballade :

> Le nom de cette chanson est appelé : « Yeux de Haddock » . - Ah, c'est ça le nom de la chanson ? fit Alice en tâchant de paraître intéressée. Le cavalier parut un peu agacé. - Non, tu ne comprends pas. C'est ainsi qu'on appelle son nom. En réalité, le nom de la chanson est : « Le Vieux Vieillard » . - Alors j'aurais dû dire : « C'est comme ça qu'on appelle la chanson ? » se corrigea Alice. - Non, pas du tout, ça n'a rien à voir. On appelle cette chanson : « La fin et les moyens ». Mais c'est seulement comme ça qu'on l'appelle, vois-tu ? Alors quelle est cette chanson ? demanda Alice, complètement perdue. - J'y venais. En réalité, cette chanson est « Assis sur la Barrière » ; et c'est moi qui ai inventé la musique. (CARROLL, 2010)

Si Carroll rompt finalement le cercle vicieux et si Alice finit par connaître le nom de la chanson, on constate que la récurrence entamée pourrait se prolonger indéfiniment. Dans ce processus de cascades de dénominations, les signes sont entraînés dans un flux où ils ne renvoient plus à des objets externes mais à d'autres signes encore, c'est-à-dire au langage lui-même. Le langage est ainsi capable de se circonscrire à son propre domaine, de désigner par des mots seulement d'autres mots, en décevant ainsi nos attentes d'un objet extérieur sur lequel porterait leur dénotation.

Dans un sens plus général relevant du procédé littéraire de mise en abyme, Lucien Dällenbach parle de récits spéculaires. Plus précisément, Dällenbach définit la réflexion comme un « énoncé qui renvoie à l'énoncé, à l'énonciation ou au code du récit » (DÄLLENBACH, 1977, p. 62). Un énoncé qui procède de la réflexivité fonctionne ainsi sur deux niveaux sémantiques, à savoir, le niveau du récit et le niveau critique où « il intervient comme élément d'une méta-signification permettant au récit de se prendre pour thème ». (DÄLLENBACH, 1977, p. 62). C'est par exemple ce qui arrive dans *Hamlet* au moment où le personnage donne une représentation théâtrale de sa propre situation. La pièce de Shakespeare vient donc à être représentée en tant que telle à l'intérieur d'elle-même.

De son côté, dans *Figure III*, Gérard Genette définit la métalepse comme l'« intrusion de narrateur ou du narrataire extradiégétique dans l'univers diégétique (ou de personnages diégétiques dans un univers métadiégétique), ou inversement » (GENETTE, 1972, p. 244). Dans *Métalepses*, il précise qu'elle est « une relation causale particulière qui unit, dans un sens ou dans l'autre, l'auteur à son œuvre, ou plus largement le producteur d'une représentation à cette représentation elle-même » (GENETTE, 2004, p. 14). Genette distingue essentiellement cinq types de métalepses. Le premier type, la métalepse auctoriale, relève de la violation du pacte de fiction par l'auteur. C'est par exemple le cas quand le narrateur feint d'agir sur ses personnages, notamment quand Virgile, au quatrième chant de l'*Énéide*, affirme avoir fait mourir Didon. Le second type de métalepse est ontologique. Il est caractérisé par une sorte de déplacement du narrateur à l'intérieur des niveaux de fiction, vers un niveau infradiégétique enchâssé dans la fiction ou vers un niveau métadiégétique. Citons en exemple le cas de la nouvelle de Woody Allen intitulé *L'épisode Kugelmass* (1980) dans lequel le protagoniste de l'histoire va séduire Madame Bovary pour la ramener ensuite à New York. Le troisième type de métalepse est également ontologique. L'auteur feint d'impliquer le narrataire au niveau diégétique ou d'impliquer un personnage comme narrataire à un niveau métadiégétique. Il s'agit d'une sorte de substitution fictionnelle entre le narrataire et le protagoniste. C'est le cas de *Continuité des parcs*, récit de Julio Cortázar où un lecteur devient la victime d'un assassinat décrit dans le roman. Le quatrième type de métalepse, que Marie-Laure Ryan qualifie de métalepse rhétorique pour l'opposer aux métalepses ontologiques, concerne une sorte de manœuvre opérée par le narrateur qui feint de différer le récit par une sorte de pause. On en trouve de nombreux exemples chez Balzac, notamment dans le passage suivant : « Pendant que le vénérable ecclésiastique monte les rampes d'Angoulême, il n'est pas inutile d'expliquer le lacis d'intérêts dans lequel il allait mettre le pied. ». De même, à la fin du premier chapitre du *Roman comique* de Paul Scarron, le narrateur affirme que « l'auteur se reposa quelque temps et se mit à songer à ce qu'il dirait dans le second chapitre. » (SCARRON, 1981, p. 67). Le passage suivant de *Jacques le fataliste* de Diderot manifeste une métalepse rhétorique en plus d'une métalepse diégétique :

> Si j'allais aussi mettre ma tête sur un oreiller, en attendant le réveil de Jacques et de son maître, qu'en pensez-vous ? Le lendemain, Jacques se leva de grand matin, mit la tête à la fenêtre pour voir quel temps il faisait, vit qu'il faisait un temps détestable, se recoucha, et nous laissa dormir, son maître et moi, tant qu'il nous plut. (D. DIDEROT, 2006)

Le cinquième type de métalepse défini par Genette est la fausse métalepse, ou pseudo-métalepse. Elle consiste en une substitution du narrateur extradiégétique au narrateur infradégiétique ou métadégiétique.

Les deux types ontologiques de métalepses peuvent être encore différenciés selon qu'ils relèvent de la métalepse extérieure, lorsque le déplacement se fait entre le niveau diégétique et extra-diégétique, ou de manière interne lorsqu'il se fait entre deux niveaux de profondeur diégétique. L'effet de mise en abyme incite en effet à considérer que la fiction ne relève pas d'un unique plan, mais qu'elle doit se différencier sur différents niveaux d'immersion. Ainsi, les duplications fictionnelles ne sont pas seulement de surface, mais se font également selon des paliers de profondeurs fictionnelles. Il y a donc deux axes orthogonaux du double, l'un sur lequel les personnages peuvent être copiés, dédoublés, et un autre sur lequel ils peuvent être « plongés » plus ou moins

profondément. Ces niveaux peuvent parfois s'appeler les uns les autres, à la manière d'un ruban de Möbius. C'est un fait que certains auteurs jouent sur la mise en abîme auto-référentielle, sur une sorte d'enroulement du langage sur lui-même. Cette possibilité montre que le problème de la fiction n'est pas seulement de la démarquer du réel dans un rapport de duplication métaleptique extérieur-intérieur, mais également de constater l'effet de *feuilletage*, de stratifications en niveaux fictionnels plus ou moins profonds à l'intérieur d'une même œuvre.

Ainsi, certaines histoires décrivent des évènements, des lieux, des personnages qui sont ne sont pas réels à l'intérieur du cadre du jeu de croyance feinte qui les décrit. Par exemple, dans *Le Père Goriot*, Balzac campe le personnage de Vautrin, en le caractérisant comme un ancien négociant. Toutefois, aux deux tiers du récit, il est révélé qu'il s'agit là d'une identité d'empreint, dissimulant un évadé du bagne dénommé Colin. Dans *Grammaire d'objets en tous genres* (DESCOMBES, 1983), Vincent Descombes souligne que la révélation de l'usurpation indique un décalage référentiel à l'intérieur même de la fiction. Car si, d'un point de vue externe, les noms « Vautrin » et « Colin » désignent un même et unique personnage, du point de vue interne, Vautrin est une invention fictionnelle de Colin. Vautrin est une fiction dans une fiction. Comme nous l'avons déjà évoqué dans le chapitre 3, on pourra ainsi dire, de manière externe, que là où le nom de Colin désigne un personnage 1-fictionnel, celui de Vautrin désigne un personnage 2-fictionnel. Et là encore, cette effet de feuilletage fictionnel peut s'interpréter par des jeux de relations entre mondes possibles.

Considérons une autre illustration de récurrence métaleptique, celle-ci cinématographique, tirée de *L'Antre de la Folie* de John Carpenter (CARPENTER, 1995). L'intrigue du film est articulée sur la problématique de la perméabilité de la fiction au réel, du réel à la fiction, mais aussi et surtout, de la perméabilité entre différents niveaux de profondeur fictionnelle. Le réalisateur joue continûment d'un effet de feuilletage sémantique, en passant d'un niveau fictionnel à un autre plus profond ou, inversement, en revenant vers un niveau plus superficiel. Esquissons le synopsis du film : le personnage de John Trent (joué par Sam Neill) est expert en assurances. Son occupation professionnelle consiste donc essentiellement à révéler à ses mandataires les tentatives de fraudes à l'assurance qui les visent. Trent doit effectuer un contrat pour une maison d'édition qui le sollicite afin de retrouver la trace d'un écrivain à succès du nom de Sutter Cane. Celui-ci a disparu peu de temps avant de remettre à son éditeur le manuscrit de son dernier récit.

Le monde fictionnel de John Trent, que je noterai M_1, contient à la fois des éléments natifs, autrement dit, des éléments qui ne relèvent pas de la réalité, comme Suter Cane, et à la fois des noms qui, dans la réalité, dénotent des personnes réelles comme Stephen King auquel est comparé Cane. Mais dans le film, les processus d'importation ne sont pas limités à une frontière entre réalité et fiction. L'un des romans de Sutter Cane met en scène un village, Hobb's End, qui est donc, dans le monde M_1, purement fictionnel. Ce village est une fiction dans une fiction, il est 2-fictionnel, appartenant à un autre monde fictionnel, que je noterai M_2, accessible depuis le monde M_1 par un certain jeu de croyance feinte. Dans ce second monde fictionnel, où les romans de Sutter Cane sont racontés comme s'il s'agissait de faits réels, le nom « Hobb's End » réfère donc à un authentique village, alors qu'il n'a aucune sorte d'existence dans le monde M_1 où vivent Trent et Cane. Or, au gré de circonstances fantastiques, les personnages du film migrent d'un monde à l'autre. Les scènes de transgressions inter-mondaines sont nombreuses et diverses, j'en citerai quatre afin d'illustrer les articulations entre les différents niveaux fictionnels. La première scène

de transgression montre, au premier plan, Trent attablé avec un collègue à l'intérieur d'un bar. Au second plan, une vue longitudinale de la rue s'étend à travers les vitres. Un aliéné, armé d'une hache immense, et qui semble tout droit sorti d'un des romans de Cane, y produit une panique générale. L'événement reste ignoré de Trent et de son collègue, complètement absorbés par la conversation qu'ils entretiennent, jusqu'à ce que la vitre du bar, frontière symbolique entre les mondes M_1 et M_2, vole en éclats, sous les coups de hache du forcené, dont on apprendra par la suite qu'il est l'agent de Cane. Dans une autre scène, Trent part à la recherche de Cane. Malgré le fait que le village de Hobb's End n'est censé exister que dans les romans, Trent est persuadé qu'il s'agit bel et bien d'un authentique village, sans doute oublié des cartographes. Trent tire une localisation probable du village à partir d'éléments du récit de Cane, et quadrille le secteur au volant de sa voiture. Las d'une recherche aussi épuisante que vaine, Trent s'endort au volant. Evitant de peu un accident de la route, il se réveille, et se découvre alors parvenu à Hobb's End. Une troisième scène du film montre la rencontre de Trent avec Cane dans Hobb's End, auteur qui habite donc le monde de ses propres récits. Lors d'une scène mémorable, l'écrivain retourne son visage qui s'avère constitué des feuillets de son propre manuscrit. Enfin, dans une quatrième scène de transgression fictionnelle, Cane remet à l'assistante de Trent le manuscrit de son roman. Celui-ci contient la description des événements qui arrivent à Trent, événements qui sont réalisés au moment même où ils sont lus. Ainsi, le monde M_2 contient un élément qui décrit entièrement le monde M_1, autrement dit, ce que Thomas Pavel appelle le *Magna Opus* de la fiction (PAVEL, 1988). On s'attend d'ailleurs à voir se prolonger ce processus du monde M_1 en direction du monde actuel, et en effet, le réalisateur du film ne déçoit pas cette attente : le générique du film contient une annonce qui intègre le monde actuel dans son monde fictionnel. Le réel est donc comme absorbé par la fiction, elle-même engloutie par les récits de ses propres personnages. La mise en abîme est circulaire, décrivant une fiction dont on ne peut plus s'échapper, à l'image des routes de Hobb's End qui ramènent toujours à leur point de départ sans possibilité d'aller voir *ailleurs si on y est*.



Après avoir critiqué les conceptions allégorique et symbolique, nous avons proposé une autre catégorisation des figures du double, clivée entre, d'une part, les doubles natifs regroupant les indiscernables ostensifs et occultifs, et d'autre part les doubles importés dont relèvent les personnages récurrents et métaleptiques. Il reste maintenant à confronter ces différents cas de l'imaginaire du double aux conceptions théoriques. J'en viens donc à la question des paradoxes que suscitent ces cas pathologiques de récurrences pour les théories qui recherchent des critères métaphysiques d'individuation des personnages de fiction.

5.2 La question de l'individuation des personnages

Dans la section précédente, nous avons mis en exergue différentes situations de récurrences de personnages, selon qu'ils apparaissent indiscriminables ou selon qu'ils sont importés de manière métaleptique entre différents niveaux de fiction. Dans cette nouvelle section, nous allons examiner

les problèmes que ces situations engendrent pour les conceptions réalistes des personnages de fiction.

5.2.1 Penser deux fois à un même personnage ?

Dans une célèbre expérience de pensée, Peter Geach propose le scénario suivant : Hob pense que la mare de Bob a été infestée par une sorcière, et Nob se demande si c'est la même sorcière qui a tué la truie de Cob. Or, il n'est guère possible de traduire ceci par la proposition *de re* « qu'il y a une sorcière dont Hob pense qu'elle a infesté la mare de Nob et dont Hob pense qu'elle a tué la truie de Cob » sans manifester d'emblée un douteux engagement ontologique envers les sorcières. Mais l'on ne peut pas non plus se contenter d'affirmer *de dicto* que « Hob pense que la mare de Bob a été infestée par une sorcière et que Nob se demande si elle a tué la truie de Cob », parce que cette fois, une référence anaphorique traverserait les deux attitudes propositionnelles, sans que nous puissions expliquer cette identité. Nous avons là ce que Geach qualifie de problème d'identité intentionnelle :

> We have intentional identity when a number of people, or one person on different occasions, have attitudes with a common focus, whether or not there actually is something at that focus. (P. T. GEACH, 1967, p. 627)

Pourtant, dans nos discours fictionnels, nous identifions couramment des personnages. Nous disons par exemple de Sherlock Holmes qu'il est un personnage *distinct* du personnage d'Hercule Poirot, et nous disons de Batman qu'il est le *même* personnage que Bruce Wayne. Nous disons aussi que le personnage de Des Esseintes identifié jusqu'à la page 64 d'*À Rebours* est le même personnage que celui des pages suivantes. Mais puisque les noms de Poirot, d'Holmes, de Batman, de Wayne et de Des Esseintes ne réfèrent à rien dans le monde actuel, comment pouvons-nous confirmer ou infirmer de telles identifications ? Comment pouvons-nous avoir des pensées singulières à propos de choses inexistantes ? L'identification d'un personnage de fiction pose de fait principalement deux difficultés :

- Comment une même personne peut-elle penser au même personnage à deux reprises ?
- Comment deux personnes différentes peuvent-elles penser au même personnage ?

Les stratégies réalistes des personnages de fiction admettent la thèse que les personnages sont des entités, autrement dit, qu'ils sont des choses *individuées*. Dès lors, les noms fictionnels ne sont plus vus comme des noms vides ne renvoyant à rien de réel, ils sont vus comme des mots qui réfèrent véritablement à des objets, objets dont il s'agit d'abord de préciser la nature métaphysique, et ensuite de déterminer les conditions d'identité. Le réalisme fictionnel a bénéficié, dans les années 70, de l'impulsion de la théorie de la désignation directe. Comme l'a défendu Saul Kripke à la suite de Stuart Mill, les noms propres ne sont pas des abréviations pour des ensembles de descriptions définies, mais désignent directement leur référence dans le monde. Alors que la sémantique de Kripke se heurte aux noms vides qui pourraient apparemment valoir d'objection à l'encontre de cette théorie, de nombreux auteurs ont au contraire assumé une référence directe à des objets,

que je désignerai sous le nom générique de *ficta*, moyennant une souscription métaphysique, nous l'avons déjà vu, à des natures assez exotiques.

Les théories des *ficta* proposent ainsi de fournir les conditions d'identité des personnages plutôt que de se satisfaire d'une identification seulement intentionnelle. De leur point de vue, la différence entre des personnages de fiction ne relève pas simplement de la manière de les imaginer comme tels, elle n'est pas seulement une identification de concepts à travers des représentations diverses. Cette différence est objective, elle existe au cœur des objets afférents aux personnages. Comment ces théories prévoient-elles alors de discriminer des personnages lorsque nous sommes invités à les représenter comme indiscernables, quoique numériquement différents ? Si la pensée d'un personnage engendre un objet associé, alors la pensée d'un double engendre une entité ayant la nature métaphysique d'un double. Il est pourtant plus sage de considérer que, en admettant des doubles doxastiques, nous ne soyons pas contraint d'admettre aussi des doubles métaphysiques.

Il est généralement admis que le problème de Geach est ainsi aisément résolu depuis un point de vue réaliste admettant des entités fictionnelles. Si Hob et Nob pensent à la même sorcière, c'est que leurs représentations sont soutenues par un objet commun, un *fictum* identique. De même pour les personnages de fiction, qui, supportés par des *ficta*, échappent aux confinements de la pure subjectivité. Ce qui distingue les différentes stratégies réalistes, c'est leur manière propre de caractériser de tels objets. Nous avons vu au cours du chapitre 3 que ces stratégies se répartissent en trois grands groupes : celles qui introduisent des objets meinongiens inexistants, celles qui font appel à des objets simplement possibles, et celles qui font appel à des artefacts abstraits. Mais réifier les personnages en objets me semble impliquer plus de difficultés que de solutions aux puzzles d'identité. Car, en tant que *ficta*, les personnages de fiction n'échappent pas à des questions d'individuation parallèles à celles des objets concrets, et ces difficultés semblent même, dans le cas des *ficta*, exacerbées.

Dans ce qui suit, je vais considérer alternativement les trois stratégies réalistes et mettre en rapport les problèmes d'indiscernabilité et de doubles qu'elles occasionnent.

5.2.2 Le réalisme interne

Du point de vue du réalisme interne, les objets fictionnels sont complètement déterminés par les faisceaux des propriétés qui leur sont attribués par les auteurs dans les textes où ils apparaissent. Nous avons vu comment les théories néo-meinongiennes répondaient aux paradoxes mis en exergue par Bertrand Russell. Pour Terence Parsons, il s'agit de particulariser les propriétés fictionnelles instanciées par le *fictum* selon le couple d'oppositions propriétés nucléaires versus propriétés extra-nucléaires. Selon Edward Zalta il s'agit de particulariser la manière dont ces propriétés sont attribuées au *fictum*, selon le couple propriétés encodées versus propriétés exemplifiées. Les néo-meinongiens affirment ainsi être en mesure de faire correspondre à chaque personnage de fiction un *fictum*, qui, en tant que tel, l'individue.

Il reste cependant à savoir comment cette individuation peut nous être donnée, puisque contrairement aux particuliers concrets dont l'individuation procède, d'une manière ou d'une autre, de l'existence matérielle, ces nouvelles sortes de particuliers, quant à eux immatériels, nécessitent des conditions d'identité suffisamment déterminantes pour assurer leur stabilité référentielle.

Principe de compréhension et propriétés nucléaires

Nous avons vu que, selon le point de vue néo-meinongien, la relation entre les propriétés attribuées au personnage et le *fictum* correspondant répond au principe de compréhension, postulat qui explique l'association systématique entre un ensemble de propriétés et un objet. Or on peut lister différents principes de compréhension selon la force d'individuation qu'ils imposent aux faisceaux correspondants. On peut ainsi les classer de la manière suivante :

Postulat faible pour toute collection de propriétés il y a un objet qui possède toutes ces propriétés.

Postulat fort pour toute collection de propriétés il y a un objet qui possède exactement ces propriétés.

Postulat extrême pour toute collection de propriétés il y a un et un seul objet qui possède exactement ces propriétés.

Seule la dernière formule est garante de l'individuation d'un personnage. Elle a été développée par Terence Parsons qui établit une corrélation « un-à-un » entre les entités et les ensembles non vides de propriétés. L'idée de Parsons est que si nous établissons la liste des particuliers concrets, et en regard de cette liste, celle des ensembles de propriétés qui leur sont corrélés, nous pourrions encore prolonger la deuxième liste avec des ensembles de propriétés ne correspondant à aucune chose existante, comme par exemple l'ensemble {être une montagne, être en or}. L'idée de Parsons consiste alors à relier cet ensemble à un antécédent. Puisque, par construction, une telle entité n'existe pas, il ne s'agirait bien sûr pas d'un objet concret. Autrement dit, la manière dont un objet est « comme ceci ou comme cela » est indépendante de son être : son *sosein* est indépendant de son *sein*.

Cependant, en appliquant ce schéma aux choses fictives, on est d'emblée confronté au problème des propriétés qui sont externes à la fiction. Considérons la question suivante : la montagne d'or qui fait l'objet des pensées de Pierre est-elle la même que la montagne d'or qui fait l'objet des pensées de Jacques ? Rappelons que l'on cherche ici à définir de tels objets seulement par les propriétés qu'ils exemplifient. Or, si nous résumons ce que nous savons du premier objet, nous pouvons affirmer qu'il est une montagne, qu'il est en or, qu'il est pensé par Pierre et qu'il est inexistant. Quant au deuxième, nous pouvons dire qu'il est une montagne, qu'il est en or, qu'il est pensé par Jacques et qu'il est inexistant. Puisqu'ils ne possèdent pas exactement les mêmes propriétés, par le principe de diversité du dissemblable, on devrait affirmer l'existence de deux objets plutôt que d'un seul, ce qui n'est sans doute pas satisfaisant du point de vue du réaliste.

Parsons admet pourtant que les objets fictifs sont bien individués par les propriétés qui leur sont attribuées. De tels objets sont exclusivement identifiés par les propriétés qui sont internes à la fiction, c'est-à-dire celles que Parsons qualifie de nucléaires. Ainsi, la montagne d'or est-elle définie par l'ensemble de propriétés être une montagne, être en or, mais non par le fait d'être pensée par Jacques ou par Pierre. Tout objet fictionnel est donc la réunion de propriétés o_1, o_2, …, o_n *nucléaires*. Autrement dit, Parsons dote les objets fictionnels de la condition d'identité suivante :

Condition d'identité x est le même objet que y si et seulement si x et y partagent toutes leurs propriétés *nucléaires* respectives.

Appliqué aux personnages de fiction, ce principe d'individuation peut être énoncé ainsi :

Principe de compréhension nucléaire (PCN) à tout ensemble de propriétés nucléaires correspond un et un seul objet fictionnel.

Il arrive qu'un auteur attribue à son personnage des propriétés contradictoires, par oubli, par négligence, voire par intention. Citons la blessure de Watson que Conan Doyle situe successivement sur deux membres différents. Cela pourrait constituer un problème pour le réalisme interne. Mais Parsons affirme qu'il n'en est rien, car son postulat de compréhension n'implique pas de fermeture logique, en particulier, celle d'adjonction. Ainsi, un objet peut bien avoir la propriété d'« être bleu » et également celle d'« être carré », sans pour autant avoir la propriété d'« être bleu et carré ». De même, un objet peut également avoir la propriété d'« être un carré » et celle « de ne pas être un carré », sans avoir pour autant celle « d'être un carré et de ne pas être un carré ».

De son côté, pour éviter de distinguer différents types de propriétés, nous avons vu qu'Edward Zalta utilise un critère d'individuation différent de celui de Parsons, distinguant les manières dont les propriétés sont attribuées aux objets fictionnels plutôt que la nature des propriétés, de la manière suivante :

(i) Pour toute condition exprimable en terme de propriétés, il y a un objet abstrait qui encode ces propriétés et seulement celles-là.

(ii) Deux objets sont identiques si et seulement si ils exemplifient tous les deux l'existence et les mêmes propriétés ou s'ils sont tous deux abstraits et encodent les mêmes propriétés.

Le réalisme interne ramène ainsi l'identification des personnages à une l'identité ontologique de faisceaux de propriétés. Il reste à voir si une telle conception métaphysique de l'identité des personnages est capable ou non de rendre compte des doppelgängers et autres personnages doubles qui apparaissent dans certaines fictions.

Les conséquences du réalisme interne

Différentes conséquences résultent de ces manières d'individuer les personnages, conséquences qui sont d'autant d'objections à la pertinence de telles théories.

La question est d'abord de circonscrire les propriétés pertinentes qui permettent d'individuer un personnage. Cela revient ni plus ni moins à en déterminer une essence individuelle. Constatons alors que la lecture progressive d'une histoire modifierait en conséquence le contenu du faisceau du personnage. Mais il est généralement hors de question, pour un réaliste interne, d'admettre qu'un *fictum* soit une entité dynamique, aux propriétés fluctuantes, parce que cela rendrait difficile son identification. Ainsi, la construction du *fictum* demande d'avoir un point de vue global sur le personnage, il ne peut être constitué en cours de lecture, mais en tenant compte de l'ensemble des assertions de l'auteur à son sujet. Cela demande donc que cet ensemble soit arrêté.

Or il est toujours possible de produire une suite du récit (séquelle, préquelle ou interquelle) dans laquelle interviennent de nouvelles attributions. Plus encore, les réquelles, qui raconte une histoire de manière différente, impliquent des attributions contradictoires, autrement dit, des propriétés incompatibles au sein du faisceau. D'un côté, si nous agglomérons dans le faisceau un nombre trop élevé de propriétés, le *fictum* que nous obtiendrons sera alors trop restrictif et ne permettra pas d'identifier le personnage qui est censé lui correspondre à travers des histoires différentes dans lesquelles ces attributions seraient contredites. Par exemple, le *fictum* correspondant à *toutes* les propriétés attribuées à Sherlock Holmes par Conan Doyle dans *une étude en rouge* ne permettrait pas d'identifier Holmes dans sa parodie *Comment Watson apprit le turc*. D'un autre côté, si nous n'en retenons pas suffisamment, le *fictum* obtenu correspondra à divers personnages distincts. Par exemple, le *fictum* {homme, détective, particulièrement intelligent} pourrait aussi bien correspondre à Sherlock Holmes qu'à Hercule Poirot.

Considérons encore une autre objection, qui, elle, ne relève pas de conflits intertextuels, mais d'incompatibilités métatextuelles. Admettons que nous découvrions qu'un individu bien réel corresponde mot pour mot à l'ensemble des descriptions que Conan Doyle a donné de Sherlock Holmes. Bien que le *ficta* serait indiscriminable de l'individu réel, ce dernier ne pourrait pas être Holmes pour la raison qu'il est tout à fait essentiel à Holmes d'être un personnage de fiction, c'est-à-dire, d'être physiquement absent du monde. La fiction ne consiste pas à parler de choses factuelles, mais à raconter une histoire comme un fait connu d'un point de vue qui n'est pas la réalité. En ce sens, un vrai-Holmes ne pourrait satisfaire au critère pragmatique de fictionnalité.

Une autre conséquence de la théorie interne est que les *ficta* ne sont pas *créés*, ils sont de toute éternité. Cette caractérisation les rapproche d'objets idéaux comme sont caractérisées parfois les entités mathématiques. Or assimiler les personnages de fiction à de telles idéalités a pour effet considérer que leurs prédicats sont nécessaires, autrement dit, vrais dans tous les mondes possibles. Certes, dans un roman publié et diffusé, ce qui est dit est dit, et donc, est définitivement attribué aux personnages. Mais il reste que la compréhension du roman repose souvent sur le double psychologique, autrement dit, sur l'idée que les choses auraient pu se passer autrement. Comme nous l'avons déjà évoqué plus haut, nous ne pouvons comprendre l'*Étranger* de Camus qu'à la condition de croire que Meursault aurait pu ne pas commettre son crime. Comme l'écrit Umberto Eco, on pourrait toujours faire l'hypothèse de quelque velléité de réécriture contre-fictionnelle :

> Nous assistons à une représentation de l'*Œdipe roi* de Sophocle. Quel est alors notre désir ? Nous voulons désespérément qu'Œdipe prenne une autre route que celle où il rencontrera et tuera son père. Nous nous demandons pourquoi il est allé finir à Thèbes, et non, par exemple, à Athènes, où il aurait pu épousé une Phryné, une Aspasie. (…) En principe, nous pouvons faire en sorte que tout cela se produise : il nous suffit de réécrire *Œdipe roi* (…). Mais qui le désire vraiment ? (ECO, 2013, pp. 123–124)

L'intrigue d'un récit est rendue plus forte par le désir que les choses puissent se passer autrement. Cet espoir, bien que vain, est néanmoins ce qui fait tout l'enjeu de la tragédie. Bien que les spectateurs sachent, dès le début de la pièce, la fin qui attend inexorablement Antigone, le tragique de la pièce de Sophocle ne peut émerger qu'à la condition que des attitudes contre-fictionnelles des spectateurs soient mobilisées. Il ne s'agit pas ici de fatalisme, qui ne serait qu'une résignation douloureuse, mais d'une révolte héroïque face à une situation perdue d'avance. Le sentiment du

tragique survient sur l'idée harcelante que les choses *aurait pu* tourner autrement, voire s'arranger.

Passons à une autre objection : les entités associées aux personnages de fiction sont incomplètes. La question des propriétés ne se limite d'ailleurs pas à celles qui sont possédées, mais s'étend aussi à celles qui ne sont pas possédées. Les objets inexistants ne sont pas seulement par-delà l'être, ils sont aussi incomplets, ils sont troués (*property-gaps*). La montagne d'or est définie par les deux seules ascriptions d'être une montagne et d'être en or. Dès lors, la montagne d'or ne possède ni la propriété d'être située à l'est d'Eden, ni celle de ne pas être située à l'est d'Eden. C'est déjà l'objection de Quine à l'égard des objets simplement possibles. Celle-ci vise le caractère indiscernable des objets intentionnels, incapables d'être individués, et ne satisfaisant donc pas le critère quinéen résumé par le slogan « pas d'entité sans conditions d'identité » :

> Take, for instance, the possible fat man in that doorway ; and, again, the possible bald man in that doorway. Are they the same possible man, or two possible men ? How do we decide ? How many possible men are there in that doorway ? Are there more possible thin ones than fat ones ? How many of them are alike ? Or would their being alike make them one ? (…) is the concept of identity simply inapplicable to unactualized possibles ? But what sense can be found in talking of entities which cannot meaningfully be said to be identical with themselves and distinct from another ? These elements are well-nigh incorrigible. (W. v. O. QUINE, 1980, p. 4)

Parsons y répond en soutenant que le possible homme gros dans l'embrasure de la porte n'est tout simplement pas une chose unique, mais tout un ensemble de nombreuses choses différentes. Car si « être un homme » et « être gros » sont bien des propriétés nucléaires, en revanche « être possible » est une propriété extra-nucléaire. La question de Quine ne cible donc pas un objet particulier mais une classe infinie de choses différentes.

> How many possible men are there in the doorway ? At least as many as there are consistent sets of nuclear properties which contain the properties of being a man and of being in the doorway. There will be at least a countably infinite number of these. Are there more possible fat man than this one ? Probably not ; there are an infinite number of each but parity suggests that the cardinality will be the same for each. (PARSONS, 1980, p. 28-29)

Parsons introduit donc une distinction entre personnages de fiction et entités fictives. De fait, les personnages de fiction sont généralement plus amplement décrits que les entités fictives. On peut certes convenir avec Parsons qu'une pensée au sujet de Sherlock Holmes soit moins indéterminée qu'une pensée à propos de l'homme dans l'embrasure de la porte. Mais ce n'est pas toujours le cas. Si certains personnages sont épais, comme Des Esseintes, Madame Bovary et Sherlock Holmes, d'autres sont fins, parfois même extrêmement fins, pensons notamment au personnage de Godot, dont on ne sait rien sinon qu'il est attendu comme le précise le titre de la pièce de Samuel Beckett. Sans doute, Parsons pourrait répondre que Godot est moins une entité fictionnelle qu'une entité fictive, mais la frontière entre les deux types d'entités deviendrait extrêmement difficile à discriminer. C'est d'ailleurs plutôt les personnages épais qui sont les moins fréquents : lorsqu'Anna Karénine traverse la foule répandue sur le quai de gare, admettre que cette foule est composée de personnages, c'est admettre une pluralité d'objets fictionnels qui sont moins déter-

minés que le roi de France chauve de Russell ou que l'homme dans l'embrasure de la porte de Quine. Ainsi, Parsons ne répond pas au problème de l'incomplétude des *ficta*.

Soulevons un autre problème, lié au précédent, et qui provient de ce que les théories internes rabattent complètement l'indiscernabilité sur l'identité. Dans *Nonexistent objects*, Parsons (PARSONS, 1980, p. 191) a évalué l'application de sa théorie aux cas problématiques des indiscernables. Il reconnaît que les difficultés que ces derniers constituent menacent sa théorie, parce qu'elle semble incapable d'admettre autant d'objets fictionnels que semble nécessiter le juste nombre de personnages à l'intérieur d'un roman. Plus précisément, Parsons souligne un problème qui apparaît lorsqu'un récit mentionne un groupe de personnages sans autre indication sur la composition de ce groupe. Dans *Le Seigneur des Anneaux*, Tolkien écrit que sept anneaux sont destinés aux Seigneurs Nains. Les anneaux s'avèrent absolument indiscernables les uns des autres, et le lecteur ne sait pas s'ils doivent ou non être imaginés ainsi. Le lecteur est libre de déterminer si les anneaux sont, dans la fiction, indiscernables, ou si l'indiscernabilité est seulement due au caractère incomplet du texte. Mais parfois, le texte mentionne des personnages indiscriminables tout en incitant le lecteur à considérer qu'ils ne le sont pas. Une telle situation intervient par exemple lorsqu'une histoire mentionne une foule rassemblée sans que soit faite la moindre allusion aux membres qui la composent. En illustration, citons un passage du *Seigneur des Anneaux* racontant la bataille de la Morannon que Tolkien décrit ainsi :

> Des collines de part et d'autre de la Morannon se déversèrent d'innombrables Orques. Les hommes de l'Ouest étaient pris au piège ; bientôt, tout autour des monticules gris où ils se tenaient, des forces dix fois plus grandes, voire plus de dix fois supérieures les enfermeraient dans une mer d'ennemis. (TOLKIEN, 1992)

Remarquons qu'il n'y ici a guère moyen de déterminer les *ficta* permettant d'individuer les différents orques, hommes, ou monticules du récit. Les foules posent le problème des indiscriminables occultifs : nous n'avons aucun moyen des les discriminer, ceci par manque d'informations.

Plus précisément, l'application du principe de compréhension nécessite que chaque personnage soit associé à un unique ensemble de propriétés nucléaires, comme par exemple « être un orque » et « avoir participé à la bataille de Morannon ». L'identité est définie de manière extensionnelle : deux objets sont identiques si et seulement s'ils ont strictement les mêmes propriétés nucléaires, autrement dit, le principe de compréhension est appuyé sur le principe fort d'indiscernabilité des identiques. L'application du principe de compréhension à la foule des orques implique, par leur indiscernabilité, de corréler chacun des orques au même *fictum*, ce qui paraît plutôt absurde. Le fait est que les *ficta* générés par des personnages indiscernables sont inévitablement eux-mêmes des objets indiscernables, donc identiques. Dans la théorie de Parsons, une foule de personnages indifférenciés les uns des autres ne génère qu'un unique *fictum*. Parsons admet cette conséquence de sa théorie, mais il considère qu'elle ne constitue pas vraiment une objection. En effet, Parsons soutient qu'elle semble cohérente avec l'incapacité du lecteur d'individuer les membres d'une foule fictionnelle lorsque ceux-ci ne sont pas décrits. Parsons souligne même qu'il est tout à fait vain de tenter de discriminer les membres anonymes d'une foule fictionnelle en faisant par exemple appel, dans l'exemple de la bataille de Tolkien, à des descriptions définies du type « l'orque le plus avancé sur la ligne de front », ou du type « l'orque que j'ai actuellement en tête » car de telles descriptions n'ont aucune sorte objectivité. Selon Parsons, personne ne peut

avoir en tête un objet particulier s'il n'est pas en mesure de dire au juste en quoi cet objet diffère des autres (PARSONS, 1980, p. 191).

La manière dont Parsons répond finalement à cette difficulté est la suivante : les orques ne sont pas véritablement des personnages, plus précisément, on ne peut pas les désolidariser de la multitude qui les contient, qui, seule, constitue véritablement un objet fictionnel :

> If a story says "a crowd gathered" but says nothing at all individually about the members of the crowd, then there is no way to get the theory to yield any particular member of the crowd as a fictional object. I think that this in fact corresponds to our practice in discussing fictional objects... Other examples look initially less clear-cut, but I think they come to the same thing. Suppose we are reading a story which mentions a pair of twins ; the story says a few things about them, but it in no way distinguishes them from one another. In reading the story, you naturally expect to learn something to distinguish them, and you expect each of them to emerge as a character in its own right-but it never happens. Do we here have two fictional objects indistinguishable withing the theory ? I think not. (PARSONS, 1980, p. 157)

Dans le cas des foules aux indiscernables occultifs, lorsque nous croyons parler de plusieurs entités réunies, nous parlons en fait d'une seule entité indivisible.

On peut toutefois objecter que tout cela semble extrêmement contre-intuitif. La notion de foule sert à caractériser la réunion de plusieurs individus, et cet usage devrait demeurer le même dans un contexte de fiction. Il paraît certes assez juste de dire qu'un lecteur n'imagine pas chacun des membres de la foule des orques. Cette réponse semble cependant éluder la difficulté plutôt qu'elle ne la résout. S'il est certain que le récit ne permet pas au lecteur de discerner des personnages, il ne prescrit pas non plus d'imaginer que les orques sont indiscernables. Le lecteur peut ainsi prendre la liberté de considérer que les orques sont bien discernables, même si les moyens de les discriminer ne lui sont pas donnés. En outre, la discernabilité des orques est présupposée par le récit [14]. En particulier, Parsons défend un littéralisme fictionnaliste : un objet inexistant possède ses propriétés nucléaires de la même façon qu'un objet réel possède les siennes. Dès lors, si un objet x inexistant est une foule, il l'est de la même manière que l'est un regroupement de « gilets jaunes » un samedi après-midi sur l'Avenue des Champs-Élysées : c'est-à-dire qu'il est inévitablement constitué de membres discernables. Aussi, contrairement à ce que prétend Parsons, il semble que dans le cadre de sa théorie, la référence à une foule fictionnelle d'individus ne peut pas éviter la multitude de références à ses membres fictionnels. Le problème des indiscernables montre ainsi les limites d'une caractérisation interne des objets fictionnels, qui demeure impropre à rendre compte de la flexibilité des intentions fictionnelles.

En outre, nous ne savons pas exactement à combien d'objets fictionnels correspondent les orques de la foule, parce que nous n'en connaissons pas le nombre. Cela engage le théoricien des objets fictionnels à des foules bien étranges en ce que leur cardinal est variable. Précisément, si un récit de fiction fait mention d'une « foule constituée de 100 à 150 personnes », cette foule engage l'ami des objets fictionnels à stipuler l'existence variable de 100 à 150 de ces objets [15]. Il

[14] Ce fait est par exemple confirmé par la mise en scène cinématographique du *Seigneur des Anneaux* de Peter Jackson, où les orques sont représentés avec des personnalités et des traits différents

[15] Voir aussi (A. EVERETT, 2013, p. 229)

faut bien faire attention ici que l'intervalle ne décrit pas une indétermination épistémique, comme lorsqu'on dit qu'il y a entre 10 et 15 personnes à une réception. Car on peut toujours compter les personnes présentes pour en connaître le nombre exact. En effet, ce n'est pas le nombre d'individus qui est variable, mais les croyances que nous en avons. En revanche, c'est sans espoir pour des individus des foules fictionnelles qui contiennent véritablement un nombre variable d'objets fictionnels. Citer les membres d'un groupe par des noms distincts ne semble même pas résoudre cette difficulté. Considérons par exemple l'histoire courte suivante : « Il était une fois un agent secret nommé Ripley. Et il était une fois un agent secret nommé Greenleaf. Personne ne savait et n'a jamais su si Ripley était ou non Greanleaf. » Même si deux noms sont introduits dans le récit, il n'est cependant pas possible de savoir si on a affaire à un ou à deux personnages. Il est alors indéterminé qu'il y ait un ou deux objets fictionnels [16].

Dans son ouvrage, Parsons examine également un deuxième cas problématique pour sa théorie, cas qui est à nouveau lié à l'indiscernabilité. Il s'agit de la situation d'un récit qui mentionnerait un couple de jumeaux sans permettre au lecteur de discriminer entre les deux personnages. Ces personnages généreraient-ils deux *ficta* indiscernables ? Parsons rejette cette possibilité. Comme dans le cas de la foule rassemblée, il ne s'agirait que d'un seul individu : le couple de jumeaux lui-même. Seule la paire de jumeaux constituerait un objet fictionnel, et nullement les jumeaux pris chacun à part. Parsons appuie son affirmation en soulignant que nous ne devrions pas nous laisser abuser par des confusions référentielles. Dans ce genre de récit, dire qu'un auteur a créé *deux jumeaux* indiscernables est une façon trompeuse de faire référence au fait qu'un auteur a créé *un couple* de jumeaux. L'auteur n'aurait pas créé l'un des jumeaux puis l'autre, mais directement le couple. Il n'y a donc, conclut Parsons, pas moyen de faire référence à chacun des individus, mais seulement au couple. Si l'on attribue *explicitement* que la seule propriété d'être un frère

[16] Gareth Evans (EVANS, 1978) a prétendu avoir démontré l'impossibilité d'objets vagues à l'aide du raisonnement par l'absurde suivant. Admettons que les termes singuliers a et b soient tels que l'identité « $a = b$ » soit indéterminée. Traduisons par $\nabla(a = b)$ l'indétermination de cette identité. Admettre la vérité de :

$$(1)\ \nabla(a = b)$$

implique de reconnaître un fait à propos de b, fait qui peut être traduit par la propriété :

$$(2)\ \mathrm{P}(\nabla(x = a))b$$

Par ailleurs, la réflexivité de l'identité implique que :

$$(3)\ \neg\nabla(a = a)$$

et donc que :

$$(4)\ \neg\mathrm{P}(\nabla(x = a))a$$

Mais de l'indiscernabilité des identiques et de (2) et (4), on devrait alors dériver que :

$$\neg(a = b)$$

Pour achever la démonstration, Evans termine par une remarque assez obscure et qui a fait couler beaucoup d'encre pour critiquer son argumentation, voir par exemple (E. J. LOWE, 2011). Evans affirme que si l'opérateur « indéfinitivement » et son dual « définitivement » (« Δ ») génèrent une logique modale aussi robuste que S5, les propositions (1) à (4) ainsi que l'indiscernabilité des identiques devraient être renforcées par le préfixe « définitivement » et permettre de dériver que $\neg\nabla(a = b)$. Même si l'on ne souscrit pas à l'impossibilité des objets vagues, il demeure toutefois que ces objets sont très peu désirables quand il s'agit de trouver une explication à nos intentions fictionnelles. Pour une réponse à l'objection d'Everett, conférer SCHNIEDER et SOLODKOFF, 2009.

à un personnage, il possède néanmoins d'autres propriétés, celles-ci implicites, comme d'être né de certains parents, d'avoir certaines croyances, etc. Si on attribue exclusivement la qualité de jumeaux, dans une histoire, à deux personnages, ils devraient toutefois également posséder d'autres propriétés, comme des propriétés extrinsèques permettant de les localiser, quoique celles-ci ne soient pas mentionnées. Dès lors, les *ficta* correspondants devraient s'avérer numériquement différents l'un de l'autre, bien que nous ne puissions les déterminer de manière précise.

Parsons examine encore une troisième situation problématique, où il imagine une pièce de théâtre adaptée du *Hamlet* de Shakespeare, dans laquelle deux personnages nommés « Rosencrantz » et « Guildenstern » agiraient et parleraient à l'unisson et qui ne seraient distingués l'un de l'autre par aucun autre personnage. Un acteur pourrait cependant représenter un personnage, et un autre, le second, ce qui semble être une preuve de démarcation. Apparemment, les personnages devraient donc générer deux objets fictionnels plutôt qu'un seul. Parsons répond que la distinction entre les deux personnages provient de l'attribution de noms distincts, qui n'est pas toujours définie de manière interne à la fiction. En effet, « Rosencrantz » et « Guildenstern » pourraient être des manières externes de faire référence à deux objets fictionnels, qui dans l'histoire, pourraient s'avérer homonymes. Dans cette situation aussi, il n'y aurait finalement, selon Parsons, qu'un seul et unique objet de référence.

Globalement, les théories internes souscrivent, dans le cas des objets fictionnels, à la théorie descriptiviste de la référence. En effet, un personnage est entièrement défini par l'ensemble des propriétés qui lui sont associés dans les ouvrages qui les mentionnent. Le point de vue est cependant très idéalisé puisqu'il consiste à faire comme si la construction du personnage était toujours achevée. Un personnage est identique à un autre si les ensembles qui leur correspondent sont extensionnellement identiques. Lorsque l'on parle de personnages comme Sherlock Holmes, qui sont censés, à l'intérieur du jeux de croyance feinte, êtres des objets matériels, donc soumis à une évolution, les théories internes semblent défaillantes en ce qu'un ensemble dont on modifie l'une des propriétés n'est plus le même que le précédent. L'individuation par ensemble de propriétés semble donc devoir systématiquement échouer lorsqu'elle a affaire à des jeux de croyance feinte dont l'enjeu consiste à faire semblant de croire en l'existence d'individus matériels, que la narration présente manifestement de manières diverses et contradictoires à différents moment du temps. La référence à l'objet fictionnel ne peut pas être fixée par un ensemble de représentations. Les compilations de données que le lecteur associe au personnage ne sont donc pas assimilables à des ensembles figés, mais devraient relever de dossiers mentaux dynamiques, évoluant au fil de la lecture et de la construction du personnage par le lecteur.

Les focalisations internes et externes

Il est important de constater que le monde fictionnel qui est progressivement construit par le lecteur n'est pas en correspondance directe avec l'ensemble des propositions explicitées ou déduites du texte. Pour reconstruire un monde imaginaire cohérent, le lecteur doit faire un tri entre les caractérisations internes et externes prescrites par le narrateur. Par exemple, dans *Le Double*, le narrateur distingue Goliadkine aîné de Goliadkine cadet. Il ne s'agit pourtant pas de croire que l'un des personnages est plus âgé que l'autre, ce qui correspondrait à une propriété interne des personnages. Il s'agit ici d'une manière dont le narrateur use pour introduire une caractéristique narrative, l'un

des personnages étant introduit après l'autre. Dans la littérature consacrée aux *ficta*, c'est un fait assez peu remarqué que le clivage entre les caractérisations internes et externes ne correspond pas exactement à une opposition entre ce qui est dit du personnage *dans le texte* et ce qui est dit de lui *à l'extérieur* du texte. En effet, il faut prendre garde que le narrateur d'un récit peut mentionner, à côté d'attributions internes comme « être fonctionnaire », « habiter Saint-Petersbourg » des attributions « externes » comme des jugements sur les actes et croyances du personnage. A coté de la caractérisation interne/externe respectivement textuelle et para-textuelle, il y a encore une autre démarcation qui elle, est propre au texte, et qui dépend de la focalisation. Je reprends par ce terme la manière dont Gérard Genette qualifie le rôle du narrateur dans *Figures III* (GENETTE, 1972, p. 206), soit que celui-ci ait un point de vue omniscient (absence de focalisation), soit qu'il s'agisse d'une focalisation externe, c'est-à-dire que le narrateur dispose de connaissances excédentaires par rapport à celles du personnage, ou soit qu'il s'agisse d'une focalisation interne si les connaissances et croyances du narrateur coïncident avec celles du personnage. Bien souvent, la focalisation interne est manifestée par à un récit à la première personne du sujet, c'est-à-dire à l'identification entre le narrateur et le personnage. Dans ce cas, le récit est en adéquation avec les mondes possibles épistémiques et doxastiques auxquels a accès le personnage. Dans *Le Double* de Dostoïevski, le narrateur omniscient est progressivement mis en concurrence avec le flux des pensées du protagoniste. Il a souvent été remarqué, dans la littérature consacrée à cette œuvre, que cette identification procède d'une similarité croissante entre la manière dont le narrateur décrit les faits et la manière dont Goliadkine s'exprime. Progressivement, une focalisation interne émerge, avec cette originalité que l'auteur maintient l'emploi de la troisième personne du sujet pour désigner son personnage. L'effet qui résulte de ce repli de données objectives du roman est que les mondes possibles accessibles sont progressivement démultiplés. Partant d'un roman réaliste, le lecteur perd progressivement des connaissances au sujet du monde fictionnel du *Double*. Il ne sait plus exactement si le personnage est seulement fou, où s'il est effectivement la victime d'un double usurpateur. Le roman peut alors se lire de deux manières, soit qu'on admette l'existence objective d'un double, soit qu'on conçoive celui-ci seulement comme une illusion entretenue par le personnage.

Ce qu'indiquent ces remarques, c'est que ce n'est pas seulement les membres de foules anonymes qui peuvent faire l'objet d'un problème de dénombrement dans les récits, mais même le nombre de personnages épais, voire le nombre de protagonistes principaux. Pour que la lecture du roman de Dostoïevski conserve toute sa richesse sémantique, celle-ci nécessite cette comptabilité ambivalente. On voit mal comment un réaliste pourrait maintenir ce vague, puisque son objectif est justement de réifier le contenu du récit en objets bien individués.

Terminons en soulignant que le réalisme interne est incapable de rendre compte d'une différence entre le personnage de Cervantes et celui de Ménard. L'identité des textes impliquant l'identité des descriptions, les *ficta* supportant les deux personnages sont exactement les mêmes. Des personnages semblables, voire exactement semblables, ne sont pas encore des personnages identiques. Et d'ailleurs, la converse est encore vraie, des personnages dissemblables ne sont pas nécessairement des personnages distincts. Et comme le souligne Thomasson, « ce problème hante toute tentative pour identifier et distinguer des personnages uniquement sur la base des propriétés qu'ils possèdent ou encodent » (THOMASSON, 2011, p. 89).

5.2.3 Le réalisme possibiliste

Nous l'avons dit, les mondes fictionnels peuvent être vus comme des ensembles de mondes possibles où sont réalisés les événements décrits par les auteurs.

Une théorie des fictions, ancrée sur les mondes possibles et qualifiée par Lubomir Doležel de théorie mimétique, procède en caractérisant un personnage fictionnel comme la contrepartie d'un personnage réel. Doležel la formalise ainsi : selon cette théorie, un particulier fictionnel représente toujours un particulier actuel. Cette caractérisation a pour mérite de donner un sens à l'importation de personnages historiques dans les fictions, comme, mettons, Napoléon dans *Guerre et Paix* de Tolstoï ou comme la ville de Londres dans les récits de Doyle. Elle explique également l'origine de certains personnages légendaires comme prototypes de personnages de Romans comme par exemple le Roi Arthur dont le personnage serait calqué sur la personne de Rothiamus, roi des Bretons (ASHE, 1985). Cette caractérisation a néanmoins une limite évidente : à qui ou à quoi peut-on faire référer les noms de Raskolnikov, de Bilbo le Hobbit ou de Gregor Samsa ? Selon Erich Auerbach, plutôt que d'y voir des noms de particuliers, nous devrions les considérer comme des abréviations pour des ensembles traits universels, autrement dit, ce qui relève dans la fiction d'un particulier représente en réalité, dans le monde actuel, un universel.

Comme le souligne Lubomir Doležel, la littérature de fiction est un système sémiotique ayant pour finalité la « construction de mondes fictionnels » (DOLEŽEL, 1985, p. 464). Selon l'auteur, la notion de personnage est primitive, elle sert à élaborer, ensuite, celle de monde fictionnel. La création d'un monde fictionnel ressortit donc essentiellement à la création de personnages. Un monde fictionnel est un ensemble d'individus fictionnels *compossibles*, autrement dit, bâti sur un ensemble de personnages témoignant d'une « aptitude à la coexistence et à l'interaction » (DOLEŽEL, 1985, p. 464). Le personnage est vu comme d'emblée individué hors de l'histoire qui le mentionne. Sous ce point de vue, le monde fictionnel d'Emma Bovary est circonscrit par l'ensemble de tous les individus qui apparaissent dans le roman de Flaubert et rien qu'eux. Emma et Charles Bovary sont compossibles dans le monde fictionnel afférent au roman de Flaubert, alors qu'Emma Bovary et Ivan Karamazov ne le sont pas. Les mondes fictionnels sont ensuite conçus comme des ensembles de mondes possibles qui regroupent les bons arrangements de personnages.

Cette théorie peut paraître assez étonnante car il semble sans doute plus intuitif de construire les mondes fictionnels à partir des récits plutôt qu'en détachant les personnages des histoires qui les contiennent. L'avantage est néanmoins de rendre les mondes fictionnels indépendants des récits : ainsi, on peut expliquer les prolongements d'un récit par la récurrence des personnages. Il n'y a donc pas de problèmes de doubles interfictionnels, puisque cette théorie est d'emblée construite pour les éviter en soutenant l'identité de personnages qui reviennent dans des récits différents. Le problème d'individuation est alors rejeté sur les mondes fictionnels. Admettons en effet qu'un auteur prolonge les mondes fictionnels respectifs d'Emma Bovary et d'Ivan Kamarozov de telle manière que soit annoncée leur rencontre dans un nouveau récit. En individuant les mondes fictionnels par la compossibilité des personnages, et non par les histoires, on aboutit inévitablement au téléscopage de mondes fictionnels, ceux de Flaubert et de Dostoïevski. Cette possibilité, toujours ouverte à toutes les fictions qui peuvent témoigner des rencontres les moins attendues, a pour conséquence d'agglomérer la multitude des divers mondes fictionnels en un seul vaste monde.

La conception possibiliste des *ficta* consiste en l'hypothèse contraire qu'il y a des individus

qui existent seulement de manière inactuelle, autrement dit, qui sont des possibilités non-réalisées (*mere possibilia*). Il y a des choses qui n'existent pas actuellement mais qui existent néanmoins dans d'autres configurations possibles du monde. Sous cette perspective, les noms des personnages de fiction réfèrent à des individus inactuels mais possibles, et non à des universaux actuels. Cette conception a été défendue par Saul Kripke avant 1972. Appliquée aux noms fictionnels, une telle théorie affirme qu'il y a un objet simplement possible, dont on peut dire, dans un monde possible, qu'il instancie les propriétés qui correspondent aux descriptions que Conan Doyle attribue à Holmes.

Une telle théorie des mondes possibles semble taillée sur mesure pour répondre aux problématiques posées par l'identification des personnages de fiction. Elle induit cependant un coût ontologique important lié à la manière dont elle caractérise le domaine des mondes. Selon la théorie pan-réaliste de David Lewis, les mondes sont des sommes méréologiques maximales d'individus, et il y a une infinité de mondes possibles dans lesquels est réalisé tout ce qui est dit ici par Conan Doyle. Chaque individu est complètement ancré à son monde (*world bound*), c'est-à-dire qu'aucun individu n'existe dans plusieurs mondes à la fois. Mais, comme nous l'avons déjà fait remarquer, le réalisme modal paraît difficile à défendre.

Une autre conception, possibiliste, initiée par Richard Routley est qualifiée de meinongianime modal. Elle est défendue par Graham Priest (PRIEST, 2005) et par Francesco Berto (BERTO, 2013). Il s'agit d'une sorte de logique libre ancrée sur les mondes possibles. Selon leur point de vue, le domaine des mondes est le même pour tous, il n'est pas adapté de manière dynamique monde à monde, mais demeure constant. Les mondes possibles ont tous le même domaine, ce qui implique que certaines entités seulement possibles qui ne sont pas manifestées de manière physique dans le monde actuel « existent » tout de même, d'une manière inactuelle. Comme nous l'avons déjà indiqué dans le chapitre 3, cela consiste à désolidariser l'existence de l'actualité. En particulier, la thèse possibiliste définit la relation entre un *fictum* et le personnage qui y correspond de la manière suivante :

- Un *fictum* x correspond à un personnage y d'une fiction f par l'agent $\mathcal{A}$ si et seulement si x est un individu d'un monde possible M qui possède dans M toutes les propriétés attribuées à y dans f par $\mathcal{A}$.

Les logiques libres admettent deux types de constantes individuelles. Le premier type regroupe celles qui désignent un objet actuel, le deuxième type, celles qui n'en désignent pas. En excluant ce deuxième type d'objets du domaine de quantification, il devient alors nécessaire d'introduire, en plus du quantificateur existentiel habituel, un nouveau prédicat spécifiant l'existence. Ce prédicat, généralement noté « E ! », est ainsi exclusivement attribué aux objets existants : $\mathrm{E!}t := \exists x\ (x = t)$. Bien sûr, l'introduction d'un tel prédicat nécessite de relativiser la validité logique de la généralisation existentielle selon le cas que la constante individuelle désigne ou non un existant. De telles sortes de logiques obéissent donc aux deux principes suivants :

- $\mathrm{P}a \nvdash \exists x\ \mathrm{P}x$
- $\mathrm{P}a, \mathrm{E!}a \vdash \exists x\ \mathrm{P}x$

Une telle interprétation de l'existence échoue cependant à rendre compte de l'unicité d'un inexistant :

$$\mathrm{P}a \not\models \exists x\, \mathrm{P}x \wedge (\forall y\, \mathrm{P}y \rightarrow (y = x))$$

Aussi, pour assurer l'individuation des personnages de fiction, une solution consiste à compléter le domaine $\mathcal{E}$ des objets existants par un second domaine, noté $\mathcal{I}$ et contenant des objets inexistants, abstraits, simplement possibles, voire impossibles. On définit alors, en plus des quantificateurs habituels « $\forall$ » et « $\exists$ » qui prennent leurs valeurs exclusivement dans l'ensemble $\mathcal{E}$, une nouvelle paire de quantificateurs notés respectivement « Σ » et « Π », portant sur la réunion $\mathcal{E} \cup \mathcal{I}$. Cette stratégie a pour effet d'interpréter formellement une phrase comme « il y a (au moins) une chose qui n'existe pas » de la manière suivante :

$$\Sigma x\, \neg\exists y\, (y = x)$$

Le quantificateur « $\exists$ » est ainsi dédoublé en un quantificateur existentiel classique auquel est adjoint un quantificateur d'« être » quant à lui indépendant de l'existence. Il s'agit donc d'une manière anti-quinéenne de considérer l'existence, puisque celle-ci induit qu'exister, ce n'est pas être simplement la valeur d'une variable associée.

Pour élaborer une sémantique des noms vides tout en renonçant à cette manière, somme toute assez étrange, de distinguer les locutions « il existe » et « il y a », le nonéisme de Graham Priest et le meinongianisme modal de Francesco Berto proposent une utilisation de la logique libre appuyée sur les mondes possibles. Je ne développerai ici que la version nonéiste, celle de Berto étant très proche.

Pour sa part, Priest élargit ainsi le domaine de quantification des constantes individuelles en admettant qu'il contient d'une part des objets actuels et d'autres part des objets inactuels, c'est-à-dire simplement possibles. Autrement dit, il associe les objets inexistants de Meinong à des objets simplement possibles. A cette fin, il introduit deux nouveaux quantificateurs, le premier noté « $\mathfrak{G}x$ » dont l'interprétation est « pour quelque (*some*) x », de deuxième noté « $\mathfrak{A}x$ », servant de quantificateur universel au même titre que « $\forall$ ».

Cette stratégie évite de dissocier l'existence de l'être. Car si la lecture que donne Priest de l'énonciation $\mathfrak{G}a\ \mathrm{P}a$ n'est pas « il existe un a tel que $\mathrm{P}a$ », cette formulation n'est pas non plus donnée à lire comme « il y a un a tel que $\mathrm{P}a$ » (PRIEST, 2005, p. 13). Il s'agit plutôt d'interpréter cette dernière formule d'une manière ontologiquement neutre : « quelque (*some*) a est tel que $\mathrm{P}a$ ». L'existence, considérée non plus comme une quantification mais comme une propriété, est ensuite assurée par le prédicat monadique « E ! ». Pour prendre un exemple, la proposition « il existe un x tel que $\mathrm{P}x$ » est ainsi interprétée de manière formelle par l'expression « $\mathfrak{G}x\ (\mathrm{E}!x \wedge \mathrm{P}x)$ ». L'idée générale de la sémantique de Priest est que l'inexistence des objets auxquels réfèrent les noms vides relève de l'inactualité, autrement dit de leur absence dans ce monde-ci et de leur réalisation dans ce monde-là. La distinction entre exister et être à laquelle souscrivait Russell avant 1905, qui instituait pour les objets fictifs une sorte d'existence ontologiquement inférieure à celle des objets physiques est donc remplacée par une distinction ontologiquement neutre entre l'actuel et l'inactuel.

Constatons que dans cette théorie, le domaine des mondes est exclusif et constant. Ce qui varie, ce sont les propriétés attribuées aux objets entre différents mondes, dont celle d'existence,

qu'un objet peut ou non instancier. Mais ce point de vue est souvent contesté. Peter von Inwagen (P. V. INWAGEN, 1986) a souligné que si les *ficta* n'existent pas de manière actuelle, cela implique des domaines qui seraient à la fois identiques et variables entre mondes possibles différents, ce qui est absurde. Les mondes ne sont pas des conteneurs d'objets. Cette conception est motivée chez Priest par le fait que les agents cognitifs se représentent leur monde d'une certaine manière, qui peut être une manière assez précise de représenter le monde actuel, mais aussi une manière précise de représenter d'autres mondes, comme par exemple les mondes d'une fiction donnée :

> [...] if I imagine Sherlock Holmes I represent the situation much as Victorian London (so, in particular, for example, there are no aeroplanes) ; but where there is a detective that lives in Baker St, and so on. The way I represent the world to be is not an accurate representation of our world, but our world *could* have been like that ; there is a world that is like that. (PRIEST, 2005, p. 84)

Un avantage considérable de cette conception par rapport à celle des néo-meinongiens classiques est d'éviter le caractère très peu intuitif des objets *ausserseiende*. De fait, dans la conception nonéiste, les objets associés aux personnages de fiction ne sont pas *heimatloss*, ils ne sont pas des objets « sans domicile fixe ». Au contraire, ils habitent des mondes fictionnels, à savoir, les mondes imaginés par leur auteur et par les récepteurs des œuvres qui les mentionnent. En ce sens, les *ficta* ne sont pas des choses inexistantes et actuelles. Il s'agit au contraire d'entités qui existent de manière inactuelle, autrement dit, qui existent dans des mondes possibles bien qu'elles n'existent pas dans le monde actuel. La question est cependant de savoir s'il y a véritablement une réduction du coût ontologique par rapport au théories néo-meinongiennes classiques.

Quoi qu'il en soit, poursuivons notre analyse de la conception nonéiste en observant maintenant la manière dont des propriétés sont attribuées aux *ficta*. Rappelons d'abord la forme naïve du principe de caractérisation meinongien :

- Pour toute propriété ou conjonction de propriétés $\mathrm{A}(x)$, il existe un objet c_{A} tel que $\mathrm{A}(c_{\mathrm{A}})$.

Comme nous l'avons déjà évoqué dans le chapitre 3, énoncé de telle manière à propos de choses qui n'existent pas, le principe débouche inévitablement sur l'explosion logique. Il est donc nécessaire soit de le restreindre aux seules propriétés nucléaires, soit de distinguer deux manières d'attribuer des propriétés, encodage et exemplification. Ces deux stratégies sont interdites à Priest, puisqu'elles impliqueraient inévitablement l'actualité de l'objet fictionnel c_{A} et Priest y renonce. Ainsi, celui-ci est amené à donner un principe de compréhension alternatif à ceux qui sont utilisés par les néo-meinongiens, de la manière suivante :

- Pour toute condition $\mathrm{A}(x)$ et tout agent a ayant une intention à l'égard d'un objet de pensée c_{A} caractérisé par $\mathrm{A}(x)$ selon l'intention correspondante Φ (croire, imaginer, concevoir, etc.),

  $$@ \Vdash^{+} a\ \Phi\mathrm{A}(c_{\mathrm{A}})$$

  où $@ \Vdash^{+} P$ est une interprétation signifiant qu'il est le cas que P dans le monde possible $@$. Autrement dit, l'objet de pensée c_{A} satisfait la condition $\mathrm{A}(x)$ dans tous les mondes w accessibles qui réalisent ce que l'agent Φ (croit, imagine, craint, etc.) dans le monde actuel $@$ (PRIEST, 2005, p. 85).

Notons que, puisque la représentation d'un monde fictionnel donné est par nature incomplète, celle-ci relève toujours, non d'un unique monde, mais de toute une pluralité de mondes possibles. En outre, Priest admet, en plus des mondes possibles, des mondes impossibles. Les mondes possibles et impossibles sont utilisés afin de résoudre non seulement le problème de l'attribution d'existence à toute chose (si Sherlock Holmes n'existe pas dans ce monde-ci, il existe bien dans cet autre monde-là) mais aussi de loger des objets contradictoires.

L'avantage de la théorie possibiliste est qu'elle donne une meilleure explication que la théorie des objets inexistants à la relation entretenue entre un *fictum* et les propriétés qui lui sont attribuées de manière interne. Lorsqu'on affirme que « Sherlock Holmes est détective », le nom propre de cette expression ne réfère plus à une entité actuelle et abstraite dont il serait bien difficile de penser qu'elle puisse exercer la profession de détective en tant que telle. Si quelqu'un est détective, alors il semble qu'il existe. Comme le souligne Priest, les conséquences logiques d'attributions de qualités spatio-temporelles à des objets inexistants ne peut mener qu'à l'explosion logique. En effet, elles débouchent inexorablement sur des conjonctions incohérentes comme « Holmes est détective (dans les histoires) et Holmes n'est pas un détective (dans la réalité) ». Dans la théorie nonéiste, le fait que le *fictum* dénoté par le nom « Sherlock Holmes » est détective n'exige pas un emploi à contre-courant de cette propriété puisque le nom réfère à un homme qui existe dans un monde seulement possible. La théorie possibiliste ne nécessite donc pas de recourir à un différenciation dans la manière d'attribuer des propriétés aux personnages de fiction, comme le fait la théorie des objets inexistants par la distinction entre exemplification et encodage. Autrement dit, du point de vue possibiliste, la manière interne d'attribuer des propriétés aux personnages de fiction est compatible et prolonge la manière d'attribuer des propriétés aux entités existantes comme les bateaux, les chats et les personnes.

Pour résumer, les principales demandes du meinongianisme modal sont donc 1) de désengager le quantificateur existentiel à la manière des logiques libres en considérant l'existence comme une propriété des objets, 2) de considérer un domaine d'objets commun à tous les mondes possibles et regroupant des objets actuels et simplement possibles, et 3) d'admettre des mondes impossibles. Voyons maintenant les problèmes d'individuation qui peuvent être soulevés dans une telle théorie.

Considérons d'abord le problème métaphysique de l'identité transmondaine. Un objet existe-t-il dans deux mondes différents ? Cela dépend bien sûr de la conception que nous avons des mondes possibles. Nous avons vu que la théorie de Lewis excluait cette possibilité : les individus sont ancrés à leur monde et n'existent pas ailleurs. Si l'on se réfère à la conception kripkéenne basée sur la structure <$\mathcal{W}$, $\mathcal{R}$, f, v> où $\mathcal{W}$ est l'ensemble de tous les mondes possibles, $\mathcal{R}$ la relation d'accessibilité usuelle entre mondes possibles, f la fonction de domaine et v une interprétation, le problème de l'identité transmondaine peut se traduire ainsi :

- Un individu i est tel que, pour deux mondes possibles distincts w_1 et w_2, $i \in f(w_1)$ et $i \in f(w_2)$.

Or, selon la conception nonéiste, ce ne sont pas les objets qui « existent » ça et là au gré des mondes visités, mais les propriétés qui sont attribuées à ces objets. Ce qui est dans un monde et non dans un autre, ce n'est donc pas un objet, mais la propriété qui lui est associée (dont celle d'existence). Le problème métaphysique de l'identité transmondaine ne semble donc pas intervenir dans la conception possibiliste.

Il demeure toutefois certains problèmes d'identité, par exemple celui d'une identification transmondaine entre les personnages et les agents cognitifs. Un lecteur pourrait s'imaginer posséder les mêmes qualités que celles que Conan Doyle attribue à son personnage dans ses histoires. En considérant un monde possible dans lequel de telles histoires sont réalisées, il faudrait alors admettre que ce lecteur est Sherlock Holmes dans ce monde là. Pour éviter cette identification contestable, Priest souscrit à la thèse haeccéististe : « The identity of an object is not determined by its properties at any one world » (PRIEST, 2005, p. 90). Ainsi, un monde w_1 peut représenter un individu i_1 caractérisé par un ensemble de propriétés $A(x)$ alors qu'un autre monde w_2 peut représenter un individu i_2 distinct de i_1 par exactement le même ensemble de propriétés $A(x)$. Puisque, contrairement à Holmes, le lecteur qui s'identifie à lui existe bel et bien dans le monde actuel, cela suffit pour les distinguer dans le monde actuel (par application de la contraposée de l'indiscernabilité des identiques). Le personnage auquel s'identifie le lecteur est donc un autre individu que Holmes, bien que ces objets soient indiscernables, au sens de doubles purs, dans les mondes en question.

Selon Priest, le critère d'individuation des objets est le suivant :

- Pour deux entités a et b, $a = b$ si et seulement si a et b ont les mêmes propriétés dans tous les mondes possibles.

Constatons par contre qu'avec l'emploi des mondes possibles, les récurrences métaleptiques trouvent, dans l'explication possibiliste, une explication par la différenciation des mondes possibles.

En revanche, on peut se demander si cette condition nécessaire et suffisante est valide dans le cas de personnages doubles. On peut certes affirmer qu'il en va ici comme il en va dans le cas d'objets actuels qui seraient des répliques pures, mais que l'on pourrait toujours différencier par localisation.

Le problème, dans le cas de la fiction, est toutefois plus complexe encore, et mène à concevoir de manière générale les *ficta* comme des doubles purs. Comme l'a bien montré Francesco Berto (BERTO, 2014, p. 83–84), le critère d'identité possibiliste des *ficta* implique l'existence d'indiscernables purement extensionnels, c'est-à-dire d'objets distincts quoique indiscernables de par leurs propriétés actuelles. Une conséquence de cette théorie est en effet qu'un objet simplement possible a et un objet simplement possible b peuvent varier intentionnellement et intensionnellement sans être *actuellement* discernables. a peut être détective dans le monde w_1 sans que b existe dans ce monde-là. Il peut être pensé par un agent dans w_1 sans que b le soit. De manière extensionnelle, on a affaire à des doubles purs, c'est-à-dire à des objets qu'il n'est pas possible de discerner par la moindre propriété intrinsèque ou extrinsèque dans le monde actuel. Or le réaliste conséquent, qui individue les *ficta* à l'aide de propriétés internes, autrement dit, par le principe d'indiscernabilité des identiques, ne peut accepter l'existence de doubles purs. Il prétendra donc distinguer de telles entités seulement par leurs propriétés modales. Mais comment expliquer une discernabilité purement modale entre deux objets, c'est-à-dire, une discernabilité qui ne soit pas fondée sur une divergence de propriétés actuelles ? Il semble au contraire que toute propriété seulement possible doive être fondée sur une propriété actuelle. En tout cas, il semble difficile de rejeter un tel principe.

5.2.4 Le réalisme externe

Il reste à considérer les problèmes de doubles qui apparaissent dans les théories alternatives, externes, autrement dit, dans les stratégies qui visent à individuer les *ficta* non pas par des descriptions auctoriales, mais par des actes de création.

L'individuation des artefacts abstraits

Amie Thomasson constate que si nous acceptons l'idée qu'un roman existe par un acte de création d'un auteur, il est difficile de ne pas accepter la même chose en ce qui concerne l'existence d'un personnage de fiction :

> If one keeps works of literature, then one gains nothing but false parsimony by 'getting rid of' fictional objects, for they are in just the same categories... (THOMASSON, 1998, p. 143)

Thomasson rapproche ainsi les conditions d'individuation des personnages de celles des œuvres, en stipulant que personnages et œuvres de fiction sont des objets de même nature.

Comme l'écrit Todorov,

> Une lecture naïve des livres de fiction confond personnages et personnes vivantes. On a même pu écrire des « biographies » de personnages, explorant jusqu'aux parties de leur vie absentes du livre (« Que faisait Hamlet pendant ses années d'études ? »). On oublie alors que le problème du personnage est avant tout linguistique, qu'il n'existe pas en dehors des mots, qu'il est un « être de papier ». (DUCROT & TODOROV, 1972, p. 286)

Or il convient de distinguer deux formes de représentations du personnage de fiction, celle interne, qui concerne, dans le cas de Sherlock Holmes, un vrai détective, et celle, externe, qui concerne un artefact culturel créé par Arthur Conan Doyle.

La théorie des artefacts abstraits prétend rendre compte de la manière dont est fixée la référence aux personnages de fiction, en particulier, elle prétend donner une explication au fait qu'un même personnage puisse apparaître dans différentes fictions (identité transfictionnelle) et une explication aux énonciations métafictionnelles. La théorie d'Amie Thomasson, nous l'avons dit, est fortement ancrée sur la sémantique kripkéenne et sur la théorie de l'intentionnalité de Roman Ingarden. Comme Kripke, elle considère que les personnages sont présentés au monde par une forme de baptême associé à l'acte inchoatif de création d'un auteur. Le nom d'un personnage est le nom du *fictum* créé par un certain auteur dans le premier texte où ce nom apparaît. De ce fait, le nom d'un personnage peut être causalement rattaché aux textes qui le mentionnent et à l'auteur qui l'a créé. Thomasson caractérise cette forme de dépendance par deux modes déclinés selon leur rapport aux entités réelles ou au mental, la dépendance générique à l'égard d'au moins un exemplaire des récits où apparaît le personnage et des pensées des lecteurs compétents, et la dépendance historique à l'égard de l'auteur et de son acte de création. Thomasson distingue par ailleurs des dépendances existentielles rigides et constantes. D'une part, Hercule Poirot entretient une relation de dépendance historique à l'égard de sa créatrice : il apparaît en raison d'un acte créatif surgissant à un moment donné. Sans cet acte, il ne serait pas. Aussi, la dépendance historique ne nécessite-t-elle qu'une forme de dépendance existentielle momentanée. Faut-il alors originer nécessairement

l'objet fictionnel dans tel acte de création plutôt que tel autre ? C'est l'avis de Thomasson qui affirme que la dépendance est rigide : Poirot ne pourrait pas avoir été créé par un autre auteur. S'il existait un monde dans lequel Agatha Christie ne serait pas l'auteur de Poirot, et si dans ce même monde, un auteur différent avait créé un personnage exactement ressemblant, un tel personnage ne serait néanmoins pas Hercule Poirot. Nous aurions deux personnages indiscernables mais néanmoins numériquement différents car ils ne répondraient pas aux mêmes intentions auctoriales. La théorie artefactuelle répond donc bien au problème de Borges, voire à des situations où les fictions respectives des deux auteurs seraient encore plus similaires (par exemple si Cervantes et Ménard avait été des voisins ayant écrit les mêmes textes sans connaissance du travail de l'autre). Par ailleurs, la subsistance d'un personnage en tant qu'entité métaphysique dépend également, de façon constante, de l'existence des textes et des attitudes mentales des lecteurs. Sans eux, le personnage en question disparaîtrait. Remarquons que cette dépendance générique à l'égard des textes et des pensées des lecteurs n'est pas rigide, puisque Poirot ne dépend évidemment pas de tel ou tel exemplaire du texte.

Le mérite principal de la théorie des artefacts abstraits est de répondre à l'intuition que les personnages sont des créations d'auteurs, et d'expliquer l'identité transfictionnelle. De façon intratextuelle, les personnages demeurent identifiés par des descriptions, mais la théorie artefactuelle propose une individuation intertextuelle basée sur les critères externes historiques de création des auteurs. Ainsi, un auteur qui voudrait faire référence au personnage de Conan Doyle, quitte à modifier certaines descriptions du personnage, pourrait le faire par l'usage du nom propre référant à l'acte inchoatif de création du personnage par Doyle. Globalement, la stratégie de Thomasson consiste donc à « attraper » le *fictum* non pas de manière interne, en l'individuant par les propriétés attribuées à l'intérieur de l'œuvre comme le font les néo-meinongiens, mais de manière externe par les actes des auteurs.

Ceci dit, nous n'avons pas encore expliqué la manière dont cette théorie associe les descriptions des auteurs aux *ficta*. Comment expliquer la vérité intuitive d'une expression fictionnelle comme « Sherlock Holmes est un détective ». Un artefact abstrait ne peut être un détective. Les détectives sont des personnes concrètes, pas des objets abstraits. Les artefacts abstraits sont pourtant décrits comme s'ils étaient des choses concrètes. Pour expliquer ce fait, Thomasson renonce aux propriétés nucléaires ou encodées des néo-meinongiens. Elle utilise la théorie de la simulation (*pretense*) : c'est dans la fiction de Conan Doyle que Sherlock Holmes est un détective, bien que cela ne soit pas littéralement vrai. L'avantage est que de ce point de vue, les *ficta* sont entièrement préservés de la contradiction, puisque l'attribution de propriétés contradictoires n'affecte pas l'objet fictionnel en lui-même, elles ne sont que des descriptions admises dans un jeu de croyance feinte.

Contrairement aux néo-meinongiens, chez Thomasson, les *ficta* ne sont pas *associés* aux personnages, la relation est ici l'identité : les *ficta sont* les personnages. Thomasson fournit une double condition d'individuation des *ficta*, interfictionnelle et intrafictionnelle. La première est une condition nécessaire mais pas suffisante, elle affirme que deux noms réfèrent au même personnage seulement s'ils co-réfèrent au même acte inchoatif de création d'un auteur. La seconde condition est une condition suffisante d'identité intrafictionnelle des personnages de fiction. Pour que le *fictum* x soit identique au *fictum* y dans la fiction f, il suffit que :

- le personnage x et le personnage y apparaissent dans la même fiction f ;
- dans cette fiction f, le personnage x a exactement les mêmes propriétés que le personnage y.

Notons que, contrairement aux néo-meinongiens, la condition d'indiscernabilité n'est pas rendue nécessaire pour assurer l'individuation interne des personnages. La théorie des artefacts abstraits s'aménage ainsi une marge de manœuvre pour expliquer l'identification des personnages lorsque leurs propriétés changent en raison de leur évolution à l'intérieur ou à l'extérieur de la fiction. Le *fictum* qui correspond à tel personnage peut avoir la propriété P au début du roman, puis ¬P ailleurs. Que l'indiscernabilité ne fournissent pas une condition complète d'individuation n'est cependant pas un problème pour Thomasson, parce que de toute façon, le *fictum* est d'emblée individué de manière externe par l'acte de création de son auteur. La tension entre l'imitation et la représentation disparaît donc.

La théorie de Thomasson ne peut se contenter de conditions d'identité intrafictionnelles. Il lui faut encore préciser des conditions d'identité interfictionnelles. Afin d'éluder le problème de l'instabilité transfictionnelle des propriétés internes, Thomasson préfère identifier historiquement les objets fictionnels en injectant l'intentionnalité à l'intérieur même des critères d'identification :

- L'auteur de la fiction f_2 doit connaître au moins d'une manière adéquate le personnage x de la fiction f_1 et avoir l'intention d'incorporer x dans la fiction f_2 en tant que y (THOMASSON, 2011, p. 96).

On voit que cette condition fait intervenir deux caractéristiques. La première est la connaissance du personnage auquel souhaite référer un auteur. Il s'agit d'une précaution qui permet d'éviter les cas où un écrivain aurait l'intention de référer à un personnage (disons Holmes) mais échouerait par méconnaissance du personnage ou par une quelconque erreur de référence (par exemple, il référerait en fait à Poirot).

La deuxième caractéristique présente dans la condition d'individuation interfictionnelle est l'intention d'incorporation. Il faut reconnaître qu'il est nécessaire d'injecter l'intention au sein des conditions d'identification. Cette intention peut d'ailleurs ressortir de manière manifeste, par exemple quand des indications paratextuelles sont ajoutées afin de rendre ostensible cette identification. L'utilisation du nom « Hercule Poirot » permet ainsi à la romancière Sophie Hannah de rendre manifeste son intention d'identifier son propre personnage à celui d'Agatha Christie en l'associant causalement à la chaîne de transmission du nom dont l'origine est l'acte inchoatif de création d'Agatha Christie.

De ce qui précède, il apparaît que la théorie artefactuelle de Thomasson est soucieuse de conférer un support ontologique aux contenus fictionnels afin de les doter d'objets de pensée. Car selon Thomasson, le contenu de pensée ne suffit pas à individuer les personnages. En accord avec la thèse de la référence directe, la thèse artefactuelle fournit une stratégie pour que les chaînes de communication des noms fictionnels débouchent sur des objets abstraits : à savoir, les objets qui représentent l'acte de création d'un auteur, acte manifesté à un certain endroit d'un texte, à une certaine date et à un certain lieu. Comme nous l'avons vu, l'avantage de cette stratégie, c'est

que les personnages ne sont pas individués par des propriétés internes. Leurs attributions peuvent conséquemment varier entre récits transfictionnels sans dommages pour leur identification. Certes, Thomasson prend bien garde de préciser que sa condition n'est que nécessaire. Mais elle affirme que ce qui empêche cette condition d'être suffisante, c'est uniquement la possibilité de changements involontaires de référence. En effet, même si un auteur utilise le nom d'un personnage, rien n'exclut la possibilité d'un changement suite auquel il ne réfère plus vraiment au personnage original, mais relève d'une nouvelle création, peut-être même à l'insu de l'auteur. C'est ici que Thomasson use de la notion de survie élaborée par Derek Parfit : elle admet que dans certains cas, nous sommes contraints à conclure en une identité vague (THOMASSON, 2011, p. 98).

Le rôle de la réception esthétique

Cependant, il semble que la théorie artefactuelle occasionne tout de même un sérieux problème provenant de l'ouverture interprétative des œuvres de fiction. On peut se demander ce que signifie « connaître d'une manière adéquate le x de f_1 et avoir l'intention d'incorporer x dans f_2 en tant que y ». Thomasson affirme qu'il s'agit là de la capacité d'utilisation compétente du nom de x. De fait, un auteur ayant rédigé un récit transfictionnel fut d'abord lecteur de la fiction dont il s'inspira. Autrement dit, il est un usager expert du nom de x. Mais qu'est-ce que cela signifie, quand il s'agit d'un non-existant ? Peut-on être un usager objectif du nom d'un personnage ? Les fictions laissent une part d'interprétation. Aussi, il semble que la condition donnée par Thomasson soit elle-même tributaire des conditions de co-identification de x entre les lecteurs de l'œuvre. Les conditions d'identité intrafictionnelle devraient prendre en compte ces variations possibles. Ce qui n'est pas le cas chez Thomasson. Néanmoins, il semble que décider de ranger une œuvre dans une série dépende d'un consensus critique issu d'une réception esthétique. Le critère d'intentionnalité qui est attaché de manière prévalente à l'auteur chez Thomasson doit être complété par les attitudes des lecteurs.

Il n'est finalement demandé au lecteur que des capacités linguistiques et épistémiques, savoir lire, comprendre de quoi il retourne, connaître le contexte dont s'inspire la fiction, etc. Pourtant, si un article critique se demande si le dernier Hercule Poirot est vraiment un Poirot, c'est que la volonté de l'auteur ne suffit pas à la réussite de l'entreprise transfictionnelle. Thomasson évacue des conditions d'identité la valeur décisionnelle critique provenant de la communauté des lecteurs. Or la question de la réception esthétique d'un texte ne peut avoir de sens qu'en faisant intervenir le jugement du lecteur. Comment s'assurer que les lecteurs d'une même fiction co-identifient bien le même objet ? La théorie artefactuelle ne répond pas à cette question. Car, en l'objectivant, elle appuie l'identification hors de la tête des lecteurs. Il se peut alors que ces lecteurs aient des contenus de pensée différents en dépit de la visée d'un même objet. Thomasson est d'ailleurs très soucieuse de ce problème puisqu'elle ne cesse de rappeler qu'objet et contenu de pensée ne se correspondent pas un à un. Outre qu'il arrive que deux descriptions définies différentes soient corrélées de manière pertinente au même référent, comme « le père de Télémaque » et « l'époux de Pénélope », il arrive aussi, inversement, que des descriptions identiques concernent des personnages différents. Dès lors, comment avoir la certitude que des contenus de pensée identiques font référence au même objet fictionnel ? On voit que l'intention auctoriale n'est évidemment pas un critère suffisant à la co-identification. Se prononcer en faveur d'objets fictionnels provoque inévitablement la question de savoir en quoi nous pouvons garantir que l'objet présent à l'esprit du lecteur A est le

même que celui présent à l'esprit du lecteur B.

Si le réalisme artefactuel néglige le rôle interprétatif du lecteur dans les jugements d'identification, c'est évidemment parce qu'il souhaite faire reposer les conditions d'identité des personnages sur l'objectivité d'un artefact dont elles dépendent ontologiquement. L'artefactualiste peut alors hypostasier l'existence d'entités abstraites sans éprouver la culpabilité engendrée par le complexe d'Occam. Le rôle du lecteur est négligé parce qu'il vient perturber cette dépendance ontologique des objets abstraits ravalée aux seuls textes. Or les intentions de la communauté des lecteurs représentent un aspect fondamental de la réussite de l'identification transfictionnelle. L'œuvre littéraire n'est pas seulement un produit dont la consommation serait confinée à l'isolement d'un lecteur soucieux d'une intimité avec le texte qu'il lit. C'est par l'effet d'un jugement esthétique, mobilisé par des discussions visant à un consensus, qu'une œuvre gagne ou non son statut d'élément à l'intérieur d'une série.

Pour résumer, selon la conception de la dépendance historique rigide des personnages aux actes de création de leur auteur défendue par Thomasson, un *fictum* est d'abord le contenu purement intentionnel d'un acte de pensée de son créateur, un *intentionale*, autrement dit, un objet immanent vers lequel est dirigé un état d'esprit. Mais pour persister en dehors de cet acte de représentation, autrement dit, pour prolonger sa vie d'objet intentionnel au delà de l'acte inchoatif de pensée et parvenir au stade de *fictum*, celui-ci doit encore dépendre constamment de l'existence d'un exemplaire de l'œuvre (où d'une activation mentale de cette œuvre), support qui lui permet de prolonger son existence. Aussi, selon cette conception, la pensée est capable non seulement de susciter l'existence d'entités mentales purement intentionnelles éphémères, dont l'existence dure autant que l'acte de pensée qui les concerne, mais également des entités beaucoup plus persistantes comme les *ficta*, dont la vie se prolonge par l'existence constante de supports. Dans cette théorie les *intentionale* sont vus comme supportés par des actes de langages, ils gagnent ainsi leur caractère objectif en raison du fait qu'ils ne sont pas réduits aux supports des actes de mentaux. On peut toutefois s'interroger sur l'étonnante capacité qu'auraient des actes langagiers à générer de nouvelles entités, et finalement, le réalisme externe des artefacts abstraits n'est pas moins ontologiquement onéreux que le réalisme interne.

Diversité des origines

En outre, associer un acte de création d'un auteur à une sorte de baptême est sans doute problématique. Il peut d'abord y avoir un problème d'origine non créative car certains personnages, comme le roi Arthur chez Chrétien de Troyes, comme Hamlet chez Shakespeare ou encore comme Œdipe chez Sophocle, sont manifestement associés à des personnages légendaires ou mythologiques dont la création semble plus avoir rapport avec des jeux sociaux communautaires qu'avec des actes de création isolés. Il sera difficile, dans ce cas, de relier le nom à une chaîne causale renvoyant à la création du personnage par un auteur. Et même lorsqu'on admet que le personnage est le fruit d'un auteur, on peut mettre en doute l'idée que celui-ci en ait l'idée d'une manière instantanée. Ne s'agit-il pas plutôt, comme toute création, d'un processus morcelé et s'étendant dans le temps, où des actes mentaux se succèdent sans nécessairement avoir des objets de pensée identiques ? Et de toute façon, même lorsqu'une inscription matérielle semble venir « cristalliser » le personnage, il reste tous les problèmes génériques d'avant-textes que nous avons soulevé dans les chapitre

consacré aux doubles des œuvres. Dans une perspective réaliste, il paraît donc absurde de rendre équivalent le *fictum* abouti et l'objet intentionnel inchoatif (la première pensée de Conan Doyle à propos de Holmes).

La critique génétique peut également révéler tout un ensemble de problèmes liés aux noms des personnages de fiction et à leur putative référence dans une optique réaliste. Nous pouvons par exemple nous demander si le personnage dont le nom de jeune fille est Madame Lestiboudois dans un premier scénario de Flaubert désigne ou non le même personnage que celui qui porte le patronyme de Rouault dans la version publiée de *Madame Bovary*. En outre, un seul et même personnage dans l'avant-texte peut donner lieu à deux personnages dans l'œuvre publiée, voire inversement, deux personnages de l'avant-texte peuvent fusionner en un seul dans la version éditée.

Dans la théorie des objets inexistants, les *ficta* associés à un personnage à un moment de la lecture et à un autre ne peuvent être les mêmes, en raison qu'après, des propriétés seront ajoutées, et donc que les ensembles qui circonscrivent le personnage ne sont pas identiques au fil du processus de croyance feinte. C'est donc tout au long de la lecture que défilent des *ficta* différents, à chaque nouveau prédicat attribué au personnage. Aussi, si elles ne font pas appel à une théorie des particuliers fins, et qu'elles considèrent au contraire que les personnages sont des faisceaux de propriétés, de telles théories impliquent alors que le *fictum* n'apparaît finalement qu'en toute fin de processus narratif, puisqu'en cours de route, il se construit progressivement par empilements successifs de propriétés.

Alberto Voltolini semble accepter complètement cette idée quand, considérant le jeu de croyance feinte en quoi consiste tout acte d'immersion fictionnelle, il soutient que :

> Sooner or later, as in the case of any other game, this game is over. Once it is over, people can start thinking of it as the process in which a certain set of properties is made believe to be such that the properties corresponding to its properties are instantiated by a certain, typically concrete individual. As soon as this happens, a fictional entity comes into existence, as made up of that process(-type) together with the very set of properties which that process is then seen as involving. Participants in the game could not think of such an entity. Since they are in the game, they simply imagine that there is a (concrete) individual having the properties corresponding to the members of the set in question. (VOLTOLINI, 2006, p. 102)

Reste à savoir ce qui constitue la fin d'un processus narratif. C'est généralement supposé être le point final d'une publication, le processus de jeu de croyance feinte entrepris par le lecteur pouvant être discontinu. Mais il reste alors à savoir ce qui assure au lecteur qui reprend une lecture interrompue d'identifier le même personnage. Le personnage de Roland fait par exemple l'objet de plusieurs histoires, entre autres et parmi les plus connues, celle du *Roland Amoureux* de Boiardo, et après elle, le *Roland Furieux* de l'Arioste. La seconde histoire est considérée comme ou une prolongation (séquelle) de celle de Boiardo.

Entre les deux fictions, avons-nous affaire à un personnage qui, devenu fou alors qu'il était avisé, a changé, ou s'agit-il au contraire de deux personnages ? En réponse à cette question, la théorie syncrétique de Voltolini, qui tente de concilier les artefacts abstraits aux entités inexistantes, définit les *ficta* comme des objets abstraits caractérisés par des propriétés internes dans des processus de narration. Or un même processus, soutient Voltolini, ne peut donner naissance à deux

ficta différents. La solution de Voltolini consiste alors à faire porter l'identité des personnages sur l'identité des processus narratifs considérés comme des jeux de croyance feinte. Pour cela, Voltolini distingue deux types de processus : d'une part des processus-phases, par exemple chacun des récits précédents, et d'autre part un processus-intégral, qui est la somme méréologique des phases précédentes. C'est que le Roland du *Roland Amoureux* et le Roland du *Roland Furieux* sont des personnages différents en rapport aux processus-phases que sont les deux histoires de Boiaro et de l'Arioste, mais qu'il s'agit du même *fictum* en rapport au processus-intégral. Cette manière de considérer les choses peut ainsi être approchée de la théorie de la perdurance qui définit les individus particuliers comme des sommes méréologiques de parties temporelles, sinon qu'ici, il ne s'agit pas d'individus concrets mais d'ensembles de propriétés qui sont *imaginées* appartenir à un particulier.

Les personnages indiscernables et récurrents

Le réalisme externe est-il capable de rendre compte des personnages indiscriminables ? De manière intratextuelle, l'identité des descriptions implique l'identité des personnages. Lorsqu'un auteur crée une foule fictionnelle, il crée un ensemble dont nous ne savons pas au juste de combien d'individus il est composé. Les mêmes problèmes d'indiscernables occultifs apparaissent, liés à la question du caractère indéterminé des objets fictionnels.

Le critère de Thomasson ne précise pas le genre de propriétés qui sont déterminantes pour assurer l'identité. S'il s'agit des seules propriétés intrinsèques, nous sommes alors amenés à ne jamais considérer que deux personnages soient indiscernables, cette similarité étant rabattue sur l'identité. De ce point de vue, Goliadkine « aîné » est identique à Goliadkine « jeune ». Cela semble restreindre l'interprétation que nous pouvons donner au roman, en évacuant de manière malheureuse la possibilité de personnages numériquement distincts, conçus comme des répliques pures.

Un autre problème, soulevé par Alberto Voltolini (VOLTOLINI, 2006, p. 49 n. 34) est celui de la récurrence métaleptique du nom. Selon la conception métaphysique de Thomasson, certaines assertions critiques, autrement dit, méta-fictionnelles, comme « (SFD) : Sherlock Holmes est un personnage de fiction créé par Conan Doyle » sont nécessairement vraies, autrement dit, vraies dans toutes les circonstances d'évaluation où Sherlock Holmes existe. Or si nous évaluons cette même proposition de manière interne à la fiction, autrement dit, selon les circonstances fictionnelles, on doit bien admettre que c'est faux. Voltolini constate qu'il y a donc un monde possible où Sherlock Holmes n'est pas un personnage de fiction, ce qui contredit la théorie des artefacts abstraits. Mais l'argument n'est sans doute pas recevable dans la mesure où dans le contexte d'évaluation fictionnel, Sherlock Holmes n'existe pas comme personnage, mais comme détective, autrement dit, comme être vivant. L'affirmation (SFD) affirme que le contexte est celui où Sherlock Holmes existe en tant que personnage, et où Doyle existe donc en tant qu'auteur, ce qui n'est pas le cas du contexte fictionnel.

La théorie des artefacts abstraits n'est pas confrontée au problème de Borges, puisque la distinction numérique des auteurs assure la distinction numérique des personnages. Radicalisant l'expérience de pensée du Pierre Ménard, Voltolini va encore plus loin, proposant de considérer l'existence de deux astres terrestres indiscernables, Terre et Terre-jumelle, sur lesquels des individus eux-mêmes indiscernables écriraient exactement les mêmes récits, avec exactement les mêmes

mots (VOLTOLINI, 2006, p. 33). Considérons en particulier l'un des personnages de tels récits, disons Holmes-de-Terre et Holmes-de-Terre-jumelle. Notons que les mathématiques, et en particulier la théorie des ensembles, devraient être uniformément les mêmes dans ces deux mondes. Les critères d'individuation internes impliqueraient dès lors que les ensembles de propriétés attribuées respectivement à Holmes-de-Terre et à Holmes-de-Terre-jumelle seraient strictement identiques. Les *ficta* seraient donc les mêmes, mais Voltolini constate que la théorie des artefacts abstraits impliquerait manifestement d'avoir affaire à des personnages numériquement distincts.

De fait, je conviens qu'il faudrait effectivement accorder cette différence numérique, parce que les actes de langage seraient bien différents, ces actes étant ancrés à des contextes distincts. Mais dans ce cas, il ne semble guère n'y avoir que deux possibilités : soit que les *ficta* décrits ici ne jouent guère de rôle dans l'individuation des personnages qui sont finalement seulement identifiés par des actes de langage, des jeux d'imagination, soit que de tels *ficta*, relèvent de doubles purs, d'indiscernables haeccéistiques.

Les conséquences de la réification des intentions

La théorie de Thomasson, sous couvert d'explications métaphysiques, attribue de mystérieuses capacités créatrices aux actes de langage, dont on peut douter qu'ils puissent peser bien lourd, contre ce que prétend cette théorie, dans l'ontologie. Toute la question est finalement de savoir ce qui permet à des *intentionale* de parvenir au statut d'artefacts. Ainsi, comme le fait remarquer Voltolini,

> If it is not clear both *whether* a certain mental act may bring into existence a purely intentional object, and *how* a purely intentional object can turn out to be an entity that is kept in existence by a literary work, that is a fictional entity, it will not be clear either how thoughts can generate fictional entities. (VOLTOLINI, 2006, p. 50)

Thomasson pourrait néanmoins répondre ici que les *intentionalia* ne font pas partie d'un catégorie ontologique différente de celle des *ficta* : au contraire, les dernières sont juste un cas particulier des premières. Mais cela ne fait que ramener le problème un cran en arrière, et il reste encore à déterminer les conditions d'identité des *intentialia*. Hob et Nob pensent-ils au même objet ? Et qu'est-ce qui différencie la sorcière de leurs pensées de Hamlet ? Faut-il admettre que de tels objets appartiennent à la même catégorie ?

Une alternative à la théorie des artefacts abstraits consisterait à substituer l'acte de création d'un auteur (en tant que phénomène quasi-instantané) au *processus* du récit, comme phénomène de représentations collectives, ce processus étant éventuellement interactif entre l'auteur et la communauté de réception. Car à la différence de l'histoire de Nob et Hob, un personnage possède un statut culturel différent, lié à sa diffusion médiatique. Si nous pouvons douter qu'un acte mental de création d'un auteur puisse générer, de manière historiquement rigide une entité métaphysique, afin de sauver la conception de Thomasson, il reste la possibilité que cette entité soit générée par la diffusion constante du support médiatique, par l'œuvre. On admet ainsi qu'il s'agit d'une entité *sociale*, culturelle, comme le sont les nations et les institutions en général. Est-ce encore compatible avec un point de vue ontologique ? Il faudrait manifestement reconnaître que la nature d'une telle entité serait extrinsèque, et non intrinsèque.

Ce qui est engendré par un processus de récit n'est pas un artefact, comme le serait un bateau, mais l'ouverture d'un ensemble de dossiers mentaux [17] associés à un même jeu de croyance feinte rapporté à un texte, voire à plusieurs textes. Il ne s'agit donc pas d'un problème d'individuation des personnages, mais d'identification des dossiers se rapportant aux mêmes processus de croyance feinte. D'un côté, le contenu de ces dossiers peut varier de sujet à sujet (nous ne nous faisons pas la même idée de Sherlock Holmes). Mais d'un autre côté, si un contenu montre des écarts trop importants par rapport à ce qui est attendu dans le processus de croyance feinte, la communauté de réception de l'œuvre refusera de l'associer au même jeu. Il s'agit donc d'une régulation culturelle de dossiers mentaux. La représentation est médiatisée par la ressemblance, sans devoir souscrire pour autant à une ressemblance parfaite, autrement dit, à des cas métaphysiques de répliques pures.

5.3 Pour une théorie dynamique de l'identification des personnages

Nous l'avons vu, les théories réalistes se heurtent à trois problèmes : soit elles refusent l'éventualité de personnages indiscernables en rabattant l'indiscriminabilité parfaite sur l'identité (cas du réalisme interne), soit elles hypostasient des doubles purs même pour parler de personnages peu ressemblants (cas du réalisme possibiliste), soit elles individuent les personnages à la manière des œuvres, repoussant ainsi les problèmes d'individuation des personnages sur ceux des actes de création des auteurs (cas du réalisme externe). Pour ma part, je défends l'idée qu'il n'y a pas d'entité qui fournisse les conditions d'individuation des personnages de fiction. Les personnages ne sont pas assimilables à des entités, et leur identification n'est que purement conventionnelle. L'identification intentionnelle n'est pas une identité métaphysique. Mais, comme je le montrerai également, cela n'implique pas nécessairement qu'on puisse identifier un personnage donné à n'importe quoi d'autre.

5.3.1 Un réquisit pour une théorie dynamique de l'identité conventionnelle

La conception dynamique de l'identité implique inévitablement une difficulté pour la théorie sémantique de la référence rigide : si le nom de Goliadkine réfère à un objet abstrait, cet objet est inévitablement le même et différent de tout autre dans tous les mondes possibles où il existe. Mais, nous l'avons dit, la mécanique du récit de double ne peut se satisfaire d'identités transmondaines aussi contraignantes.

En outre, les théories de la désignation directe renoncent à la médiation que le sens frégéen portait entre les expressions singulières et les références. Le descriptivisme frégéen est en effet motivé par la thèse que toute différence cognitive provient d'une différence sémantique. Ainsi, si l'on pense à *deux* objets, la différence numérique des pensées doit provenir d'une différente de signification, soit par la différence de référence, comme lorsqu'on pense à la planète Vénus et à la planète Mars (références différentes et donc sens différents), soit par la différence du mode de présentation (la différence de sens) comme lorsqu'on pense à Hesperus et à Phosphorus (même référence mais modes d'accès différents). Lorsqu'il s'agit de noms vides, seule la différence de

[17] Sur la conception des dossiers mentaux, voir (RECANATI, 2007a).

mode de présentation permet encore de distinguer numériquement les pensées. Selon cette conception, une même signification implique une même pensée. Les théories de la désignation directe, en évacuant le mode de présentation de la référence, perdent l'explication que Frege donnait à la différence de valeur cognitive entre des énoncés du type $a = b$ et des énoncés triviaux comme $a = a$.

La question qui nous concerne plus particulièrement est de savoir comment nous pouvons intuitivement penser la différence numérique de choses au demeurant parfaitement indiscriminables, en particulier quand de telles choses n'existent pas de manière spatio-temporelle. Cette difficulté est en effet particulièrement aggravée dans le cas de fictions où l'on demande de penser à deux personnages indiscernables. Comme l'indique Wittgenstein dans le *Tractatus*, l'un des moyens de différencier des concepts est de marquer formellement l'identité numérique par l'identité des signes, une différence de signes impliquant alors une différence de concepts. Mais comme le révèlent les cas de Goliadkine et de William Wilson, les personnages doubles portent bien souvent les mêmes noms. Comment, dès lors, marquer encore une différence numérique des pensées ?

La question est précisément de déterminer ce qui nous permet d'identifier deux concepts nominatifs. D'abord, il ne s'agit pas d'une référence commune, puisque nous pouvons penser à une même chose de deux manières différentes sans nous rendre compte qu'il s'agit de la même chose, comme dans le cas des astronomes babyloniens à propos d'Hesperus et de Phosphorus. Ensuite, ce n'est pas non plus leur contenu, puisque nous pouvons avoir des contenus de concepts parfaitement identiques sans identité de ces concepts, par exemple parce qu'ils référeraient à des choses différentes lorsque nous croyons, à tort, être en présence d'une seule et même chose là où s'en tiennent en vérité deux. Ce n'est pas non plus la relation info-générative qui est à l'origine de leur "remplissage" parce qu'une même relation info-générative peut être à l'origine de concepts nominatifs différents.

Une réponse consisterait à en nier la possibilité de penser à deux personnages indiscernables, sous couvert d'illusion : nous croyons avoir des pensées différentes, mais nous nous méprenons. C'est une forme de théorie de l'illusion comme en a proposé Saul Kripke dans le cas des identités métaphysiquement impossibles.

Fodor (FODOR, 1987) a réintroduit, dans les théories externalistes, sous l'appellation de « contenu étroit », le rôle épistémique joué par le mode de présentation, mais de manière indépendante de la constitution de la référence. C'est aussi pour cette raison que Putnam (PUTNAM, 1975) utilise la notion de stéréotype, qui assure la fonction de mode de présentation de la référence. Oscar-T et Oscar-TJ ont le même contenu de pensée, à savoir, le même stéréotype, mais ils ne réfèrent pas aux mêmes choses parce qu'ils ne sont pas placés dans le même contexte. Il faut donc une (vraie) double sémantique, désolidarisant les deux fonctions frégéennes du sens et de la référence : une fonction sémantique reliant directement les expressions au monde, mais aussi une fonction représentationnelle assurant un choix d'aspects par lesquels la référence est identifiée dans l'acte de pensée. Si l'on ne veut pas hypostasier des *ficta*, il semble alors nécessaire d'admettre une sémantique bifide en considérant que les noms des personnages de fiction sont bien des modes de présentation en dépit du fait qu'ils sont associés à des noms vides.

Voici donc un réquisit : des théories de la désignation rigide, l'on doit garder l'idée que la constitution de la référence doit être indépendante de la manière de se représenter cette référence.

En revanche, on doit retenir de Frege l'idée que la référence est toujours visée sous un certain aspect. De manière synthétique, on peut alors retenir que la fixation de la référence, c'est-à-dire l'accès sémantique à la référence, doit être désolidarisée de l'identification de la référence, c'est-à-dire de l'accès épistémique à la référence (ce réquisit sera désormais noté « R »).

Pour satisfaire à ce réquisit, certains auteurs proposent de réajuster la théorie millième en affirmant que le véhicule de la représentation est suffisant pour leur différence numérique [18]. Des propositions ayant exactement la même signification peuvent relever de pensées numériquement différentes [19]. En particulier, un concept nominatif a peut s'avérer être numériquement différent d'un concept b, sans différences sémantiques entre a et b. Ce n'est plus seulement comme conceptions distinctes médiatisées par des différences de significations que les pensées peuvent différer numériquement à partir de leur contenu, mais également par le véhicule même du concept : « the vehicles of content can affect cognition in ways that do not depend on their semantic properties. » (R. M. SAINSBURY & TYE, 2013, p. 26) Il est donc inutile, comme le fait Frege, de compléter la dénotation par le sens, puisque la représentation suffit d'elle-même à convoyer la différence numérique des pensées.

Ainsi, la où la théorie de Frege limite radicalement la pensée du double en l'ajustant de manière systématique sur une différence de significations, les théories « du véhicule » permettent de rendre compte des doubles épistémiques en recourant à la médiation du mode de présentation.

5.3.2 La mêmeté comme relation de ressemblance

Dans *Objects of Thought* Arthur Prior (PRIOR, 1971, pp. 3) a distingué deux types d'objets de pensée, propositionnels et objectuels, ressortissant respectivement à ce « que l'on pense » et à « ce que l'on pense de ». Les objets de pensée ne concernent selon Prior que la seconde forme, c'est-à-dire, ce vers quoi les états mentaux sont dirigés. Ces choses peuvent être aussi bien des individus du monde que des choses seulement *représentées comme* étant des individus du monde, mais sans en être réellement. Dans *The Objects of Though*, Tim Crane reprend cette idée, insistant sur le fait que les objets intentionnels n'ont nullement la nécessité d'exister. La représentation de choses qui ne sont pas dans le monde n'a pas besoin de s'appuyer des entités exotiques, sur des objets intentionnels non-existant ou simplement possibles : la réalité ne contient « pas plus que ce qui existe » (CRANE, 2013, p. 5).

Pour autant, Crane constate qu'il y a bien, entre les attitudes intentionnelles de Hob et de Nob dans l'exemple de Geach, une même représentation portée par un nom vide. Mais Crane refuse de considérer que le sens de « même » se rapporte ici à l'identité. Car celle-ci ne concerne, insiste Crane, que les entités réelles. Des pensées singulières qui portent sur des fictions sont les « mêmes » dans un sens plus lâche, signifiant « très ressemblantes ». La sorcière dont Bob est victime n'est pas, au sens littéral, identique à celle de Cob, parce qu'elle n'existe pas. En revanche, la sorcière de Bob est la même que celle de Cob au sens où elle ressemble complètement à la sorcière dont Cob est victime, tout comme l'on dit que telle voiture est la même que telle autre du

[18] Ainsi, sont admis que les représentations fournissent un aspect (*guise*) (Salmon 1986), un mode de présentation (Schiffer 1992), ou une notion (Crimmins and Perry 1989).

[19] Cette conception est défendue notamment par Richard Sainsbury et Michael Tye (R. M. SAINSBURY & TYE, 2013).

point de vue de l'identité qualitative. Crane se réfère ici à ce qu'écrivait déjà Daniel Dennett en 1968 :

> Whenever we wish to disavow belief in the actual existence of any intentional objects, 'same' can at best mean 'exactly similar in characteristics'. (D.C. DENNETT, 1968, p. 337)

On a en quelque sorte, concernant les noms fictionnels ou fictifs, un renversement du principe d'identité des indiscernables : c'est maintenant l'identité intentionnelle est complètement rabattue sur la ressemblance :

> My proposal is that what is at issue in cases of mere intentional identity is not identity at all, but similarity of representation. We can talk about sameness and difference here, but what makes this talk true, when it is true, is the similarity or sameness of properties. We count Mercury and Hermes as the 'same' god, because the stories that are told about them are relevantly similar : they say similar things about these two characters. There is no fact of the matter about whether Mercury and Hermes are identical [...] What I am proposing is not that similarity of represented properties suffices for identity, but that it explains our talk of objects being 'the same'. Identity of non-existent objects is a confused idea. (CRANE, 2013, p. 164)

Nous pouvons aussi complètement détacher l'identification de la ressemblance, par une identification purement intentionnelle, déclarative. Qu'est-ce qui pourrait nous conduire à découvrir l'identité entre des représentations distinctes d'un même personnage ? Crane constate que la seule manière de découvrir que le docteur Jeckyll est Mister Hyde ne consiste pas à découvrir empiriquement qu'un certain $\textit{fictum}_1$ est le même qu'un $\textit{fictum}_2$, mais de l'apprendre de la plume de l'auteur. Dans le cas des fictions, les seules identités *découvertes par les lecteurs* sont les identités *stipulées par les auteurs*. Il n'y a aucun autre moyen, insiste Crane, de déterminer des identités intentionnelles. Nous ne découvrons pas par nous-mêmes ou par une quelconque enquête sur le monde que tel personnage est ou n'est pas identique à tel autre. Nous l'apprenons en lisant les mots de l'auteur.

Crane arrête là son analyse des identités stipulées, mais c'est sans doute un peu trop tôt. Car attribuer des identités comme l'on attribue des propriétés n'implique pas de dire n'importe quoi. Souvenons-nous par exemple des polémiques qui ont accompagné la sortie du premier *Batman* de Tim Burton, dans lequel les parents de Bruce Wayne étaient assassinés par Jack Napier, alias le Joker, alors que dans la série de son créateur Bob Kane, il s'agissait d'un malfrat du nom de Joe Chill n'ayant rien à voir avec le Joker. L'identité « Joe Chill = Joker » stipulée par Burton a été contestée et rejetée par les plus fervents amateurs de la série, tant et si bien que les épisodes suivants des aventures du super-héros ont abandonné cette identification apparemment malvenue. Nous retrouvons ici le même débat que celui que nous avons entrepris avec l'interprétation des œuvres : tout comme il y a des limites à l'interprétation des histoires, il y a des limites aux identités stipulées qui sont celles que les récepteurs des fictions sont prêts à accepter. Car s'il est vrai que rien n'empêcherait un auteur de stipuler que Batman est la même personne que le professeur Tournesol, rien ne garantirait cependant le succès de cette identification, qui finirait à n'en pas douter dans les marges de la fiction, car cette stipulation n'aurait sans doute aucun intérêt, et serait bien vite oubliée. L'identité intentionnelle, en fiction, est essentiellement une identité de consensus. C'est

une identité qui persévère dans et par les jeux de croyance feinte des récepteurs. Or une identité stipulée qui s'avérerait idiote rencontrerait sans doute peu d'écho chez les récepteurs de la fiction, et ne parviendrait pas à gagner une quelconque pérennité. Les identités fictionnelles sont certes des identités stipulées, mais pour que l'attribution puisse parvenir à l'état d'attribution conférée à un objet culturel, elles doivent faire, en quelque sorte, autorité, c'est-à-dire, être reconnues par la communauté de réception esthétique.

Je souhaite maintenant comparer deux positions prometteuses quant au réquisit R, à savoir, la théorie des dossiers mentaux et la théorie des lignes de mondes. Dans l'une et dans l'autre apparaît l'idée que les énoncés d'identité sont contextuels, et qu'ils doivent être jugés de manière pragmatique depuis le monde (épistémique) où ils sont formulés. Les mondes épistémiques accessibles pouvant varier, on retrouve ainsi une dynamique permettant des rapports changeants à l'identification.

5.3.3 La théorie des dossiers mentaux

Remarquons que le fait qu'un lecteur puisse s'interroger sur l'identité entre deux personnages suggère qu'il utilise deux modes de présentation distincts pour le (les) représenter. En effet, le doute quant à une identité implique inévitablement la diversité des modes de présentation des personnages afférents. La diversité des modes de présentation pour une même référence, voire pour une absence de référence, a été développée par Peter Strawson avec la théorie des dossiers mentaux [20]. Un dossier mental est un particulier mental qui ressemble, dans son fonctionnement, à un dossier cartonné. Il sert à stocker un contenu dont la référence est un objet du monde. Le dossier est engendré par une relation « épistémiquement profitable », ou relation info-générative, comme la perception, la proprioception, la communication d'informations par un agent fiable, etc. Contrairement au mode de présentation frégéen dont dépend la référence d'une expression, le rapport du dossier mental à la référence ne dépend pas de son contenu, ce qui explique que dossier puisse faire référence à un objet donné sans qu'aucune des informations qui sont stockées dans ce dossier ne s'avère pertinente. De fait, il ne s'agit donc pas d'une théorie sémantique relevant de la satisfaction des descriptions afférentes à la référence visée, mais d'une théorie relationnelle : le dossier manifeste la relation d'un agent par rapport à un objet du monde. La théorie des dossiers mentaux permet ainsi de conserver, de la sémantique de Frege, le fait que le contenu mental explique le comportement d'un agent, mais sans admettre pour autant que la référence soit déterminée par ce contenu.

La conception relationnelle des concepts nominatifs

Un dossier se rapporte ainsi à un concept, celui-ci pouvant être aussi bien général (un tigre) que singulier, autrement dit nominatif (Napoléon, Sherlock Holmes). Reste à savoir comment est produite la relation info-générative qui engendre des dossiers mentaux dans le cas particulier des personnages de fiction, puisqu'il ne s'agit clairement pas d'une relation empirique entre un agent et un objet du monde. Selon François Recanati, le dossier concernant un personnage de fiction est informé par le canal d'une relation de faire-semblant. Dans les schémas narratifs les plus simples,

[20] La littérature anglo-saxonne privilégie plutôt l'appellation de « fichiers mentaux (*mental files*) ».

le narrateur d'une fiction fait comme-s'il rapportait des faits réels. De son côté, le lecteur qui découvre un personnage fait comme s'il entendait parler de quelqu'un de réel. Par cette source d'informations feintes, il ouvre un dossier mental dont il nourrit le contenu à partir des informations révélées par le narrateur. Il s'agit donc, pour le lecteur, de se rapporter de manière causale, moyennant le rapport du narrateur considéré comme fiable, à un individu fictif par le moyen d'une relation de croyance feinte.

La conception syntaxique de Jerry Fodor, dont l'objet est d'expliquer le comportement d'un agent au moyen de ses représentations mentales, met en parallèle la représentation mentale et la représentation syntaxique. Cette théorie affirme elle aussi se passer du niveau du sens frégéen tout en maintenant un mode de présentation, celui-ci étant entièrement véhiculé par la représentation elle-même, sans appel à un autre mécanisme. De ce point de vue, certains dossiers mentaux sont indiscriminables. Admettons en effet qu'une personne rencontre un joueur d'une équipe de sport, Bob, une première fois, puis une seconde fois, mais sans se rendre compte qu'il s'agit de la même personne [21]. Cette personne générera ainsi deux dossiers mentaux correspondant respectivement à l'occurrence Bob_1 et à l'occurrence Bob_2. Ces deux dossiers seraient parfaitement ressemblants, non seulement quant à leur contenu et à leur référence, mais aussi quant à leur source info-générative. Ces dossiers seraient néanmoins numériquement distincts, puisque l'agent concevrait deux personnes là où, en fait, il n'y en aurait qu'une seule. Dans ce cas, l'agent ne pourrait pas faire intervenir un dossier plutôt qu'un autre pour se représenter une occurrence plutôt qu'une autre. De fait, il n'y aurait pas ici de différence d'attitude du sujet à l'égard des deux Bob, autrement dit, cela ne pourrait aucunement expliquer des différences de comportement du sujet à l'égard de l'un d'entre eux. Cette conception semble taillée pour la conception d'indiscernables, puisqu'elle affranchit complètement la ressemblance parfaite de l'identité numérique. Mais, du fait que les dossiers mentaux sont de caractère épistémique, il n'y a guère de raison d'avancer ici des indiscernables ontologiques, des doubles purs.

Les dossiers mentaux comme particuliers

Les dossiers mentaux sont caractérisés par le type de relations info-génératives qui les nourrissent, autrement dit, par un certain type de relation causale que l'agent entretient avec la référence. Ces relations peuvent varier, relever de la perception, de la proprioception, du lieu, du temps, du rapport déictique avec un individu du monde, etc. Chaque type de dossier présuppose en effet une certaine relation à l'objet auquel se rapporte le dossier. Ces présuppositions expliquent le comportement du sujet, elles jouent un rôle crucial proche du mode de présentation présent dans la sémantique de Frege, tout en étant des éléments purement mentaux. Mais Recanati souligne qu'on peut ouvrir deux dossiers mentaux par la même source info-générative, dossiers ayant la même référence et le même contenu. Cela peut arriver par exemple lorsqu'on regarde un arbre autour duquel court un écureuil. On peut croire qu'il s'agit en fait de deux écureuils qui se pourchassent autour du tronc, alors qu'il n'est est rien [22]. On pourrait objecter une différence dans l'ordre d'apparition des écureuils, mais cette information pourrait être amenée ensuite à disparaître du dossier. On aurait dans ce cas deux dossiers mentaux numériquement distincts ne différant ni par le contenu, ni par

[21] L'exemple est de James Pryor (PRYOR, 2016).

[22] L'exemple est adapté de Peter Pagin (PAGIN, 2013, p. 140).

la référence, ni même par la relation info-générative.

Dans la conception des dossiers mentaux de Recanati, c'est l'origine du dossier mental qui sert à fixer la référence du dossier [23]. La distinction numérique des dossiers, seraient-ils indiscernables en tout autre point, est due aux différences de dates de création des dossiers.

Considérons les choses depuis le point de vue de la fiction : le concept nominatif de Sherlock Holmes relève de l'ouverture, chez le lecteur, d'un dossier mental particulier, comme s'il s'agissait d'une personne existant réellement, quoique ce dossier ne fasse, en réalité, référence à rien. En revanche, précise Recanati (RECANATI, 2018, 1 :00 :00), dans les cas métafictionnels, cette fois, il y a bien référence, non à un personnage, mais à un artefact culturel : on déploie un dossier qui ne porte pas sur le personnage de chair et d'os, mais sur une création littéraire. Ainsi, la référence d'un personnage de fiction est assurée par un double dossier. Dans le dossier fictionnel, on range des informations développées par des jeux de croyance feinte. Dans le second dossier métafictionnel, on range les informations se rapportant à l'artefact culturel. Mais Recanati souligne par ailleurs que la seule façon externe de se représenter ce qu'il y a dans une fiction, c'est de s'immerger dans la fiction elle-même. Le dossier métafictionnel doit donc contenir à la fois les informations externes et les informations internes. Plutôt que de considérer que le dossier fictionnel est dédoublé dans le dossier méta-fictionnel, Recanati considère que le dossier métafictionnel doit contenir un vecteur renvoyant en direction du dossier fictionnel. Cette théorie permet d'expliquer l'apparition de personnages indiscernables, ne serait-ce que l'ordre de leur apparition dans les histoires : les dossiers mentaux afférents aux personnages indiscernables sont particularisés par l'ordre de leur apparition dans le récit. C'est d'ailleurs un mécanisme qu'emploie Dostoïevski dans *Le Double* pour distinguer Goliadkine « aîné », le premier des deux personnages qui apparaît au cours du récit, de Goliadkine « le jeune », qui apparaît plus tardivement dans le récit. Au début du roman de Dostoïevski, nous sommes amenés à convenir d'une relation particulière avec le personnage de Goliadkine, *comme si* l'on parlait d'un personnage réel duquel nous pourrions être en accointance. Nous avons donc une relation épistémiquement profitable qui génère un dossier mental. L'apparition du double relève alors de la création d'un second dossier mental où nous pouvons stocker des informations absolument identiques aux précédentes. Ce qui distingue numériquement les dossiers ne sont pas alors la représentation elle-même, mais les moments numériquement distincts de leur génération.

La norme de fusion strawsonienne

Dans sa conception des fichiers mentaux servant à désigner les concepts nominatifs, Peter Strawson a introduit une norme consistant à associer à chaque objet du monde *un et un seul* dossier. Ainsi, la découverte que deux concepts nominatifs correspondent en fait à un unique objet, comme par exemple la découverte qu'Emile Ajar *est* Romain Gary, cette découverte implique la fusion des deux fichiers en un seul, suivie de la disparition des deux fichiers originaux.

Ruth Millikan souligne que cette conception implique une conception particulière de l'identité passant par un appariement un à un des objets dans l'esprit et des objets dans le monde. Chez Strawson, l'identité d'un particulier est ainsi représentée par l'identité d'un particulier mental.

[23] L'individuation des concepts nominatifs par l'origine est une idée qui est également défendue par Mark Sainsbury et Michael Tye (R. M. SAINSBURY & TYE, 2013).

Ce qui, dans l'esprit, représente l'identité numérique d'un objet, c'est l'identité numérique du symbole mental qui représente cet objet. La théorie de l'identité est vue ici comme imitation, dans la représentation, d'une identité réelle dans le monde. On a donc affaire à une théorie iconique de l'identité [24], comme le souligne Ruth Millikan :

> So long as you haven't made any mistakes, everything you know about your mother is attached to the same particular mental representation of your mother, the same token. Your understanding that all these facts are facts about the same woman consists in the representations of the logical predicates of each of these facts being attached to numerically the same "dot" in your mind or brain. Call this the "Strawson model" of how identity or sameness is thought. (MILLIKAN, 1997, p. 500)

La conception de l'identité de Strawson entre directement en conflit avec les différents modes de présentation frégéens selon lesquels un objet peut être conçu sous un aspect ou un autre. En effet, dans une conception privilégiant un mode de présentation, lorsqu'on se rend compte que ce qu'on croyait être deux objets n'en sont en réalité qu'un seul, les différents aspects, habituellement séparés, sous lesquels apparaissent ces deux objets sont simultanément éprouvés. Il n'y a donc pas lieu de considérer que les concepts nominaux fusionnent en un seul. Par ailleurs, ce qu'on peut objecter à la conception de la fusion des dossiers, c'est qu'elle s'avère irréversible. Elle a en effet le tort de faire disparaître les deux voies distinctes d'accès à la même référence, ce qui paraît contre-intuitif parce que même avoir appris une identité, on peut toujours revenir à l'un des dossiers antérieurs, comme par exemple quand l'on considère le point de vue des astronomes babyloniens qui distinguaient Phosphorus d'Hesperus.

Dans William Wilson, nous apprenons, en fin du récit l'identité des deux personnages. Cela exige de mettre en jeu deux concepts nominatifs différents qui nous permettent, dans un second temps, d'être identifiés au même personnage. Il y a donc deux dossiers mentaux en jeu. Lors de la découverte d'une identité (dans le cas particulier de cet exemple, cette identité est bien sûr stipulée par Edgard Poe), Recanati propose de considérer que les fichiers mentaux ne sont pas fusionnés, mais liés. La liaison assure la circulation des informations entre les dossiers, tout en maintenant leur distinction numérique comme représentations. L'intérêt de conserver les deux dossiers est bien sûr de pouvoir revenir en arrière en imaginant à nouveau deux personnages distincts.

5.3.4 La théorie des lignes de monde

Suivant la voie frayée par Jaakko Hintikka, Tero Tulenheimo et Shahid Rahman ont offert une caractérisation des individus vus comme des « lignes de mondes ». Leur stratégie consiste à rapprocher la sémantique pluri-domaines d'Hintikka de la théorie aspectuelle des contreparties de David Lewis. La notion de ligne de monde apparaît d'ailleurs aussi chez cet auteur à travers la notion de stade, puisque Lewis laisse une place aux individus transmondains, ce qu'exige sa quantification non restreinte, bien que dans la conception de Lewis, ces entités prennent une teneur ontologique qui semble par trop coûteuse :

[24] Ce modèle est, notons-le, proche de celui exposé par Wittgenstein dans le *Tractatus*, assurant l'identité des choses par l'identité des signes.

> Je conteste non pas l'existence d'individus transmondains - si je quantifie sans restriction - mais l'attention qu'on leur porte. Je ne refuse pas l'existence des individus transmondains, bien que je dise qu'en un certain sens qu'il ne leur est en aucune manière possible d'exister. Il est possible pour une chose d'exister si et seulement s'il est possible à la totalité de cette chose d'exister. C'est-à-dire si et seulement s'il y a un monde dans lequel la totalité de cette chose existe. (...) Les parties de monde sont des individus possibles ; les individus transmondains sont des individus impossibles (LEWIS, 2007, p. 322)

On peut par exemple considérer l'individu impossible qui est la somme méréologique d'Emmanuel Macron perdant les élections et de telle tasse à café actuelle. La notion de *stade* est ainsi définie : « si un individu possible x fait partie de l'individu y transmondain, et si x n'est pas une partie propre d'un autre individu quelconque (possible) faisant partie de y, alors x est appelé un *stade* de y. » (LEWIS, 2007, p. 323)

De la même façon, Tulenheimo considère qu'un monde possible est constitué d'objets locaux, *worldbound*. Du point de vue d'une quantification de premier ordre, un objet est considéré comme n'importe quel élément du domaine de quantification d'un monde. Dans un cadre modal, un individu est quelque chose qui « apparaît » à de multiples reprises, ce que Tulenheimo appelle les « manifestations de l'individu », mais qui ne coïncide avec aucune d'entre elles. Un individu est alors défini comme une fonction, appelée « ligne de mondes » sélectionnant des objets dans des mondes possibles. La fonction est partielle du fait qu'il est possible que dans un monde, l'individu n'ait pas de manifestation. A la différence de Lewis, ces objets ne sont considérés que de manière sémantique, et non de manière métaphysique. Un objet est défini, d'un point de vue non modal comme un objet local, et d'un point de vue modal comme une ligne de mondes qui relie des objets locaux de plusieurs mondes. Aussi, le point de départ de Tulenheimo est de considérer que les mondes et les individus sont mutuellement indépendants. Les individus ne sont pas réductibles aux mondes ni les mondes aux individus. En revanche, individus et mondes peuvent s'expliquer en terme d'objets locaux qui les présupposent. Les objets locaux peuvent être sensément référés à l'intérieur d'une proposition de type extensionnelle, mais pas à l'intérieur d'une proposition intensionnelle. Un objet local n'existe pas dans différents mondes possibles. Ainsi, un monde possède un certain nombre d'objets locaux qui le constituent, ceux qui sont comprésents dans ce monde. De la même manière, un individu possède un certain nombre d'objets locaux qui le constituent, ceux qui permettent la manifestation de cet individu, qui sont les réalisations de cet individu. Si deux mondes différents n'ont aucun objet local en commun, en revanche, les individus ne sont pas sujet à cette restriction : deux individus différents peuvent partager les mêmes objets locaux.

Hintikka utilise les lignes de mondes de deux façons différentes : de manière épistémique comme manière d'identifier un même individu dans des mondes différents, mais également de manière transcendantale pour affirmer qu'il s'agit d'une condition préalable à nos discours contrefactuels. Il s'agit de construire les individus et les mondes comme des choses du même type, à savoir, des corrélations entre des objets locaux. Une manière de partitionner la totalité des objets locaux donne naissance aux mondes, et une autre manière de les partitionner donne naissance aux individus. Les lignes de mondes apportent ainsi une nouvelle interprétation à la sémantique modale proposant un traitement identique entre les objets physiques et les objets intentionnels. A la différence d'Hintikka, Tulenheimo précise que les lignes de mondes ne sont pas définies de manière métaphysique mais de manière purement sémantique :

> for Hintikka individuating functions are not any God-given entities but depend on our conceptual assumptions: the available individuating functions vary with the kind of modality we are interested in (knowing, believing, desiring, hoping, remembering, perceiving, being necessary) and with the totality of scenarios associated with a context of semantic evaluation. Individuals are circumscribed by relative, conventional, man-made boundaries. Now, evidently the emergence of individuating functions in our actual conceptual framework is tied to the overall 'ontological structure' of scenarios (properties and interrelations among objects in domains). (TULENHEIMO, 2009, p. 385)

La théorie des lignes de mondes défendue par Tulenheimo affirme qu'un contenu intentionnel est la résultante de deux types de représentations, à la fois celle d'un monde et celle d'un ensemble d'objets, que ce monde soit actuel ou non et que ces objets soient ou non réalisés. En outre, l'auteur considère que ce n'est donc pas la représentation, mais l'unité modale qui est primitive :

> For me, the notion of modal unity is a basic notion. Worlds are one variety of modal unities, world lines another. Not all world lines correspond to representations in the sense in which Crane uses this word. Intentionally individuated world lines do ; physically individuated world lines do not. A content, in my semantic sense, is always relative to an agent and to a fixed current world, though not for this reason private or unrepeatable. Its world representation depends on the mode of the state in Crane's sense. A world representation is never alone capable of representing intentional objects. To this end, object representations are needed, as well. (TULENHEIMO, 2017, p. 151)

A la différence de Crane, Tulenheimo considère que c'est l'unité modale et non la représentation phénoménale qui particularise les représentations concernant un même objet. La représentation est considérée comme une abstraction, un type, et non une occurrence psychologique primitive :

> When Crane speaks of representations, his comments are meant to apply to concrete phenomenological contents. [...] Representations, in my sense, are the result of a theoretical analysis of types of representations in Crane's sense. Unlike the latter representations, the former are abstract, and they are not psychologically primary. Such abstractions are unavoidable : as a matter of fact, already the notion of two Cranean representations having the same object requires moving away from the concrete reality of an experiential episode. (TULENHEIMO, 2017, p. 151)

L'attribution d'une propriété à un objet suit deux modes différents. Dans le cas de la prédication d'un objet physique, ce n'est qu'à actuel objet qu'est associée la propriété afférente. Mais dans le cas d'un objet purement intentionnel, ce n'est évidemment plus à un particulier donné, mais à toute une ligne de mondes couvrant les mondes épistémiques compatibles avec l'expérience de l'agent qu'est attribuée la propriété :

> Intentional objects are by their nature defined over a set of scenarios compatible with an agent's intentional state. Those states are typically compatible with a variety of ways in which the world could be. This is why the modal margin of an intentional object normally contains a

> number of worlds. Talking about intentional objects makes sense only relative to such a margin, dependent the agent and the world in which she is located. The unity to be characterized when characterizing an intentional object is the totality of its realizations. (TULENHEIMO, 2017, p. 86)

Contrairement aux objets physiques qui ne nécessitent pas systématiquement le recours à un autre monde pour parler d'eux, lorsque nous évoquons des objets seulement intentionnels, nous charrions en fait toute une série de mondes compatibles avec les pensées de l'agent. L'individuation des objets purement intentionnels est donc toujours bâtie sur tout un ensemble de possibilités reliées entre elles par les connaissances et les croyances de celui qui y pense. Si l'on considère que cette théorie des objets intentionnels peut s'appliquer aux personnages de fiction, il apparaît, comme dans la sémantique lewisienne, qu'elle règle le problème de l'indétermination, chaque entité monde-ancrée étant complète. La théorie des lignes de monde occasionne ainsi la substitution des *ficta*, objets fictionnels particuliers mais exotiques car indéterminés, en une pluralité d'objets tout à fait déterminés et reliés par des conditions épistémiques relative à l'agent (voire à différents agents).

La théorie des lignes de mondes encourage ainsi l'idée que l'identification transfictionnelle ne consiste pas en la reconnaissance d'une identité métaphysique, mais en la connexion mentale de concepts nominatifs ressortissant à des individus possibles situés dans des domaines de mondes différents, cette identification étant rapportée à un agent, ou à un ensemble d'agents.

L'approche de l'identification par des lignes de mondes est conforme à une conception selon laquelle les individus sont au moins partiellement construits par l'agent qui les appréhende. En reconnaissant leur rôle de fonctions d'indivudation, les lignes de mondes permettent un rapprochement avec la notion de duplique fonctionnelle introduite dans le chapitre précédent.

Comparaison entre les deux théories

La théorie des dossier mentaux est sans doute moins souple que celle des lignes de mondes quant aux identifications imaginaires contrefactuelles. Admettons que Lucy ne sache pas que Samuel Clemens est Mark Twain et qu'elle croie que Twain est célèbre et que ce n'est pas le cas de Clemens. Bien sûr, il s'agit seulement là d'un banal cas frégéen qui implique deux modes distincts de présentation du même individu. Selon la théorie des dossiers mentaux, Lucy s'est constitué deux dossiers mentaux au sujet du même individu parce qu'elle ne sait pas qu'elle a en réalité affaire à une seule et même personne. Dans le dossier Twain figure la mention « célèbre », alors que le dossier Clemens contient la mention « pas célèbre ». Dilip Ninan (NINAN, 2015) propose alors de considérer le cas suivant. Admettons maintenant que Lucy imagine le contraire, autrement dit, que Clemens soit célèbre et pas Twain. Nous devrions donc mobiliser à nouveau deux dossiers mentaux. Mais comment le dossier Clemens de Lucy pourrait-il capturer son état imagination que Clemens est célèbre, si le dossier contient l'attribution d'être non célèbre ? Une profonde asymétrie semble donc éloigner les attitudes de croyance et les attitudes de croyance-feinte.

Ninan suggère que la théorie des dossiers mentaux peut sans doute répondre à cette difficulté à condition de subdiviser l'esprit en étagères distinctes de dossiers [25]. Il y aurait alors une étagère

[25] En réalité, Ninan parle de « boites » (*box*).

spécialisée pour les dossiers imaginaires et une étagère spécialisée pour les dossiers de croyances, quitte à ce que des dossiers présents sur des étagères différentes puissent avoir des relations entre eux. Recanati (RECANATI, 2015) rétorque qu'une telle conception est bien possible, mais qu'elle implique qu'on ne peut plus considérer les dossiers mentaux comme des représentations dont la fonction première est de véhiculer des informations à propos des objets.

La sémantique initiée par Hintikka et développée par Tulenheimo est en revanche ici plus souple, parce qu'elle permet de diversifier les fonctions d'individuation par le type de relations reliant les mondes possibles. Un certain opérateur permet de relier des objets entre certains mondes possible selon une ligne donnée, et un autre opérateur permettra d'en relier d'autres.

Tout comme la théorie des lignes de mondes, la théorie des dossiers mentaux a l'avantage de ne pas produire de doubles ontologiques par identités croisées ou par contreparties. Mais la première semble plus adaptée aux situations imaginaires. Nous adopterons donc, dans ce qui suit, la théorie des lignes de mondes, en montrant dans la section suivante qu'elle peut s'employer de manière profitable dans le cas de pensées contre-fictionnelles, en évitant de stipuler des mondes impossibles dans certains contextes d'identité imaginaires contradictoires.



L'enjeu que pose le double métaleptique est finalement du même genre que celui de l'identité croisée, à savoir, l'idée d'une entité qui existerait dans plusieurs mondes possibles. Mais plutôt que de reconnaître des chevauchements de mondes possibles considérés comme des entités ontologiques, on peut, ici encore, faire intervenir la notion de ligne de mondes pour identifier un individu en liant, de manière toute épistémique, diverses occurrences monde-ancrées dans différents mondes possibles. De ce fait, la conception de doubles métaleptiques n'a finalement plus rien de contradictoire.

Nous pouvons résumer notre analyse en affirmant que les doubles fictionnels ne se rapportent pas nécessairement à la contradiction logique : nous pouvons penser aux doubles de manière rationnelle. Il est vrai que nous avons, pour l'instant, mis de côté le cas d'un récit de Maupassant, *Lui !*, où le double prend un aspect encore plus pathologique. Nous allons voir maintenant, dans la dernière section de ce chapitre, une manière de le traiter sans céder à l'absurde.

5.4 Logiques de l'impossible

5.4.1 Un exemple littéraire

Dans la nouvelle intitulée *Lui ?*, Maupassant (MAUPASSANT, 1979b) relate l'angoisse d'un personnage isolé mais néanmoins troublé par le sentiment d'une vague présence étrangère. L'auteur conclut son récit par les mots suivants : « Car il est là parce que je suis seul, uniquement parce que je suis seul ! ». La proposition exprimée par cette phrase semble inciter le lecteur à imaginer que le personnage est à la fois seul et accompagné, autrement dit, qu'il est *seul*, à *deux*. Cet exemple

montre que l'imagination, stimulée par la fiction, outrepasse parfois le domaine du simplement possible pour côtoyer le champ du difficilement concevable [26].

Jean Bessière fait remarquer que le double met en exergue la fonction remplie par la littérature fantastique, qui n'est pas seulement un jeu avec l'imaginaire, mais plus spécifiquement la « figuration des conditions de la construction du jeu représentationnel. Le thème du double est ainsi un moyen exemplaire d'exposer ce que fait un récit littéraire dans le monde, au regard du monde. » (BESSIÈRE, 2018, p. 17). Selon Bessière , le fantastique en littérature, et dans les autres arts, n'a pour vocation une exploration débridée de l'imagination, mais la mise en avant de « l'impossibilité de la représentation »(BESSIÈRE, 2018, p. 19). Le double se caractérise principalement par le fait d'être un « impossible » (BESSIÈRE, 2018, p. 17). En particulier, Bessière reprend le concept de *doubling*, que Wolfgang Iser utilise pour rompre la tradition de critique basée sur l'opposition entre réalité et fiction :

> Instead of harking back to the old fiction/reality dichotomy, I propose at the outset to conceive of the fictive as an operational mode of consciousness that makes inroads into existing versions of the world. In this way, the fictive becomes an act of boundary-crossing which, nonetheless, keeps in view what has been over stepped. As a result, the fictive simultaneously disrupts and doubles the referential world. (ISER, 1993, p. xv)

De son côté Roger Caillois affirme que le fantastique a pour objet de provoquer une sorte de panique mentale, liée à l'affranchissement des lois logiques. La principale caractéristique du fantastique, c'est qu'il émerge subitement dans l'histoire :

> Le fantastique, [...], manifeste un scandale, une déchirure, une irruption insolite, presque insupportable dans le monde réel. (CAILLOIS, 1966, p. 15)

Mais de quelle sorte d'irruption s'agit-il ? Il s'agit de l'émergence subite de l'inconcevable dans le monde réel. Si le féérique s'écarte des lois physiques du monde réel, il demeure logiquement possible, ou comme l'écrit Caillois, « Le féerique est un univers merveilleux qui s'ajoute au monde réel sans lui porter atteinte ni en détruire la cohérence » (CAILLOIS, 1966, p. 15). Ce n'est pas le cas du fantastique qui semble justement s'affranchir de la possibilité logique :

> [...] dans le fantastique, le surnaturel apparaît comme une rupture de la cohérence universelle. [...] Il est l'impossible, survenant à l'improviste dans un monde où l'impossible est banni par définition. (CAILLOIS, 1966, p. 17)

Pour que la « déchirure » soit possible, écrit Caillois, il faut qu'elle ait lieu dans un monde similaire au nôtre. Point de fantastique dans un monde merveilleux, « le cadre du fantastique n'est pas la forêt enchantée de la Belle au Bois Dormant, mais le morne univers administratif de la société contemporaine. Transposé au Moyen Âge ou dans l'antiquité, un récit fantastique y perdrait de sa puissance. C'est que le surnaturel y est, si je puis dire, plus naturel. » (CAILLOIS, 1966, p. 21).

[26] Il pourrait être objecté que la phrase de Maupassant n'a rien de contradictoire, en ce qu'elle dirait simplement la chose suivante : « Il est là quand je ne suis pas accompagné de quelqu'un d'autre (que lui) ». Mais la nouvelle privilégie une interprétation fantastique que ne recouvrirait pas suffisamment cette traduction : le thème évoqué par Maupassant est évidemment celui du double, cette nouvelle anticipant celle, plus connue, intitulée *Le Horla.*

Selon Marie-Laure Ryan, les cas de personnages récurrents par effet de boucles temporelles ou par effet métaleptique montrent que :

> l'adhérence stricte à la logique est une condition trop draconienne en ce qui concerne les mondes fictionnels. Il existe des textes qui enfreignent la logique mais qui présentent tout de même des mondes imaginables. Parmi ces textes : les récits de voyage dans le temps qui créent des boucles causales ; la métalepse, qui permet aux personnages de transgresser les frontières ontologiques et d'accumuler des propriétés contradictoires (à la fois fictionnelles et réelles) ; les mondes où le passé est soudainement changé ; ou encore les mondes où l'espace est impossible. On peut donc distinguer trois types de textes : (1) Textes à mondes logiquement cohérents. L'équivalent en peinture serait une œuvre qui représente un sujet de manière relativement réaliste. (2) Textes à mondes dit "de fromage suisse" : il y a des gouffres où la logique est transgressée, mais ces gouffres sont clairement délimités, comme les trous dans le fromage, et le lecteur peut appliquer le principe de l'écart minimal pour les régions du texte qui correspondent au fromage. C'est l'équivalent littéraire des peintures d'Escher ou de Magritte. (3) Textes sans monde, qui sont l'équivalent de peintures abstraites : poésie concrète ; textes bâtis systématiquement sur le gommage et la contradiction ; textes où les objets décrits se métamorphosent continuellement. (RYAN, 2006)

Comment une logique de la fiction peut-elle répondre à ces difficultés ? Et s'agirait-il encore d'une logique ? Peut-il y avoir des logiques réellement différentes de la logique classique, admettant une affirmation mais aussi sa négation ? Les mondes fictionnels, en particulier ceux qui se rapportent au fantastique, sont-ils des mondes où règne une logique très différente, voire, qui ne sont pas du tout régis par la logique ? Quel traitement sémantique donner au jugement trouble, d'étrangeté inquiétante, ou comme le qualifie Sigmund Freud de sentiment *unheimlich* (FREUD, 1985c) ressortissant de la situation paradoxale où est placé le personnage de Maupassant ? Il semble que la nouvelle soit paradigmatique des récits fantastiques basés sur le thème du double. Le fonctionnement général des logiques d'imagination basées sur la sémantique des mondes possibles se heurte en particulier au problème de l'interprétation à donner au texte de Maupassant. Certaines logiques, dites paraconsistantes ou encore dialéthéistes, proposent de formaliser nos attitudes imaginatives nonobstant le rapport à la contradiction qu'elles peuvent parfois occasionner. Des logiciens (Graham Priest (PRIEST, 2005), Francesco Berto (BERTO, 2017)) ont étendu les structures modales normales de Kripke en les complétant par l'extension de mondes impossibles. Leurs systèmes formels admettent des contradictions tout en rejetant le principe d'explosion de la logique classique. Cette formalisation se fait néanmoins au coût d'un réalisme d'objets abstraits tout droit sortis des profondeurs de la jungle meingonienne.

Dans ce qui suit, j'indique cependant des raisons montrant que ce type d'explications métaphysiques de nos actes d'imagination manque sa nature essentiellement dynamique et pragmatique. En particulier, je soulignerai encore une fois l'importance interprétative du lecteur dont témoigne, dans l'acte de lecture, la caractéristique d'hésitation déjà indiquée par Tzvetan Todorov dans son *Introduction à la littérature fantastique*.

5.4.2 Concevoir des scénarios alternatifs

Imagination et possibilités

Traditionnellement, l'imagination est définie par les philosophes comme une disposition à rendre présent à l'esprit quelque chose d'absent, que cette chose soit réelle ou non. On peut la caractériser par la capacité de susciter une image mentale d'une chose en l'absence de cette chose. Il ne faut toutefois pas restreindre l'imagination aux seules images visuelles, puisqu'elle relève aussi des autres sens : on peut « imaginer » une mélodie, une odeur, une saveur, un contact tactile. Puisqu'il s'agit de susciter quelque chose en son absence, l'imagination a donc autant rapport à la mémoire qu'à la perception. Certains philosophes ont d'ailleurs assujetti l'imagination à la première ou à la deuxième, réduisant les actes d'imagination aux souvenirs (c'est le cas de Bergson (BERGSON, 1946)), ou à des perceptions appauvries (c'est le cas de Hume (HUME, 1968b)).

Néanmoins, on peut aussi conférer à l'imagination un rôle plus autonome à l'égard de la mémoire et des perceptions, comme l'a d'ailleurs indiqué Sartre en 1936 dans *L'imagination* (SARTRE, 2012) puis en 1940 dans *L'imaginaire* (SARTRE, 1940). Elle peut également être considérée comme un ensemble d'actes intentionnels dirigés vers des objets, des configurations d'objets, des situations et des circonstances et plus généralement des *scénarios*. Cette caractérisation de l'imagination implique *de facto* qu'elle est liée à la conception [27] et au possible. David Chalmers souligne que l'imagination d'une situation s'accompagne généralement de nombreux détails, en particulier d'inférences déduites de la situation initiale. Ainsi, l'analyse logique d'une situation imaginée peut montrer que, pour une certaine proposition S, il est le cas que S, autrement dit, que la situation vérifie S, ou plus précisément encore, qu'il est imaginé que S. Chalmers appelle *conception positive* la capacité d'imaginer que S est le cas à partir d'une situation donnée :

> to positively conceive of a situation is to in some sense imagine a specific configuration of objects and properties. It is common to imagine situations in considerable detail, and this imagination is often accompanied by interpretation and reasoning. When one imagines a situation and reasons about it, the object of one's imagination is often revealed as a situation in which S is this case, for some S. When this is the case, we can say that the imagined situation verifies S, and that one has imagined that S. Overall, we can say that S is positively conceivable when one can imagine that S : that is, when one can imagine a situation that verifies S. (D. CHALMERS, 2002, p. 5)

Par ailleurs, l'examen de nos actes d'imagination montre qu'ils sont réglés par une manifeste teneur logique. C'est un fait que nous inférons certaines conséquences à partir des données exprimées dans le récit qui est lu. Et c'est aussi un fait que nous n'imaginons pas n'importe quoi à partir du texte. Ainsi, et plus généralement, l'étude de l'imagination relève d'une analyse des relations entre un *contenu explicite*, ce qui raconté, dit dans le texte, représenté sur l'image ou dans le film, et un *contenu imaginé*. En particulier, lorsque nous imaginons un scénario à partir d'un *contenu explicite*,

- nous tirons des inférences qui ne sont pas complètement anarchiques ;

[27] On distinguera la conception de la concevabilité, définie comme possibilité de conception.

- nous ne tirons pas toutes les conséquences dérivables du contenu explicite ;
- nous imaginons également des choses qui ne sont pas logiquement dérivables du contenu explicite ;
- et nous pouvons imaginer-concevoir des choses métaphysiquement ou physiquement impossibles.

La logique classique (disons celle de Frege-Russell) est peu disposée à endosser une analyse des actes d'imagination pour les raisons que :

1. du point de vue sémantique, elle considère que les expressions contenant des noms fictionnels sont dénuées de vérité (Frege) ou fausses (Russell), mais cela va à l'encontre de notre intuition que la proposition exprimée par « Sherlock Holmes est un détective londonien » peut être imaginée comme vraie ;
2. du point de vue logique, elle obéit au principe d'explosion : de toute contradiction on peut déduire n'importe quoi.

Comment échapper à l'explosion lorsqu'on souhaite analyser le caractère logique des attitudes d'imagination en dépit de leur rapport à la contradiction ? Comment rendre compte, de manière logique, de notre attitude imaginative à l'égard du récit de Maupassant ?

Le caractère imitatif des mondes imaginaires est, en grande partie, logique. Les mondes fictionnels sont des mondes clos sous l'implication. Dans le cas idéal d'une fiction dont le narrateur est fiable et qui ne comporte pas de contradictions, si Φ et $\Phi \rightarrow \Psi$ sont vraies dans le monde fictionnel, alors il vient que Ψ est également vraie dans ce monde. Bien sûr, le lecteur n'est pas contraint *d'imaginer* que Ψ, voire, il peut tout à fait ignorer que Ψ. Par exemple, il est évident qu'un monde fictionnel dans lequel les axiomes de l'arithmétique sont vrais n'implique pas que le lecteur en connaisse toutes les propriétés, ni même qu'il doive imaginer celles qu'il connaît. L'imagination est contextualiste. Dans le cas d'une intrigue policière, le lecteur pourrait mobiliser des raisonnements logiques pour anticiper les résultats de l'enquête. Mais dans d'autres formes de fiction, il pourrait ne pas mettre en activité ses capacités logiques, ou du moins, pas aussi intensément.

C'est notamment la façon dont procède John Woods dans *Logic of Fiction*. Il utilise le critère de vérité auctorial (*author's sayso*) consistant à affirmer que dans le cadre d'une fiction, ce qui est vrai est ce qui est dit par l'auteur. Or, comme il le souligne, en fiction, « tout est permis (WOODS, 1974) ». Nous ne pouvons brider, sinon arbitrairement, le principe de stipulation auctoriale.

Constatons cependant qu'il arrive aux auteurs d'écrire des choses incohérentes, de manière volontaire ou non. Deux sortes de contradictions peuvent surgir. La première forme concerne des assertions qui sont inconsistantes de manière internes, autrement dit des énoncés auctoriaux contradictoires, comme c'est le cas par exemple dans *Anna Karenine* que Tolstoï fait débuter à la fois un lundi et un mardi. Un écrivain pourrait écrire une histoire dans laquelle son personnage pourrait résoudre la quadrature du cercle.

Le deuxième type d'incohérences relève des contradictions entre le monde actuel et le monde de la fiction, comme par exemple le fait qu'il est dit que Sherlock Holmes habite au 221B Baker Street à Londres, alors que personne n'habite réellement cette adresse [28].

Les exemples qui précèdent concernent des incohérences circonscrites. Woods imagine le cas paroxystique d'un auteur qui affirmerait le contraire de tout ce qui est vrai dans le monde réel. Comment concilier le critère d'énonciation auctoriale et la valeur de vérité des énoncés fictionnels ? Selon Woods, il convient de considérer que la stipulation auctoriale n'est pas une condition de vérité, mais une condition d'assertabilité :

> Since more things are correctly assertable than are true, why not reconstrue auctorial say-so as a condition on assertability, not truth ? (WOODS, 2002, p. 210.)

Ainsi, Doyle individue son personnage non par des vérités, mais bien par des mensonges. Dès lors, il devient nécessaire de comprendre comment on accède à une théorie du vrai par des mensonges. Cela est d'ailleurs une pratique que l'on retrouve en physique (par exemple quand on néglige les forces de frottement).

Comme le souligne Woods, les objets fictionnels sont des objets *nécessairement possiblement contradictoires*. En effet, puisque la notion de vérité en fiction repose sur des actes de stipulations auctoriales, il est toujours possible qu'un auteur attribue à son personnage une propriété et, dans le même temps, la négation de cette propriété. Tout objet possiblement impossible est un objet impossible. Autrement dit, les personnages de fiction sont (nécessairement) des impossibles.

Concevoir l'impossible

De nombreux philosophes ont souscrit au principe de concevabilité, énoncé par Hume, stipulant que si une chose est concevable, alors elle est également possible, ou sous sa forme contraposée, que si une chose est impossible, alors elle n'est pas concevable : « This is an established maxim in metaphysics, that whatever the mind clearly conceives includes the idea of possible existence, or in other words, that nothing we imagine is absolutely impossible. » (HUME, 1968a, Livre I, partie II, section 2). Ce principe réduit donc le concevable au possible : ce qui est concevable est aussi possible et l'impossible est hors d'atteinte de la représentation. Suivant en cela Hume, certains auteurs [29] affirment que l'imagination est une forme particulière de conception qui elle-même est réductible au possible (figure 5.3).

Cette thèse, qui conditionne l'épistémique au métaphysiquement possible, a été contestée de multiples façons. Car de fait, on peut se représenter, et parfois même percevoir des figures impossibles, comme le triangle de Reutersvard-Penrose, ou la cascade de Maurits Cornelis Escher (Waterfall, 1961). On peut admettre une hypothèse et en déduire par l'absurde qu'elle n'est pas valide.

[28] Cet exemple, maintes fois cité dans la littérature consacrée, engage cependant des difficultés particulières. Premièrement, le numéro 221B n'existait pas au temps de Conan Doyle. Deuxièmement, au XXe, le culte voué à Holmes s'est développé au point de fournir un véritable logis au personnage, lui consacrant un musée-boutique, situé au 235 Baker Street. Mais cette adresse a été rebaptisée 221B. Aujourd'hui, le numéro existe donc à deux endroits différents de la rue de Londres. À ce sujet, on pourra consulter l'entrée « 221B Baker Street » de l'encyclopédie en ligne Wikipedia.

[29] Parmi les travaux relativement récents en logique de la fiction, on peut citer Ilkka Niiniluoto (NIINILUOTO, 1985), Alexandre Costa-Leite (COSTA-LEITE, 2010) et Heinrich Wansing (WANSING, 2017) qui réduisent tous le concevable au possible

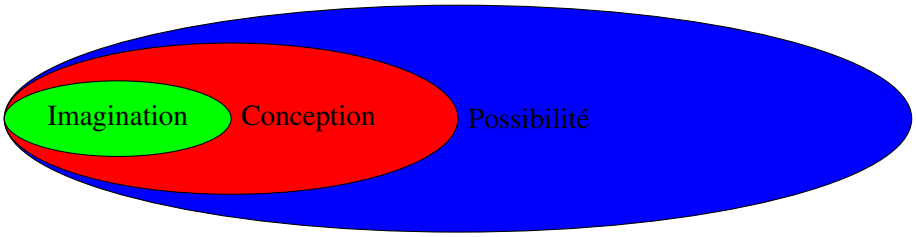


FIGURE 5.3 : Interprétation classique des relations entre possibilité, conception et possibilité

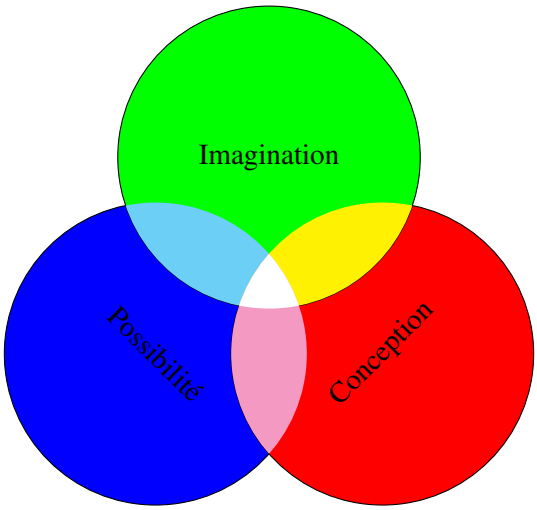
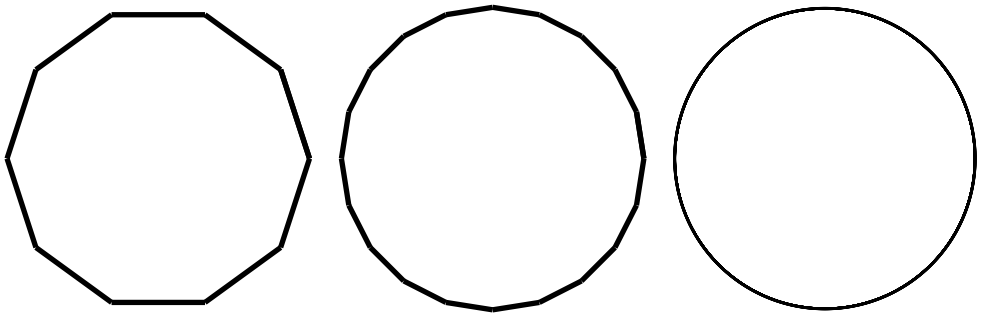


FIGURE 5.4 : Relations entre possibilité, conception et possibilité (d'après Jean-Yves Beziau)

Enfin, on peut encore citer les identités *a posteriori* mises en exergue par Kripke affirmant qu'il est (métaphysiquement) impossible que l'eau ne soit pas H_2O même si cela est néanmoins concevable [30]. A cet égard, Jean-Yves Beziau (BÉZIAU, 2016) a représenté les relations réciproques entre imagination, conception et possibilité par un diagramme de Venn que nous reproduisons sur la figure 5.4.

Contre le principe de concevabilité, ce diagramme illustre non plus trois mais sept éventualités dont Beziau a déterminé pour chacune d'entre elles des exemples de choses y répondant. Parmi les zones coloriées, la blanche, située à l'intersection des trois cercles, est manifestement habitée. On peut par exemple citer des choses à la fois possibles, concevables et imaginables comme les camemberts et les triangles. Pour ce qui est de la zone magenta des choses concevables et possibles mais pas imaginables, elle est souvent illustrée par le célèbre exemple du chiliogone de Descartes (DESCARTES, 1953c), polygone à mille côtés dont il est prétendu qu'il ne peut être imaginé.

[30] Pour une discussion sur les relations entre possibilité et concevabilité, on pourra se référer à David Chalmers (D. CHALMERS, 2002), Graham Priest (PRIEST, 2005), Ruth Byrne (R. BYRNE, 2005), Alex Byrne (A. BYRNE, 2007), Oreste Fiocco (FIOCCO, 2007), Mark Jago (JAGO, 2014), Peter Kung (KUNG, 2010), Francesco Berto (BERTO & SCHOONEN, 2018).



Décagone régulier Icosagone chiliogone régulier

FIGURE 5.5 : Du décagone régulier au chiliogone régulier

L'exemple me parait manifestement très discutable, puisqu'on peut bien imaginer un cercle, ce qui donne une très bonne image du chiliogone régulier, et qu'une construction de ce polygone, dans sa forme régulière, est extrêmement aisée par programmation 5.5. Néanmoins, d'autres exemples semblent bel et bien témoigner de l'occupation de la zone magenta, comme, mettons, la bouteille de Klein [31] ou les espaces vectoriels à 12 dimensions.

La figure 5.6 pourrait illustrer un habitant de la zone verte : objet imaginable mais ni concevable, ni possible. Remarquons que les zones dont l'occupation pose réellement question sont plutôt la bleue et la rouge. Précisément, on peut réellement se demander s'il y a des objets dans la zone bleue, autrement dit des objets seulement possibles mais ni concevables ni imaginables. Beziau cite l'exemple de l'univers comme somme totale de tout ce qui est [32].

5.4.3 Mondes étranges et mondes merveilleux

Dans quelle zone placer la situation exposée par Maupassant ? De fait, le récit semble témoigner d'une impossibilité plus contraignante que celles, nomologiques, des récits de science-fiction consistant à imaginer des voyages à des vitesses supra-luminique ou des récits fabuleux qui mettent en scène des dragons et des transformations d'êtres humains. Quelles alternatives avons-nous pour entreprendre une interprétation logique d'une situation paradoxale comme celle exposée par Maupassant ?

Dans son étude sur la littérature fantastique, Tzvetan Todorov (TODOROV, 1970) a distingué les genres de l'étrange et du merveilleux. Le premier est caractérisé par la présence, dans le récit, d'éléments étranges, mais qui se ramènent en bout de course à une explication rationnelle et compatible avec ce que nous savons du monde réel (ce que Todorov appelle une explication *naturelle*).

[31] Conférer l'excellent site *Mathcurve* pour une description de la bouteille de Klein et une représentation paramétrée d'une projection de celle-ci en 3D : https://www.mathcurve.com/surfaces/klein/klein.shtml

[32] C'est peut-être aussi la perspective de Clément Rosset (ROSSET, 1997) lorsqu'il définit le Réel comme quelque chose qui ne peut être ni conçu ni imaginé, parce que la seule manière de concevoir le réel est entièrement réduit à la pure tautologie : le réel, c'est le réel. En outre, notons que l'idée d'une totalité conçue comme un objet nécessite une caractérisation méréologique puisque d'un point de vue purement ensembliste, l'axiomatique de Zermelo-Fraenkel rejette l'existence d'un ensemble de tous les ensembles. En revanche, on peut toujours définir la somme méréologique de tous les objets et elle constitue elle-même un objet.

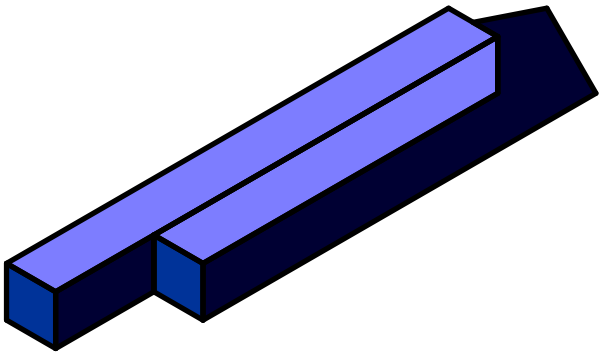


FIGURE 5.6 : Une ou deux poutres ?

En revanche, un récit dans lequel apparaissent des éléments absolument incompatibles avec les lois de la physique, de la chimie et de la biologie (contenant par exemple des métamorphoses, des animaux fabuleux ou des voyages supraluminiques) appartient au genre du merveilleux (explication *surnaturelle*). Précisons que le genre du merveilleux est donc physiquement, voire métaphysiquement impossible, mais pas logiquement impossible. Dans ce qui suit, j'étends les deux adjectifs du genre aux mondes possibles, autrement dit, je nommerai monde étrange un monde (nomologiquement) possible interprétant une lecture du récit relevant du genre de l'étrange, et monde merveilleux un monde (logiquement) possible interprétant une lecture relevant du genre merveilleux.

Comment surmonter la difficulté posée par la contradiction qui ressort de la situation où est placé le personnage de Maupassant ? Il pourrait être prétendu qu'il est tout simplement fou. En admettant exclusivement cette possibilité, on rangerait ainsi la nouvelle dans le registre de la littérature réaliste, sinon naturaliste : il ne s'agirait ni plus ni moins que la description des affres d'un esprit dont la perte de raison mène à des points de vue plutôt étranges. Cette hypothèse est d'ailleurs au centre de bon nombre d'interprétations psychanalytiques et psychiatriques. Celles-ci analysent généralement les trois nouvelles de Maupassant les plus fortement liées au thème du double (à savoir *Lui*, *Le Horla* et *Qui sait ?*) comme la minutieuse chronique d'hallucinations autoscopiques dont souffrait l'auteur. Ainsi, à propos du *Horla*, Johann Jung se réfère à la notion de double transitionnel pour expliquer la détresse du personnage de Maupassant qui ne serait que l'évocation des propres souffrances de l'auteur : « nous pouvons faire l'hypothèse que le double constitue l'interface à partir de laquelle se constitue l'identité psychique à la limite du dedans et du dehors. » (JUNG, 2010). L'incapacité où serait placée le personnage de céder au double transitionnel le mènerait à sa propre perte. Mais comme l'écrit justement Louis Vax, « personne ne confondra littérature fantastique et littérature psychiatrique. (...) Le médecin cherche à apaiser son malade, le conteur veut troubler son lecteur. » (VAX, 1963).

Une deuxième idée consiste à considérer à l'inverse que le personnage n'est pas fou, et donc qu'il est accompagné dans sa solitude, sans pour autant admettre de contradiction logique. Si cela semble assez paradoxal, c'est pourtant ce que proposent certains philosophes pluralistes qui

distinguent l'identité d'une chose de sa constitution matérielle. Depuis la théorie essentialiste de la dépendance sortale développée par David Wiggins dans *Sameness and Substance* (WIGGINS, 1980), des philosophes [33] distinguent l'existence d'une statue de l'existence du bloc de matière qui la constitue. Cette distinction peut être déduite de différences historiques lorsque la carrière de la statue ne coïncide que temporairement avec celle du bloc de matière, par exemple quand un même bloc de matière est remodelé en une statue différente. Elle peut aussi être déduite de différences modales lorsque leur coïncidence est permanente : une statue pourrait subir la perte d'un bras contrairement au bloc qui la constitue. Ces différences historiques et modales impliquent alors, par contraposée du principe d'indiscernabilité des identiques, qu'il y a deux choses et non une : d'une part une statue et d'autre part un bloc de matière. L'identité des statues, mais plus généralement aussi des artefacts, des organismes et des personnes, est ainsi désolidarisée de leur constitution matérielle. Aussi, là où le langage naturel nous convie habituellement à ne voir qu'une seule chose, il se tiendrait en réalité deux choses concoïncidentes. Cette thèse suppose de rejeter le principe métaphysique d'identité des coïncidents défendu par John Locke (LOCKE, 2001), stipulant que deux choses ne peuvent se tenir simultanément au même endroit. On pourrait ainsi admettre que le personnage de Maupassant est non seulement sain d'esprit, mais effectivement dédoublé par un coïncident, sans contradiction logique. Dans ce cas de figure, le monde du personnage de Maupassant serait possible au sens logique du terme, mais il serait néanmoins un monde merveilleux, autrement dit, un monde ne répondant plus aux conditions nomologiques que prescrit la physique du monde actuel. En outre, un tel monde ne nécessiterait pas l'admission de contradictions logiques. Cette solution consisterait donc à rabattre le récit sur le genre du merveilleux.

Ces deux hypothèses sont évidemment contradictoires, mais il appert qu'opter seulement pour l'une ou pour l'autre impliquerait d'appauvrir la dimension interprétative du récit fantastique. Il nous faut donc conserver les deux options nonobstant leur aspect paradoxal. C'est d'ailleurs, à en croire Montague Rhodes James, propre au registre du fantastique dans lequel « It is not amiss sometimes to leave a loophole for a natural explanation ; but, I would say, let the loophole be so narrow as not to be practicable. » (M. R. JAMES, 1924). Il serait donc souhaitable de conserver, pour l'instant, l'aspect contradictoire du récit. On se heurte alors d'emblée au principe d'explosion logique, ou comme le disaient les logiciens médiévaux, *ex falso sequitur quodlibet* : de la fausseté résulte n'importe quoi.

Qu'en est-il de cette déduction dans le langage de l'opérateur d'imagination Ⓡ ? Il prend la forme suivante :

(1')	$Ⓡ\varphi \wedge Ⓡ\neg\varphi$	par hypothèse
(2')	$Ⓡ\varphi$	par simplification de la conjonction dans (1')
(3')	$Ⓡ\neg\varphi$	par simplification de la conjonction dans (1')
(4')	$\neg Ⓡ\varphi$	par (3') et par l'axiome de consistance pour Ⓡ
(5')	$Ⓡ\varphi \vee Ⓡ\psi$	par (2') et disjonction
(6')	$Ⓡ\psi$	par (4') et (5') et règle de résolution pour la disjonction

Imaginer des choses contradictoires impliquerait donc d'imaginer n'importe quoi. Mais cela ne

[33] Parmi eux, citons Lyne Ruder Baker (BAKER, 1997), Kit Fine (FINE, 2000, 2003), Mark Jonhston (JOHNSTON, 1992).

semble pas compatible avec la part de rationalité de l'imagination car lorsqu'on imagine l'état de *solitude accompagnée* du personnage de Maupassant, il n'en découle pourtant pas qu'on imagine que la lune est habitée de licornes broutant des carrés-ronds. Les actes d'imagination ne sont pas complètement anarchiques même s'ils peuvent contenir des éléments irrationnels. Comment éviter alors la dérivation explosive ?

5.4.4 Contradictions non-adjonctives

Si on souhaite offrir un traitement curatif au schéma déductif précédent, il s'agit alors d'en bloquer le mécanisme en rejetant la donnée de départ ou l'une des conséquences. On pourrait par exemple refuser l'implication entre $\text{®}\neg\varphi$ et $\neg\text{®}\varphi$.

Une autre tentative de résolution consisterait à bloquer le schéma menant à l'explosion logique dès le départ, en objectant que $\text{®}\varphi \wedge \text{®}\neg\varphi$ est une traduction impropre de ce qui est prescrit par le récit de Maupassant. Pour rendre compte de l'analyse logique de cette stratégie, il faut dès lors rejeter le principe d'adjonction par rapport à l'opérateur d'imagination : du fait d'imaginer de manière contradictoire $\text{®}\varphi, \text{®}\neg\varphi$ on ne doit pas inférer de contradiction vraie $\text{®}\varphi \wedge \neg\text{®}\varphi$. Comment procéder ? Nicholas Reicher et Robert Brandom (RESCHER & BRANDOM, 1980) ont développé une logique paraconsistante basée sur l'intersection ou la réunion de mondes possibles. La réunion de deux mondes possibles constitue un troisième monde qui peut être localement inconsistant [34]. Admettons par exemple que w_1 soit un monde possible (du genre étrange) dans lequel le personnage de Maupassant est seul (propriété que nous noterons ici « s »). Admettons également un deuxième monde possible w_2 (du genre merveilleux) dans lequel le personnage n'est pas seul. Considérons le monde w constitué par ce que les auteurs nomment « processus de superposition », à savoir, la réunion de w_1 et w_2 : $w = w_1 \bigcup w_2$. Dans un tel monde w, il est le cas que φ si et seulement s'il est le cas que φ dans w_1 ou dans w_2. Ainsi, dans w, il est à la fois le cas que le personnage est seul et qu'il n'est pas seul. Afin d'éviter l'explosion logique, les deux auteurs ont modifié l'interprétation sémantique de l'opérateur de conjonction en rejetant la condition nécessaire (principe d'adjonction) prescrite par la règle classique. Dès lors, la vérité de s et la vérité de $\neg$s n'impliquent plus celle de s $\wedge$ $\neg$s. Dans le monde w, il est à la fois vrai que le personnage est seul et à la fois vrai que le personnage n'est pas seul. En revanche, il n'est pas vrai que le personnage est seul et n'est pas seul. Il appert donc que les mondes construits ainsi ne sont pas adjonctifs : $\varphi, \psi \not\models \varphi \wedge \psi$. Ces mondes sont donc soumis à la clôture logique pour les implications aux prémisses atomiques, mais ont un comportement anarchique et ouvert pour les implications à prémisses multiples [35].

Cette construction parait néanmoins assez artificielle, dans la mesure ou elle emboîte des mondes possibles. Par ailleurs, le principe d'adjonction semble bien utilisé dans nos actes d'imagination : lorsqu'il est vrai que nous imaginons que Sherlock Holmes est détective et lorsqu'il est vrai que nous imaginons que Sherlock Holmes habite Londres, il semble indéniable qu'il soit vrai que nous imaginions aussi que Sherlock Holmes soit détective et habite Londres. Le rejet du

[34] Les deux auteurs suivent en cela la tradition liée à la logique discursive (*discussive logic*, qualifiée aussi parfois *discursive logic*) initiée par Stanislaw Jaśkowski (JAŚKOWSKI, 1948).

[35] Les mêmes auteurs ont également développé un modèle de mondes localement incomplets par le processus de schématisation de mondes possibles : $w' = w_1 \bigcap w_2$. Pour reprendre l'exemple précédent, dans w', il n'est pas le cas que s car s est fausse dans w_2, et il n'est pas non plus le cas que $\neg$s car $\neg$s est fausse dans w_1.

principe d'adjonction semble donc être un coût trop important.

On peut tenter de donner un autre traitement à la distinction entre l'acte d'imaginer que le personnage est seul et d'imaginer qu'il est accompagné (autrement dit $\text{®}s \wedge \text{®}\neg s$) de celui d'imaginer que le personnage est à la fois seul et à deux (autrement dit $\text{®}(s \wedge \neg s)$), sans pour autant invalider de manière générale l'adjonction logique. Alors que la deuxième caractérisation nécessite inévitablement d'imaginer une contradiction, la première peut trouver une interprétation non-contradictoire, à condition de trouver une interprétation moins naïve de l'opérateur ®. De fait, l'expression « seul à deux », inciterait notre imagination à accoler deux situations différentes mettant en scène le même personnage, un peu comme une juxtaposition de deux clichés photographiques présentant la même personne dans des situations différentes. C'est notamment la stratégie qu'indique Heinrich Wansing, dans son étude sur la logique de la fiction, à propos d'une situation invitant à imaginer qu'une personne du nom d'Igmar est simultanément en train de danser et de ne pas danser :

> The mental picture conception, however, supports the distinction between the conflicting imagination we encounter if (i) I imagine that Ingmar is dancing and at the same time imagine that Ingmar is not dancing ($I_a\text{A} \wedge I_a\neg\text{A}$), and (ii) the contradictory imagination that I would have if I imagined that Ingmar is both dancing and not dancing at the same time $I_a(\text{A} \wedge \neg A)$. (WANSING, 2017)

Il s'agit d'ailleurs d'une technique narrative éprouvée par certains peintres de la Renaissance italienne qui font apparaître un même personnage à différents emplacements afin de représenter différentes situations. C'est notamment le cas de la fresque murale intitulée *Le paiement du tribut*, peinte par Masaccio entre 1424 et 1427 dans l'église Santa Maria del Carmine de Florence où l'on peut apercevoir le personnage de Saint-Pierre à deux emplacements différents de la fresque. Mais cette représentation ne demande pas au spectateur de supposer que le personnage ait la capacité de se dédoubler. L'effet contradictoire relève seulement d'une technique visant à produire un certain effet : raconter différentes choses sur le même support. Pour le spectateur, cela nécessite de relativiser la représentation de manière temporelle : à tel moment, Saint-Pierre fait telle chose, à tel autre moment, telle autre. Et il n'y a donc aucune contradiction à dénicher dans le tableau de Masaccio.

C'est cette voie que je souhaite explorer plus avant. Il semble qu'elle ne nécessite pas forcément de distinguer entre conjonction interne et externe à l'opérateur d'imagination en refusant le principe d'adjonction logique. Cette stratégie nécessite en revanche de relativiser les opérateurs intentionnels, non à des moments différents comme le fait Masaccio, mais à des *mondes possibles différents*. Dans ce cas, l'opérateur d'imagination doit être différent selon la perspective choisie : $\text{®}_1\varphi$ ou $\text{®}_2\neg\varphi$.

5.4.5 Hésitation du lecteur et accessibilité à des mondes différents

De ce qui précède, il ressort que la question est de déterminer où placer la contradiction. Dans les données du récit ? Dans l'acte d'imagination du lecteur ? Dans la technique de narration ? Dans un monde impossible ? De fait, le texte de Maupassant ne précise pas si le personnage est fou ou non, autrement dit, si l'on a affaire à un récit relevant du genre de l'étrange ou à celui du merveilleux.

La question reste indéterminée. Cette marge d'incomplétude est d'ailleurs essentielle aux récits fantastiques, que Tzvetan Todorov caractérise justement par l'hésitation du lecteur qui ne sait pas exactement à quel genre obéit le texte :

> Ou bien le diable est une illusion, un être imaginaire ; ou bien il existe réellement, tout comme les autres êtres vivants (...). Le fantastique occupe le temps de cette incertitude ; dès que l'on choisit l'une ou l'autre réponse, on quitte le fantastique pour entrer dans un genre voisin, l'étrange ou le merveilleux. Le fantastique, c'est l'hésitation éprouvée par un être qui ne connaît que les lois naturelles, face à un événement en apparence surnaturel. (TODOROV, 1970, p. 29)

Le personnage de Maupassant vit-il dans un monde proche du notre, monde dans lequel il est fou et seul ? Ou vit-il dans un monde merveilleux où deux êtres peuvent matériellement coïncider ? Le « sentiment du fantastique » (VAX, 1965) provient de cette indétermination entre deux genres opposés. Sitôt que le récit reçoit une explication naturelle ou surnaturelle, il ne s'agit plus du genre du fantastique, mais de l'étrange ou du merveilleux. Le fantastique est donc un genre instable dont la fragile carrière dépend de l'impossibilité de ranger le récit dans le genre de l'étrange ou du merveilleux. Il n'est pas absolu dans le sens où il pourrait caractériser l'ensemble d'un récit, mais relatif aux attitudes du lecteur, qui alterne ainsi entre des ensembles de mondes possibles différents. Comme l'écrit Louis Vax :

> En nous demandant ce que c'est que l'étrange, nous cherchons moins à préciser le sens d'un mot dont nous usons correctement qu'à découvrir les causes d'une impression et à saisir une essence. (...) Méthode équivoque et bâtarde ! Entreprise fondée sur les illusions des données immédiates et de la table rase ! Au fond de moi-même, ce n'est pas un fondement métaphysique que je découvre, mais le résidu d'une culture. Pour l'avoir oublié, je suis amené à juger fondamentale une expérience mouvante et bornée, à ignorer que le sens des mots varie selon les intentions du parleur, le contexte de discours et l'évolution des institutions. (VAX, 1965, p. 6)

Au fur et à mesure de la progression d'une lecture, certaines hypothèses du lecteur sont abandonnées, quand d'autres sont avancées. Aussi, pour que la sémantique des mondes possibles soit un outil efficace à l'interprétation des actes d'imagination dans le cas d'univers fictionnels appuyés sur des œuvres, les relations d'accessibilité qui traduisent ces attitudes semblent nécessiter une structure dynamique dont la forme se voit modifiée au fur et à mesure de la lecture de l'œuvre. C'est particulièrement saillant dans le cas particulier des œuvres fantastiques, parce que celles-ci répondent à des hésitations d'interprétations basculant d'un type à un autre au gré des informations prélevées.

Si les relations d'accessibilité peuvent être relativisées au temps, elles peuvent également être relativisées aux mondes. Ainsi, plutôt que de concevoir des mondes où le personnage est à la fois seul et accompagné, il est également possible de considérer des mondes accessibles selon les interprétations hésitantes du lecteur. Dans le premier type de mondes, le protagoniste est fou et seul, dans le deuxième, le principe d'identité des coïncidents n'est plus valable et le personnage est bel et bien accompagné d'un deuxième. Les mondes de la fiction sont alors vus moins comme des

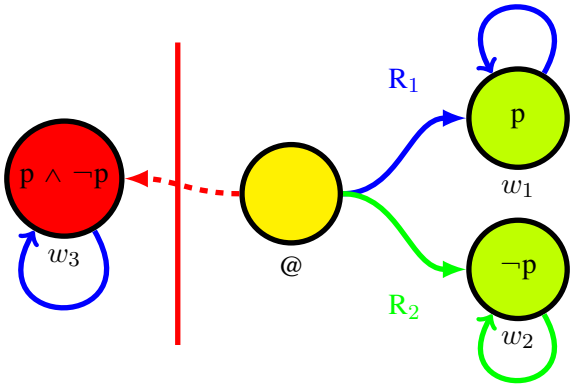


FIGURE 5.7 : Plutôt que d'accéder à un monde impossible (opérateur rouge) on peut concevoir un opérateur accédant à des mondes possibles selon les hésitations du lecteur (opérateur bleu-vert)

mondes inconsistants, mais comment des ensemble de différents mondes possibles diversement reliés par des relations d'accessibilité différentes (bleues et vertes sur l'exemple illustré par la figure 5.7).

Ainsi, ma proposition, pour le dire sous forme de boutade, est donc d'étendre la possible folie du personnage de Maupassant à l'esprit du lecteur. Plus précisément, d'étendre la manière de voir double qui est celle du personnage au contenu mental du lecteur qui oscille de manière permanente entre l'élaboration de deux types différents de mondes : l'un, étrange, où le personnage de Maupassant est seul, l'autre, merveilleux, où il ne l'est pas. Cela nécessite donc de souscrire à une version pragmatique de la logique de la fiction, en donnant une place plus importante, dans la compréhension des attitudes imaginatives au processus dynamique de la lecture, en constatant que cette dernière procède par construction-destruction de mondes possibles. Cela évite l'alternative entre l'hypothèse de contradictions internes ($\circledR(\varphi \wedge \neg\varphi)$ et l'hypothèse de contradictions externes ($\circledR\varphi \wedge \circledR\neg\varphi$) au profit d'une interprétation double $\circledR_1\varphi \wedge \circledR_2\neg\varphi$ pour laquelle les relations d'accessibilité entre mondes possible de la fiction dépendent de l'interprétation que le lecteur donne au texte. Le genre du fantastique serait ainsi caractérisé par la présence d'un opérateur fictionnel double, forme d'interrupteur articulant entre eux les accès à des mondes contenant des hypothèses contradictoires et traduisant l'hésitation du lecteur entre les genres de l'étrange et du merveilleux.

On peut alors considérer que le personnage décrit par Maupassant est individué par une ligne de mondes qui le relie aux individus possibles habitant les domaines de mondes accessibles depuis les différentes relations d'accessibilité $\circledR_1$ et $\circledR_2$.

5.5 Conclusion

La lecture de récits fantastiques mettant en scène des doubles impose un mode de lecture basé sur l'hésitation. Les fictions du double produisent un processus de double vue, où le statut de la co-

référence est mise en doute sans jamais être complètement démentie. Dans ces mondes fictionnels, les mondes possibles que nous considérons basculent d'un modèle à un autre, leurs domaines respectifs contenant plus ou moins d'individus. La pensée du double suppose ainsi une dynamique propre permettant à nos croyances d'attribuer de manière successive l'identité et la distinction numérique à des personnages de fiction. Or la théorie de la référence rigide ne permet pas de comprendre les attitudes d'un agent dont la vision du monde supposerait une identité s'avérant fausse, ou qui rejetterait une identité véritable.

Dans ce chapitre, nous avons montré en quoi les différentes théories réalistes, autant internalistes qu'externalistes, échouent à rendre compte des phénomènes littéraires impliquant des indiscernables. Nous avons vu que l'identification des fictions ne nécessitait aucunement une substance persistante dans le temps pour supporter la vérité des assertions fictionnelles. L'identification des personnages est contextuelle, elle n'est pas tributaire d'une identité objectuelle véhiculée par des *ficta*.

Conclusion de la seconde partie

Dans cette seconde partie, nous avons commencé par distinguer l'irréel de l'illusoire. Un jugement illusoire est un jugement toujours faux. Les fictions, qui relèvent de l'irréel, peuvent en revanche véhiculer le vrai. En outre, nous avons vu les mérites de la sémantique des mondes possibles pour autoriser une telle désolidarisation entre sémantique et ontologie.

Ensuite, en analysant la notion d'œuvre textuelle sous celle de duplique fonctionnelle, nous avons avancé des éléments pour désengager l'œuvre d'une caractérisation substantialiste.

De la question de l'identification des œuvres via des doubles fonctionnels, nous sommes ensuite passés à la question de l'identification des personnages. Là encore, nous avons vu les difficultés auxquelles se heurte l'identité objectuelle. Nous avons favorisé un point de vue conventionnaliste de l'identité en fiction, en nous aidant pour cela de la théorie épistémiques des lignes de mondes.

Le résultat global que nous pouvons tirer de ce qui précède est que le contenu d'une représentation ne suffit pas à fixer la signification. l'imaginaire du double nous confronte parfois à des pensées dont les contenus sont indiscriminables, pensées que nous pouvons toutefois encore distinguer par leur véhicule, par leur mode de présentation. Les doubles ontologiques sur lesquels reposent les *ficta* ne sont quant eux, individués que par des contenus de pensée, et lorsque ceux-ci s'avèrent indiscernables, ces doubles se révèlent ou directement contradictoires, ou nécessitant des propriétés haeccéitiques encore plus étranges que celles attribuées aux entités matérielles.

Si ce qui précède est vrai, l'imaginaire du double semble ainsi nous confronter à deux possibilités ; soit admettre que le double en fiction n'est pas contradictoire, mais alors rejeter les entités fictionnelles doubles, les doubles ontologiques qui supportent les personnages doubles, soit admettre de telles entités mais renoncer dans ce cas à une conception rationnelle du double.

Notre réflexion nous a porté à considérer les œuvres textuelles sous la notion de dupliques fonctionnelles, et à considérer l'identification des personnages à partir de lignes de mondes. Cela nous engage à exporter plus généralement la question de la reproduction aux cas des oeuvres esthétiques et à son rapport à l'authenticité. De là, nous aborderons finalement la question de l'identité personnelle. Car si dire des choses vraies à propos des personnages n'implique pas de faire référence à des substances immatérielles, cela pourrait nous engager a dire la même chose à propos des personnes, c'est du moins la question vers laquelle nous nous dirigeons maintenant.

Troisième partie

Des doubles esthétiques aux identités imaginaires

Introduction à la troisième partie

Selon Paul Ricœur, la littérature est un terrain d'expérimentations essentiellement destiné à mettre à l'épreuve la relation d'identité narrative (RICŒUR, 1990, p. 176) La présente partie est consacrée à la question de l'authenticité, et de là, à celle de l'identité personnelle. J'articule les deux problématiques à travers la question de l'objet esthétique et de l'identité spécifique.

Dans le chapitre 6, je montre comment la question de l'authenticité de l'objet esthétique, dans la critique artistique, a été progressivement remplacée par la question de l'authenticité de l'artiste, qui tient de son identité sociale.

Dans le chapitre 7, il s'agit de critiquer les théories substantialistes de l'identité personnelle au profit d'une théorie de l'identité fictionnelle. Pour cela, nous montrerons l'importance des doubles épistémiques dans la construction de l'identité sociale.

Chapitre 6

Objets esthétiques et authenticité

> Time's glory is to calm contending kings,
> To unmask falsehood, and bring truth to light,
> To stamp the seal of time in aged things,
> To wake the morn, and sentinel the night,
> To wrong the wronger till he render right.
>
> Shakespeare, *The rape of Lucrece*

Dans le chapitre 4, nous avons mis en relief l'intérêt de la notion de duplique en rapport avec l'identification des œuvres textuelles. Les dupliques permettent une identification qui est essentiellement fonctionnelle, basée sur des relations intentionnelles extrinsèques, et non sur des relations qualitatives comme les répliques pures. Ce nouveau chapitre est consacré à leurs rôles respectifs en ce qui concerne les objets dits « esthétiques ».

Nous avons vu que ce que nous appelons œuvres textuelles appartient à un registre de choses qui admettent des dupliques, des doubles fonctionnels, et que toute reproduction correcte vaut pour une nouvelle occurrence. Qu'en est-il de la reproduction de particuliers concrets que sont les objets esthétiques ? Si certains sont accusés d'être des contrefaçons, peut-on aussi parler de contrefaçon d'espèces ?

Dans le paragraphe 42 de la *Critique de la faculté de juger*, Kant examine le cas d'une personne qui prendrait plaisir à écouter ce qu'il supposerait être le chant d'un rossignol alors qu'en vérité, il s'agirait des talents d'un jeune « espiègle » dissimulé dans un buisson. Kant affirme qu'une fois la personne informée de la supercherie, elle y perdrait tout intérêt. Le son ne mérite une attention non en vertu de ses seules qualités esthétiques, mais à la condition qu'il émane d'un véritable oiseau, et non d'une imitation, aussi parfaite soit-elle. Ce que Kant apprécie à travers le son émis, ce n'est donc pas proprement son caractère intrinsèque, mais au contraire extrinsèque, le fait qu'il ait une certaine origine, à savoir, qu'il provienne d'un oiseau *authentique*. Cette remarque indique qu'il n'est parfois accordé de qualités esthétiques à un phénomène qu'à la condition que celui-ci s'accorde avec une forme d'authenticité. Mais poussons l'histoire plus loin, en admettant que le simulacre ait pris une forme particulièrement sophistiquée : imaginons que, comme le célèbre canard construit par Jacques de Vaucanson (1734), il s'agisse d'un automate reproduisant de ma-

nière très ressemblante non seulement le chant, mais également l'apparence et le comportement d'un véritable oiseau. Imaginons qu'on ne puisse pas même distinguer cet automate d'un oiseau authentique, sinon par la confession de son créateur. Comment, dès lors, lui refuser encore les qualités délaissées par Kant ? Nous avons là un problème qui articule trois notions importantes : celui de la ressemblance parfaite, celui de l'identité spécifique et celui de l'authenticité.

Dans la première partie de ce mémoire, nous avons discuté du problème général des indiscernables, en particulier, de l'idée que, puisque deux choses peuvent partager une même qualité, il est logiquement possible qu'elles partagent l'ensemble de *toutes* leurs qualités. De ce point de vue, l'identité numérique est complètement indépendante des identités spécifiques. Nous ne saurions évaluer la singularité d'une chose qu'en raison des différences qualitatives qu'elle exemplifierait à l'égard de choses numériquement différentes (par contraposée du principe d'indiscernabilité des identiques), mais nous ne pourrions jamais nous assurer de l'identité numérique d'une chose en raison d'une totalité de qualités partagées. Comme l'écrit Luis Prieto dans *Le mythe de l'original* :

> l'observation d'un objet ne permet de vérifier la présence ou l'absence dans cet objet que de caractéristiques définissant les identités *spécifiques* dont il est pourvu. En tenant compte de ces caractéristiques et, donc, des identités spécifiques des objets, on réussit cependant, grâce à l'expérience que l'on a des conditions dans lesquelles la réalité matérielle se transforme (ou ne se transforme pas) à calculer la *probabilité* que l'objet auquel on a affaire à un moment donné soit ou non un certain objet numériquement déterminé. [...] La probabilité ainsi calculée peut être plus ou moins grande, mais elle peut être totale seulement pour l'hypothèse négative. (PRIETO, 1992, p. 137-138)

Si la reconnaissance d'une identité numérique relève de probabilités plus ou moins grandes, son évaluation est donc affaire d'expert, c'est-à-dire qu'elle est du ressort d'un individu entraîné à percevoir suffisamment de détails pour confirmer (avec un degré inévitable d'erreur) ou infirmer l'hypothèse d'identité numérique. Un expert ne pourra jamais établir avec certitude qu'un objet est identique à un objet numériquement déterminé, en revanche, il pourra évaluer si l'objet rempli ou non les conditions qu'il juge nécessaires pour l'être. De la question métaphysique de l'identité numérique, on passe alors à la question épistémique de l'authenticité : quels sont les prédicats que l'on doit pouvoir attribuer à une chose pour garantir son authenticité ?

On l'aura compris, ce chapitre est dévolu aux relations entre reproductions, originaux et authentiques. Les reproductions tombent de fait sous trois régimes : celles dont les occurrences sont reproductibles à l'envi sous le même type comme par exemple le mot « mot » ; celles dont les occurrences sont reproductibles, mais auxquelles s'ajoute une valeur d'unicité qui ne tombe pas directement sous son type (comme la valeur marchande de la première lithographie d'un tirage), et enfin celles qui semblent échapper à toute reproductibilité, autrement dit, dont l'occurrence coïncide directement avec le type (comme par exemple une toile de maître). Dans ce qui suit, je montrerai que la notion d'originalité devrait être associée à celle d'identité numérique, autrement dit, de singularité, alors que la notion d'authenticité devrait être associée à la reproductibilité. Je souhaite montrer en particulier qu'il n'y a pas d'authenticité sans duplication. Un original, c'est-à-dire un antécédent temporel d'une chose, ne parvient au statut d'authentique que sous la condition d'être le sujet d'une duplication.

Déroulons le plan du chapitre. Après avoir mis en évidence, dans une première section, la

distinction entre imitation et reproduction, nous commencerons par examiner, dans une seconde section, la notion de double appliquée aux objets matériels, sous l'angle des diverses pratiques que constituent les reproductions, les dupliques, les contrefaçons, les faux. Dans la troisième section, nous nous concentrerons sur le cas des objets dits esthétiques, qui intéressent particulièrement la notion d'authenticité. De fait, l'objet d'art pose non seulement la question, comme tout objet, de son identification, mais également celle de sa nature d'objet esthétique, autrement dit, la question de son identité spécifique. Nous avons donc là, oserai-je dire, un *double* problème de double. D'un côté, nous avons la question des doubles épistémiques que constituent les reproductions d'une seule et même chose, qui interroge sur la forme d'identité qui se tient entre la copie et son modèle, autrement dit, qui demande en quoi « la copie est-elle une autre occurrence de la *même* œuvre ». D'un autre côté, nous avons la question de sa nature dédoublée, autrement dit, de son identité spécifique démultipliée : entre telle copie et son modèle, en quoi l'une peut être dite être une œuvre et l'autre pas ? Dans la quatrième section, nous considérerons le rêve de l'œuvre d'art idéale, l'automate parfait, dans son rapport à l'imitation et à la simulation.

6.1 Reproduire, imiter, simuler

Dans *Pour une reformulation du concept de signe iconique*, Umberto Eco souligne qu'une copie n'est généralement pas un signe de l'original. La copie n'est donc pas une instance de représentation d'un original :

> On définit comme réplique absolument duplicative une occurrence qui possède toutes les propriétés physiques d'une autre occurrence. Étant donné un cube de bois d'un format, d'une couleur, d'un poids, etc. déterminés, si l'on produit un autre cube ayant les mêmes propriétés (et qui forme donc le même continuum, de manière exactement identique), on ne produit pas un signe du cube mais tout simplement un autre cube, qui tout au plus peut représenter le premier, comme de nombreux objets représentent la classe dont ils sont membres, étant choisis en tant que signes ostensifs. (ECO, 1978, p. 143)

Ce que représente la copie, c'est donc le dénominateur commun qui le relie à la collection où figurent d'autres occurrences, mais qui ne représente pas une autre occurrence en particulier.

Néanmoins, certaines duplications d'objets renvoient quand même à un original, en ce sens qu'elles forment un témoignage d'une occurrence antérieure. C'est notamment le cas des objets dits esthétiques, pour lesquels une copie dénote un original, en plus d'en être une duplique.

Il existe plusieurs travaux philosophiques où est abordée de front la question des indiscernables esthétiques. Elle apparaît dans *Langages de l'Art* (1968) de Nelson Goodman, dans *La Transfiguration du Banal* (1981) d'Arthur Danto et dans les discussions successivement apportées par Gérard Genette dans *L'Œuvre de l'Art* (1994), puis par Jacques Morizot dans *Le Problème de Borges* (1999). Mais dans leurs contextes respectifs, ces essais ne posent pas la question du double pour elle-même, l'hypothèse de l'existence d'indiscernables étant essentiellement rejetée (Goodman, dans le cas des arts autographiques) ou admise (Danto) afin de servir d'argumentaire ontologique et de moyen heuristique pour déterminer une différenciation des entités esthétiques considérées. Dans ce qui suit, nous souhaitons au contraire reconsidérer plus spécifiquement le

problème du double en esthétique.

6.1.1 Reproduction et imitation

La reproduction des objets est sans doute aussi vieille que l'humanité : rien ne ressemble plus à un silex biface préhistorique qu'un autre silex biface. Le développement de la technologie, de la taille des pierres à l'impression 3D en passant par tissage au métier Jacquard, a progressivement accru la vitesse de duplication des exemplaires, tout en amoindrissant au maximum leurs dissemblances.

Le domaine du vivant témoigne de différents mécanismes de réplications. Depuis le développement de la génétique, la biologie explique la reproduction de qualités au sein d'une famille ou d'une espèce, autrement dit, l'hérédité, par la reproduction d'une partie de la chaîne acido-désoxyribo-nucléique des individus parents. De ce point de vue, la réplication est essentiellement un phénomène de *transmission* évolutive. L'aspect de transmission présente un caractère conservatif, par la reproduction des gènes, mais cette transmission est altérée par les effets de mutations, et les individus subissent un effet de sélection dû aux contraintes environnementales. La génétique secondée par la théorie darwinienne permet de comprendre le phénomène dynamique et évolutif des reproductions.

A côté des cas de répliques génétiques, la biologie et l'éthologie ont également affaire à la reproduction par imitation, où cette notion désigne la capacité de certains organismes vivants à leurrer leurs prédateurs par le comportement ou par le camouflage [1] (*mimicry*). On a donc ici une sorte de transmission de qualités inter-spécifiques. Cette forme d'imitation ressortit à une ressemblance manifestée entre le comportement, la forme (homotypie) ou la couleur de l'animal (homochromie) et son environnement. Un crapaud adopte une immobilité qu'on pourrait qualifier de « minérale » durant une grande partie du jour, ne commençant la chasse que lorsque la lumière vient à décliner. Sa robe, d'un vert grisâtre parsemé de tâches brunes, lui permet de se fondre dans les herbes, et les nombreuses verrues saillantes qui le recouvrent reproduisent l'aspect granuleux de la terre. Certains camouflages sont particulièrement complexes, comme ceux des phasmes qui reproduisent les traits des branches, des feuilles ou des écorces des végétaux qui leur servent de nourriture. L'imitation chromatique n'est pas toujours passive. D'autres animaux sont capables de modifier de manière active leur coloration en rapport avec celle de leur environnement, comme les truxales qui adoptent successivement une coloration vert clair puis jaunâtre au fil des saisons, imitant ainsi la brûlure progressive des végétaux roussis par le soleil. Chez les caméléons, stars du camouflage, mais aussi chez les turbots, le phénomène est encore plus spectaculaire, en raison des variations chromiques s'adaptant de manière quasi-instantanée au milieu où ils sont placés. Les calliptames d'Italie, sortes de gros criquets méditerranéens, se placent toujours face au soleil, de manière à se poser sur leur ombre afin de dissimuler l'effet de relief que produit leurs corps sur le terrain où ils s'exposent. À côté des effets cryptiques de coloration et de forme, d'autres sont au contraire disruptifs à l'égard du milieu, quand les colorations ne sont pas homogènes à l'environnement, mais au contraire, particulièrement saillantes. Bariolées, elles ont pour effet de rendre la forme générale de l'insecte difficilement reconnaissable, produisant un effet d'indiscernabilité occultive par impossibilité de comparaison. Par ailleurs, l'aposématisme consiste en la

[1] Pour un développement de la question du camouflage chez les animaux, on pourra se référer à VANDEN EECKHOUDT, 1970 et à CAILLOIS, 1938.

manifestation de signaux, généralement des couleurs vives se détachant nettement, dont l'effet est d'avertir le prédateur d'une toxicité réelle ou feintée. Pour finir, rappelons que le camouflage de certains insectes copie les traits des prédateurs des animaux dont ils sont les proies. Les ocelles noires et symétriques dont disposent les Eudia pavonia sur leur ailes, imitent la forme des yeux des rapaces, et produisent ainsi un effet répulsif sur les oiseaux. Et c'est l'imitation de nombreux traits, chez les syrphes, qui les rend presque indiscernables des guêpes, tant par l'apparence que par le comportement.

Bien sûr, de telles capacités mimétiques ne sont généralement pas acquises, mais innées, dues au processus de sélection naturelle des individus. Les individus présentant, de manière plus ou moins aléatoire, des qualités mimétiques trompant les prédateurs persisteront plus longtemps et donc auront plus de possibilités de reproduction que les individus n'en présentant pas. En ce sens, l'absence d'un caractère intentionnel empêche de parler véritablement de « copies » ou du « dupliques », et se rapproche de la similarité intrinsèque des répliques.

A côté de la réplication mimétique par sélection génétique, il convient aussi de souligner la reproduction des comportements qui relèvent de traits culturels répétitifs. Le propre des reproductions culturelles est qu'elles sont liées à des agissements mimétique servant à des phénomènes de transmissions que Richard Dawkins, dans son essai *The Selfish Gene*, désigne par l'appellation de *mèmes*. Les mèmes sont des phénomènes culturels qui, historiquement parlant, succèdent et coexistent aux gènes. Les airs de musique, les slogans politiques, la mode vestimentaire, la manière de modeler des bols ou de construire des ponts, et plus généralement les idées sont des exemples de mèmes. Là ou les gènes sont propagés par reproduction sexuée, les mèmes sont propagés de cerveau à cerveau par imitation. Les mèmes expliquent non seulement les phénomènes de transmission culturelle, mais également leur caractère évolutif, soumis à des altérations. Certains mèmes disparaissent là où d'autres persistent. Les mèmes sont donc l'équivalent mimétique des gènes, et la théorie darwinienne prend donc aussi un sens en dehors de la seule génétique. Dawkins remarque d'ailleurs que, s'il y a quelque chose de commun à toutes les formes de vie qui existent au sein de l'univers, c'est certainement le principe suivant : toute vie évolue par différence de persistance des entités réplicatrices (DAWKINS, 1989, p. 248). Mais cette capacité n'est sans doute pas limitée à la seule reproduction génétique, et doit s'étendre à l'imitation. Contrairement aux gènes qui sont responsable de la réplication, les mèmes doivent être rapportés aux phénomènes de duplications, ils relèvent de reproductions intentionnelles, de similarités extrinsèques.

De fait, comme le souligne Jean-Marie Schaeffer dans *Pourquoi la fiction ?* (SCHAEFFER, 1999, p. 65), de tous les animaux, ce sont sans doute les hommes qui ont le plus développé les stratégies de leurres comportementaux, dont le plus sophistiqué est certainement le mensonge, qu'on pourrait qualifier de leurre linguistique. C'est parce que leur comportement n'est pas génétiquement fixé comme l'est celui d'un crapaud ou d'un serpent que les hommes ont la possibilité d'enrichir indéfiniment leurs stratégies imitatives. Ici intervient l'apprentissage, la pédagogie et plus généralement, la cognition sociale, ce qu'Aristote constatait déjà en soulignant que les hommes sont naturellement enclins à l'imitation. C'est une caractéristique anthropologique dont témoignent les jeux d'enfants qui alimentent les premiers apprentissages :

> Dès l'enfance, les hommes ont, de manière innée, à la fois une tendance à mimer - et l'homme se différencie des autres vivants parce qu'il est spécifiquement enclin à mimer et qu'il a re-

> cours à la *mimésis* dans ses premiers apprentissages - et une tendance à trouver plaisir aux imitations. (ARISTOTE, 1996, p. 82)

On peut distinguer deux types de comportements mimétiques. Le premier type relève des phénomènes d'imitation pour lesquels le comportement reproduit est *déjà acquis*, par exemple ceux que les éthologues qualifient d'amorçage observationnel (*observational priming*). Il s'agit par exemple des tendances qu'ont des individus réunis à adopter mutuellement le même comportement, comportement ayant pour effet de rendre plus cohésive l'appartenance au groupe [2]. La thèse qui est généralement admise actuellement est que l'amorçage observationnel est biologiquement favorisé par les neurones miroirs, catégorie de neurones découverte en 1990 par Giacomo Rizzolatti. Ces neurones sont activés par l'observation d'un comportement, et déchargent un potentiel d'action, qu'il s'agisse de sa propre action ou de l'observation de celle d'autrui. Ces neurones semblent d'ailleurs très spécialisés, puisqu'il semble que chacun ne réponde qu'à un seul type d'action particulière.

Le second type de comportement, appelé apprentissage par imitation (*observational learning*), consiste, pour un individu donné, en l'imitation de gestes produits par un autre individu, et dont le but est l'acquisition d'un comportement inédit. La copie d'un artefact par un apprenti ne consiste généralement pas en un leurre ou une contrefaçon, mais en l'appropriation de la technique nécessaire pour reproduire le modèle. Lorsque l'imitation prend la forme d'une formation dispensée par un maître, elle peut alors revêtir la forme de l'imitation des gestes qui donnent naissance à l'objet. Ce qui est visé est alors moins la copie de l'objet que la copie des gestes techniques indispensables à la production de l'artefact.

Ces remarques indiquent, on le voit, que le rapport des vivants, et en particuliers des humains, à la copie et à la reproduction est profondément ancré dans la constitution biologique et psychologique des individus. L'imitation n'est pas un processus homogène, mais doit être distingué entre imitation par réplique, lorsqu'il s'agit de comparer des similarités intrinsèques, et imitation par duplique, lorsqu'il s'agit de similarités extrinsèques d'ordre fonctionnel.

6.1.2 Imitation et simulation

A côté de l'imitation, il nous reste encore à considérer une autre forme particulière de duplication, la simulation.

Nous avons vu qu'imiter signifie ou reproduire le comportement d'un être ou feindre l'aspect qualitatif d'une autre chose. Qu'est-ce que la simulation ? Quel rapport entretient-elle avec l'imitation ? On peut commencer par répondre que si dissimuler consiste à cacher quelque chose qui existe, simuler revient alors à mettre en évidence quelque chose qui n'existe pas. La simulation est donc une sorte de feinte destinée à produire un simulacre. Par exemple, la simulation informatique permet de générer une réalité dite virtuelle. Elle procède d'une phase de modélisation, mais n'en reste pas là puisqu'elle n'est pas basée directement sur un échantillon de réalité, comme lorsqu'on schématise une situation donnée.

Soulignons pour commencer que ce qu'on nomme « modèles » et « simulations » dans les

[2] Ces deux formes d'imitation étant importantes quant à la construction de l'identité personnelle, nous y reviendrons plus particulièrement dans la dernière partie de ce mémoire.

sciences et dans la technologie présente une telle diversité de formes qu'il n'est pas évident de les regrouper dans un cadre épistémologique unique. Malgré tout, on peut affirmer une première caractéristique : la simulation relève non pas de l'imitation par réplication, mais d'une forme d'imitation par représentation. Contrairement à la reproduction d'un texte qui est elle-même un texte, la simulation d'un mécanisme financier n'est pas elle-même un mécanisme financier. La simulation est irréductible à la réplication et à la duplication.

D'autre part, la simulation, en tant que mode de production du virtuel, a une origine mathématique, à savoir, l'acte de modéliser une situation physique. Gilles Gaston Granger (GRANGER, 1995) rappelle d'ailleurs que son usage, dans les sciences empiriques, remonte à Jean Bernoulli qui parlait de « vitesse virtuelle », mais aussi à D'Alembert qui substitue à la notion newtonienne de force l'idée de « mouvements virtuels ». Mais au passage, nous avons ainsi identifié la simulation au modèle, qu'il reste nous donc encore à décrire.

La simulation se rapporte, nous l'avons dit, au modèle, mais entendu dans un sens différent de celui que nous employons lorsque nous parlons d'une copie et de son modèle. Dans le cas de la simulation, le modèle n'est pas entièrement identifié à l'original, mais confronte à la fois la source réelle de la simulation (l'antécédent) et le résultat, virtuel, du processus de simulation (l'image).

Rappelons que le mot français vient du latin *modulus*, qui signifie unité de mesure, l'étalon utilisé pour estimer les grandeurs architecturales. La maquette et le modèle réduit conservent encore cette notion de grandeur modulo un rapport d'échelle. La notion même de modèle réduit laisse encore la préséance à l'objet copié, qui est bien dissocié de sa représentation à laquelle est accolée l'adjectif « réduit » : le modèle qui a été réduit est bien l'original, et sa représentation (la réduction) a été obtenue par similitude, c'est-à-dire, ramenée à une figure plus petite quoique semblable, autrement dit, conservant les mêmes rapports de proportions. Le modèle réduit n'est donc pas le modèle tout court.

Cependant, l'apparition des premières simulations informatiques a induit une certaine confusion dans la démarcation entre l'objet représenté et l'objet représentant. Dans la simulation informatique d'un phénomène ou d'un objet physique, le modèle est à la fois présent dans l'objet d'où il est prélevé, et dans la représentation qui s'y rapporte. Il est une sorte de dénominateur commun entre les deux objets.

Les choses se compliquent lorsque le modèle, au sens d'antécédent, n'est pas réel, mais fictif. Le modèle réduit d'un AT-AT de Star Wars vendu par la société Kenner est un objet matériel en plastique, alors que son antécédent n'existe tout simplement pas. Un autre exemple encore : au musée des Arts et métiers de Paris sont représentées, sous la forme de réseaux de fils, des surfaces mathématiques dont on serait sans doute bien en peine de trouver des antécédents matériels.

La simulation, qui est le régime général pour lequel le modèle *est* le représentant plutôt que le représenté, semble ainsi imposer une sorte de représentation auto-suffisante, se passant d'un antécédent réel [3]. Il est vrai que l'usage des ordinateurs pour simuler non seulement des courbes abstraites, mais plus encore un espace physique à travers un monde virtuel, semble doter la simula-

[3]Soulignons d'ailleurs que cette confusion est fréquemment reprochée aux scientifiques. Au gré des polémiques concernant les mesures sanitaires nécessaires lors de la propagation de la Covid-19, on a par exemple rappelé aux épidémiologistes de confondre leur modèle avec la réalité (référer à https://www.lesechos.fr/idees-debats/sciences-prospective/covid-19-les-modeles-de-prevision-en-question-1208739).

tion d'une forme d'autonomie, d'indépendance à l'égard du monde matériel, au point qu'en 1965, l'informaticien Marvin Minsky proposa la définition suivante du modèle : « Pour un observateur B, un objet A* est un modèle d'un objet A, dans la mesure où B peut utiliser A* pour répondre à des questions qui l'intéressent au sujet de A » (MINSKY, 1965). Le modèle est, comme on le voit, ici, entièrement circonscrit par la simulation, et non par l'objet original, prenant donc le contresens d'authentique et d'original.

La simulation produit généralement les apparences d'une situation seulement possible, autrement dit, sans original. En ce sens, la simulation « à vide » est en quelque sorte l'opposé de la copie : elle ne copie rien, elle n'est pas le signe de quelque chose d'autre qu'elle, mais produit au contraire sa propre référence, sa propre réalité, *virtuelle*.

6.2 L'artefact esthétique et son double

Le verbe « copier » est transitif. À chaque fois que nous copions, nous copions quelque chose. Autrement dit, une copie dénote un artefact original, c'est-à-dire un artefact qui le précède dans le temps, qui constitue le modèle de la copie. Mais un original n'est pas nécessairement un authentique. Une copie n'est pas nécessairement la copie d'un authentique, puisqu'on peut copier une autre copie. Nous appellerons donc original d'une copie l'antécédent temporel dont il résulte. Il n'y a ici, semble-t-il aucun problème autre que terminologique. Mais il reste cependant à considérer les relations que l'authenticité noue avec l'originalité.

Précisons que, dans ce qui suit, j'appellerai « objet esthétique » un objet dont le mode de fonctionnement relève d'instances culturelles spécialisées, à savoir, le monde de l'art, et dont le mode de consommation est essentiellement esthétique. Il est évident que tous les objets présentent des qualités esthétiques (symétrie, rapports de dimensions, textures, colorations, élégance, rapport relationnel aux autres objets, effet disruptif par rapport au milieu, etc.). Mais ce qui distingue les objets proprement esthétiques des autres objets, que j'appellerai « ordinaires », c'est que l'attention qu'on leur porte est *exclusivement* d'ordre esthétique, soit qu'ils sont directement produits pour cela (comme une toile impressionniste), soit que leur histoire soit plus ou moins incidemment portée vers cette fin (comme par exemple une statue grecque antique ou un bidet de toilette dont la raison première n'est clairement pas esthétique).

La préservation de l'authenticité à travers la duplication est, pour Goodman, ce qui distingue la littérature de la peinture et de la sculpture où la copie n'accède jamais au même statut d'authenticité que l'original. Par sa distinction entre régimes allographique et autographique, l'approche de Goodman n'est pas ancrée à la question de l'imitation, des liens de ressemblance entre la réalité et ses représentations, mais à celle de l'authenticité des œuvres. La question de la ressemblance mimétique n'est cependant pas complètement rejetée, mais plutôt déportée ailleurs : elle ne suit plus l'axiologie verticale réel/fiction mais porte horizontalement sur les ressemblances que peuvent manifester des artefacts entre eux. La distinction de Goodman entre arts allographiques et arts autographiques n'est cependant pas toujours acceptée. Luis Prieto considère qu'une œuvre esthétique est toujours l'invention d'une identité spécifique (PRIETO, 1975), c'est-à-dire une abstraction relevant d'un ensemble particulier de qualités sélectionnées. De ce point de vue, l'œuvre étant un type abstrait, toute reproduction manifestant la même identité spécifique serait moins un

duplicata, un double de cette œuvre qu'une nouvelle réalisation. Ainsi, comme l'écrit Prieto dans *Le mythe de l'original*, toutes les œuvres, qu'elles usent de langages symboliques ou non, offrent une résistance nulle à la duplication, à la copie : la reproduction d'un objet esthétique n'est pas une autre œuvre, mais une nouvelle réalisation de la même œuvre, plus ou moins réussie.

Voilà donc notre question : comment évaluer la réplicabilité des œuvres esthétiques ? On peut de fait questionner la place occupée par la reproduction d'un objet esthétique. La reproduction fonctionne-t-elle comme une copie ou comme une duplique, c'est-à-dire, occupe-t-elle ou non, à l'égard de l'œuvre, un rôle qui lui est substituable ?

6.2.1 La reproduction des œuvres esthétiques

Il y a bien des manières de reproduire des artefacts, et en particulier, de reproduire des objets esthétiques. On peut les reproduire à des fins de conservation, comme ont été reproduites par exemple les Grottes de Lascault, bâties à destination des visiteurs. D'autre part, la restauration peut être vue comme une manière de reproduire une œuvre endommagée. On peut aussi reproduire les œuvres dans un but heuristique, par exemple afin de s'approprier la technique de l'artiste, comme le faisaient les apprentis peintres qui « recopiaient » les œuvres exposées dans les musées. On peut aussi les reproduire dans un but pédagogique afin d'en communiquer l'existence et les qualités. D'autres manières sont lucratives. Elles peuvent avoir pour raison la commercialisation d'un souvenir provenant de la visite d'un monument, comme les porte-clés de la Tour Eiffel. Ce genre de reproduction est bien sûr très infidèle, contrairement à l'affaire des faussaires, dont l'enjeu est de faire croire en l'authenticité d'une œuvre qui, copiée, ne l'est pas véritablement. Nous allons passer en revue ces diverses pratiques de reproduction des objets esthétiques afin d'en faire saillir la notion d'authenticité qui les sous-tend.

La reproduction par restauration et par rénovation

Il peut être intéressant de mobiliser la relation de dupliques dans le cas des objets restaurés. La question étant alors de déterminer si oui ou non il y a persistance de l'œuvre à travers sa restauration, où si cette dernière constitue une sorte de double de remplacement. On aura d'ailleurs différents problèmes selon qu'on considère le cas de restaurations partielles ou totales. Quand la conservation d'un artefact nécessite de très nombreux remplacements, voire la reconstruction complète parce que l'original n'est plus du tout accessible, on parle alors plutôt de reconstitution que de rénovation [4]. La Basilique Sainte-Marie-Madeleine de Vézelay, rénovée, à partir de 1840, par Eugène Viollet-le-Duc, est-elle la même que l'abbatiale bâtie au XIIe siècle ? Lorsqu'on observe la reconstitution de la porte d'Ishtar, au Pergamon Museum de Berlin, a-t-on ou non affaire à la même porte que celle qui protégeait l'entrée de Babylone ? Quand nous parlons d'objets qui n'ont pas de prétention essentiellement esthétique, comme le bateau de Thésée, nous sommes plutôt enclins à affirmer que les réparations ne changent pas l'identité numérique, que la chose demeure la même. La plupart du temps, nous l'acceptons aussi en ce qui concerne l'identité de personnes. La

[4] On désigne aussi par ce terme la répétition d'un objet temporel unique, comme un événement, par exemple lorsqu'on parle de reconstitution historique. Mentionnons également le cas des reconstitutions dont le modèle est purement fictif, comme la maison de Juliette à Vérone ou comme celle de Sherlock Holmes à Londres. Nous aurons l'occasion d'y revenir dans ce qui suit.

perte d'un bras compensée par une prothèse n'implique pas la destruction de l'individu. Pourquoi, alors, serait-ce différent pour des objets, et en particulier, des œuvres d'art ?

Du point de vue de l'identité numérique, il paraît assez intuitif de répondre qu'une œuvre partiellement restaurée est la même œuvre que l'original, qu'elle persiste à travers ses changements matériels. Les changements qu'elle subit ajoutent d'ailleurs à l'original de nouvelles qualités esthétiques qu'il n'avait pas au départ : le jaunissement des photographies par l'action du soleil, la blancheur virginale produite par la desquamation des statues, les béances des ruines, l'oxydation des cuivres et des bronzes et la patine des bois et des marbres manifestent un supplément de qualités qui participent activement à l'appréciation esthétique. Si les statues persistent malgré leurs lentes dégradations, elles persistent sans doute aussi après leur restauration.

A côté de la question de la persistance, la question de l' « authenticité » est davantage discutée. Selon Mark Sagoff, qui associe l'authenticité à une qualité esthétique, une œuvre restaurée n'est pas authentique :

> Le fait est que l'authentique et l'inauthentique sont esthétiquement différents non en raison d'apparences différentes mais parce qu'ils sont des choses numériquement différentes. Certains objets ont des parties remplaçables comme le sont par exemple les pièces automobiles. Un silencieux neuf peut prendre place sur une ancienne voiture. En revanche, les œuvres d'art sont créées une fois pour toutes, par un certain artiste à un certain moment. Michel-Ange créa la *Pietà* à la fin du quinzième siècle ; de Campos peut-il en réaliser des parties aujourd'hui ? Non, il ne peut que remplacer une partie d'une chose, la *Pietà*, avec la partie d'une autre chose, une copie. Si les dommages persistent, peu importe le degré de perfection de chaque réparation, la statue de Michel Ange est inévitablement détruite comme s'il n'y avait eu aucune réparation ; en fin de compte, cette statue peut être réduite à un fragment, alors qu'une reproduction, qui est construite lentement autour d'elle, finira par prendre sa place. C'est justement ce que sont les prothèses aussi bien pour les personnes que pour les œuvres d'art. Elles sauvent les apparences de l'œuvre, mais en changeant sa substance. Elles la tournent, petit à petit en quelque chose d'autre. En un sens, la *Pietà* ne peut être réparée intégralement. Nous pouvons créer une nouvelle pièce, nous pouvons la lui attacher, mais nous ne pouvons pas faire de cette pièce une partie de cette œuvre d'art. [5] (SAGOFF, 1978, p. 458-459)

Cette conception est donc ancrée à une sorte de méréologisme matérialiste essentialiste : la préservation de la totalité des parties matérielles est nécessaire à la persistance temporelle de l'œuvre. Cette conception demande sans doute trop, puisqu'elle affirme que la perte de quelques atomes de marbre, voire d'un seul, implique la destruction d'une statue. Une réponse à cette difficulté consisterait alors à introduire un seuil de tolérance constitutive, comme nous l'avons déjà indiqué dans la première partie de ce travail. Mais l'essentialisme méréologique qui est prescrit par Sagoff, et qui paraît tout bonnement intenable, provient sans doute d'une hypothèse malvenue, à savoir, l'identification de l'authenticité à l'identité numérique. Pour ma part, il me semble qu'il faut au contraire distinguer identité numérique, originalité et authenticité. On peut en effet admettre à la foi que la statue est restée numériquement la même après sa restauration, tout en réfutant son authenticité, un peu comme le cas d'un document administratif qui n'aurait plus cours parce que le contexte

[5] Traduction de l'auteur.

d'évaluation aurait été modifié. Ce sont nos modes de présentation de la statue, comme entité matérielle, qui ont été modifiés, de telle manière que nous n'en n'admettons plus son caractère authentique.

La reproduction pédagogique

La photographie est un moyen, relativement économique, rapide et efficace, pour reproduire l'apparence d'objets, et en particulier, celle des objets esthétiques autographiques. André Malraux en a célébré le mérite, soulignant que les documents photographiques constituent des « musées imaginaires » (MALRAUX, 1996), qui permettent la connaissance d'œuvres qui seraient, sinon, difficilement accessibles en raison de leur éloignement spatial, voire inaccessibles en raison de leur emplacement (fonds de musées, statues disposées en hauteur, etc.). Le qualificatif d'*imaginaire* attribué par Malraux doit être entendu en deux sens. En un premier sens, c'est qu'un document réunit, par la photographie, dans une même collection virtuelle, des œuvres qui se trouvent dispersées, autrement dit, qui ne coexistent dans aucune collection réelle. De ce fait, les reproductions photographiques permettent ainsi des comparaisons entre des œuvres séparées physiquement. Elles permettent aussi des interprétations nouvelles dues à l'attention portée sur certains détails (photographies de détail). En un deuxième sens, le musée est dit « imaginaire » dans un sens plus général, comme l'est finalement tout musée, du fait qu'il présente des objets dont la nature d'œuvre d'art est, elle-même, imaginaire. Car, comme le rappelle Malraux dès les premières lignes de son essai, « un crucifix roman n'était pas d'abord une sculpture, la *Madone* de Cimabue n'était pas un tableau, même l'*Athéna* de Phidias n'était pas d'abord une statue. » (MALRAUX, 1996, p. 11). Ce sont seulement parce que de tels objets sont exposés dans des lieux consacrés qu'ils jouent le rôle d'œuvres d'art, autrement dit, ils ne gagnent ce statut que par la médiation d'une institutionnalisation muséale. Cependant, dans leur contexte d'origine, ces objets relèvent exclusivement de fonctions attachées à des rituels ou à formes de représentations d'individus absents. Les musées « ont contribué à délivrer de leur fonction les œuvres d'art qu'ils réunissaient ; à métamorphoser en tableaux, jusqu'aux portraits » (MALRAUX, 1996, p. 12). Quand il s'agit d'artefacts esthétiques, le double intervient donc non seulement en manifestant l'identité numérique par sa reproduction médiatique, mais aussi au niveau de son identité spécifique, c'est-à-dire, au niveau de la nature de cet objet. Car, comme l'écrit Malraux, les objets esthétiques ne parviennent à gagner leur plein statut d'œuvre d'art que par une sorte d'arrachement à la réalité concrète où ils apparaissent d'abord comme des objets matériels pratiques ou rituels. Cette désolidarisation avec le contexte de leur production leur confère un traitement particulier, leur arrogeant une seconde nature, celle-ci imaginaire, virtuelle, la nature d'œuvre d'art. Dans son écrin muséal, la matérialité de l'objet, surprotégée, tend à être éclipsée au profit de la nature culturelle de l'œuvre, délivrée de sa fonction pratique.

Avant la photographie, les œuvres d'art étaient représentées par des dessins, des gravures, des lithographies. On peut se demander s'il y a une différence ontologique entre ces deux formes, c'est-à-dire entre l'image dessinée et la photographie. La seconde est-elle plus « réaliste » que la première ? Est-elle plus indépendante de son sujet ? La copie d'une œuvre d'art par l'image nécessite un travail de sélection des critères pertinents à reproduire. La reproduction d'une œuvre peinte, d'une sculpture, d'un édifice architectural, consiste à déterminer les traits pertinents qui doivent être reproduits pour en faire une copie fidèle. Mais bien sûr, le copiste peut se tromper en

ne sélectionnant pas tout à fait les critères idoines. La photographie permet, quant à elle, un enregistrement d'un grand nombre de qualités visuelles de son modèle, enregistrant ainsi, de manière plus passive, l'identité spécifique de l'œuvre.

Quel rapport entretiennent les copies à leur modèle ? Que serait une œuvre sans reproduction ? Constatons que c'est le sort des œuvres d'encourager leur reproduction. Les photographies permettent ainsi d'actionner (Goodman) l'œuvre dans des contextes où elle n'aurait pu l'être sans ce recours médiatique. Dans *Conserver/Restaurer*, Jean-Pierre Cometti souligne que les reproductions font intégralement partie de la « vie » de l'œuvre :

> [...] les photographies de sculpture, celles du land art ou d'installations, voire de performances, participent de la connaissance des œuvres, en particulier (mais pas seulement) lorsqu'elles ont cessé d'être accessibles ou lorsque leur nature est telle qu'elles ne le sont qu'une fois. Or, sauf à dissocier les œuvres de leurs modalités de reconnaissance et de la connaissance que nous en avons, ces modalités doivent être considérées comme appartenant à leur mode d'existence : les « œuvres » existent dans la perception que nous en avons (directe et indirecte), dans les connaissances que nous en obtenons, en un sens plus conceptuel, critique et réfléchi, grâce aux moyens que nous mobilisons pour cela (de manière verbale, picturale, conceptuelle, etc.) et grâce aux dispositifs qui en ouvrent l'accès. (COMETTI, 2015, p. 132)

Les reproductions jouent, en quelque sorte, le rôle « d'ombres » des œuvres : elles les accompagnent, témoignant de leur existence effective, elles en manifestent la présence. La connaissance des œuvres est donc une forme de connaissance par le double, les copies d'une œuvre sont des excroissances de celles-ci.

C'est donc bien parce qu'il peut être comparé à ses reproductions qu'un original (c'est-à-dire l'antécédent temporel d'une copie) parvient au statut d'authentique. Une chose ne peut paradoxalement gagner un statut d'authentique qu'à la condition d'être reproduite. Autrement dit, et contrairement à l'identité numérique, il n'y a pas de sens à assigner une authenticité sans reproductions.

6.2.2 La photographie et son double

Dans *L'œuvre d'art à l'époque de sa reproductibilité technique*, Walter Benjamin constatait que « vers 1900, la reproduction technique avait atteint un tel niveau qu'elle était en mesure, désormais, non seulement de s'appliquer à toutes les œuvres d'art du passé en d'en modifier, de façon très profonde, les modes d'action, mais de conquérir elle-même une place parmi les procédés artistiques » (BENJAMIN, 2005, p. 141). Et de fait, certaines photographies ont ainsi, en tant que telles, été érigées au rang d'œuvres d'art. Se pose, ici, la question de la catégorie ontologique à laquelle appartiennent les photographies. Prennent-elles part aux choses matérielles (les tirages papiers), ou à un type abstrait ? Les tirages papiers, les reproductions sur écran, avant d'être des représentations d'autres choses, sont elles-mêmes des choses. En ce sens, elles partagent toutes les privilèges de la réalité, comme le souligne Clément Rosset :

> Photographie, reproduction sonore, peinture sont des produits de l'art, c'est-à-dire des réalités à part entière qu'il serait par conséquent vain de vouloir distinguer de la réalité en général dont

> elles partagent tous les privilèges. (ROSSET, 2006, p. 11)

Elles se partagent donc entre la carrière d'être une chose, et d'être là pour autre chose.

On a souvent insisté sur le caractère de fidélité qui émanerait de la photographie : celle-ci serait un témoignage reflétant la réalité mieux que tout autre procédé de représentation. Comment alors rendre compte de cette objectivité qu'atteindrait un procédé n'appartenant par ailleurs à aucune catégorie ontologique ? C'est sans doute qu'un préjudice d'objectivité est attaché à la photographie, comme si celle-ci se contentait de reproduire passivement la réalité. En ce sens, elle serait un témoignage objectif de ce qui est et de ce qui a été, une sorte de miroir tendu vers le monde. Le caractère spéculaire de la photographie la rapprocherait, dit-on, de la réalité, non de façon approximative, mais manière exacte. L'aptitude présumée de la photographie à se saisir du réel est d'ailleurs inscrite dans la dénomination même de l'appareil qu'est « l'objectif » photographique. La thèse que la photographie désigne un réel « objectif », est tenace. On en trouve une caractérisation aiguë chez André Bazin, qui, dans *Qu'est-ce que le cinéma ?*, écrit que

> L'objectivité de la photographie lui confère une puissance de crédibilité absente de toute œuvre picturale. Quelles que soient les objections de notre esprit critique, nous sommes obligés de croire à l'existence de l'objet représenté, effectivement re-présenté, c'est-à-dire rendu présent dans le temps et l'espace. La photographie bénéficie d'un transfert de réalité de la chose sur sa reproduction. (BAZIN, 1958, pp. 15-16)

Dans *La Chambre Claire*, Roland Barthes défend également un réalisme photographique radical, affirmant l'identité stricte entre l'objet photographié et l'objet représenté sur la photographie :

> Telle photo, en effet, ne se distingue jamais de son référent [...]. Par nature, la photographie [...] a quelque chose de tautologique : une pipe y est toujours une pipe, intraitablement. (BARTHES, 1980, pp. 16-17)

Ainsi, la photographie saisirait directement le réel, sans passer par le biais d'une représentation. La photographie relèverait donc plus du régime de la reproduction de la réalité, d'un double passif, que d'une représentation construite. C'est cette idée qui fait dire à Bazin que l'apparition de la photographie est un événement sans précédent dans l'histoire des techniques, parce que

> pour la première fois, une image du monde extérieur se forme automatiquement sans intervention créatrice de l'homme, selon un déterminisme rigoureux. (BAZIN, 1958, pp. 15)

C'est encore le cas de Kendall Walton (WALTON, 1984), qui affirme que les photographies sont transparentes : elles permettent directement de voir les objets qu'elles *reproduisent* plus qu'elles ne *représentent*.

On peut immédiatement mettre en doute une telle objectivité en mettant en avant le cas des photographies truquées qui ont pour vocation de tromper leur destinataire, comme ce fut le cas à de nombreuses reprises dans l'histoire de la politique. Dans le *Le livre du rire et de l'oubli* (KUNDERA, 1985), Milan Kundera rappelle le cas de la photographie truquée qui présentait le dirigeant communiste Klement Gottwald en train d'haranguer la foule Pragoise. La photographie

initiale témoignait de la présence, aux cotés de Gottwald, de Vladimir Clémentis, qui fut accusé par la suite de trahison puis pendu. Aussi, il avait été jugé bon de faire tout simplement disparaître celui-ci de la photographie. L'histoire montre, au passage, la stupéfiante confiance sur la capacité que la photographie aurait à modifier la perception du réel. Rappelons en effet que la foule Pragoise avait été directement témoin de la scène réelle. On peut aussi mentionner, pour clore ce passage sur les photographies truquées, celles de spectres et autres émanations spiritistes douteuses qui circulaient à la fin du XIXe siècle. La photographie était estimée révéler une réalité échappant à la matérialité physique et aux perceptions communes, dans une sorte d'hallucination du réel.

Mais il n'est pas certain que le cas du trucage soit une objection suffisante contre le réalisme photographique. Il pourrait bien au contraire servir d'argument pour en défendre la thèse. Si la photographie demande à être truquée pour parvenir à ses fins, dira-t-on, c'est que, sans ce trucage, elle serait bien fidèle à la réalité. Il me semble pourtant que nos intuitions plaident en faveur d'une conception directement contradictoire au réalisme photographique : les photographies ne reproduisent pas la réalité, mais représentent une fiction. Pour illustrer mon propos, je souhaite évoquer une exposition photographique qui fut organisée au jardin du Luxembourg de Paris il y a quelques années [6]. Les grilles du jardin, qui servent régulièrement de cimaises à des expositions photographiques provisoires, furent mobilisées cette fois pour des photographies du jardin du Luxembourg lui-même, sans introduire d'originalité marquante de point de vue. L'exposition connu un certain succès : les parisiens se rendirent devant les grilles du jardin pour contempler sa représentation en photographies. Ils prenaient ainsi le temps d'observer ce que nombre d'entre eux avouaient ne pas ou ne plus regarder, en dépit du fait que ces choses s'offraient directement à eux par l'expérience d'une promenade à l'intérieur du jardin. Cette situation permet d'appliquer à la photographie la remarque que Pascal tirait de la peinture : « Quelle vanité que la peinture, qui attire l'admiration par la ressemblance des choses dont on n'admire point les originaux » (B. PASCAL, 1950, aph. 134). Une telle attitude, qui engageait les spectateurs à redécouvrir le jardin, c'est-à-dire à renoncer à l'indifférence du regard, semble indiquer que *la photographie ne présente pas le réel*. En effet quel aurait été l'intérêt de s'intéresser à quelque chose qui ne ferait que redoubler la réalité présente, directement accessible ?

Comme le démontre de manière convaincante Jiri Benovski (BENOVSKY, 2012) dans *Qu'est-ce qu'une photographie*, croire en l'objectivité des photographies, c'est oublier qu'elles relèvent d'un processus intentionnel de sélection parmi de nombreuses possibilités de cadrage, d'angle de vue, de mise au point, etc. qui sont loin d'enregistrer passivement le monde. Un instrument photographique, aussi rudimentaire soit-il, est toujours placé et réglé pour obtenir un certain effet, même si le réglage est pré-programmé (c'est-à-dire déjà établi par les techniciens ayant configuré l'appareil). Et même en laissant de côté l'intervention humaine dans le déclenchent produisant le cliché, la prétendue impassibilité reproductive attribuée à l'appareil photographique reste très contestable. Les photographies produisent en effet, inévitablement, des déformations, ce qui implique que ces représentations ne sont pas isomorphes à leurs modèles réels. Comme l'indique Benovski, les photographies n'appartiennent ni à la catégorie des choses matérielles que sont les tirages papiers, ni même à une catégorie abstraite : les photographies n'existent pas, ou plus précisément, elles n'ont aucun poids ontologique.

Une autre position anti-réaliste, consiste à affirmer que les photographies ne désignent pas

[6] Je n'ai malheureusement pas réussi à retrouver les références précises de cette exposition.

directement leur sujet, mais relatent des mondes possibles dans lequel un objet particulier se trouve représenté. Doit-on en déduire que l'individu représenté sur la photographie est autre que celui qui a été photographié, parce que devenu moyennant la représentation photographique, une chose purement fictionnelle ? Comme l'indique Manuel Rebuschi, en usant de la modalité, on n'est pas contraint de renoncer à l'identité. L'idée est de concevoir les représentations physiques que sont les photographies comme des « fenêtres modales », des représentations d'un même individu dans diverses circonstances possibles, qu'elles soient passées ou imaginaires.

En outre, la photographie favorise la duplication sur plusieurs niveaux. S'il est possible de reproduire des artefacts, il est évidemment possible de reproduire des reproductions. La place des répétitions, en rapport aux originaux, a été questionnée par les artistes eux-mêmes, qui ont admis qu'une répétition pouvait, par elle-même, produire une œuvre nouvelle, c'est-à-dire, une autre œuvre. Avec le *copy art*, appelé encore encore *Xerox art*, du nom de la marque du photocopieur, né autour de 1960, la reproduction devient une fin en soi. Le phénomène n'est donc pas l'apanage de la photographie. On connaît les nombreuses ré-appropriations d'Andy Warhol, des reproductions matérielles de boites de savon Brillo aux portraits stylisés de la Joconde. Mais il vaut la peine de s'attarder sur la photographie en raison des cas pathologiques qu'elle recouvre. Mentionnons à cet égard l'artiste appropriationniste Sherrie Levine qui a exposé des photographies de photographies d'art célèbres. Son exposition *After Walker Evans* à la Metro Pictures Gallery en 1981 montrait des photographies de Walker Evans, rephotographiées par Levine à partir d'un catalogue d'exposition d'Evans, et présentées en tant que telles. C'est le même genre de procédé, appliqué cette fois à la culture populaire, qui est développé par Richard Prince, qui a exposé une photographie de la photographie publicitaire du *cowboy* de la publicité Malboro. La photographie permet ainsi, non seulement la duplication infinie, et le travail sur la série, mais également un jeux de citation métaleptiques incluant la totalité ou la quasi-totalité d'une œuvre dans une autre. La question de l'identité ne peut être facilement tranchée : puisque les photographies de Levine sont indiscernables de celles d'Evans, doit-on inférer qu'il s'agit de la même œuvre ou s'agit-il au contraire d'une œuvre différente ? De fait, lorsqu'elle n'est pas clairement référencée, la répétition obtenue par citation peut aboutir au plagiat. C'est notamment ce qui a été récemment reproché à Étienne Klein parce que son ouvrage de vulgarisation scientifique *Le pays qu'habitait Albert Einstein* est ponctué de citations sans références. Imprimé dans le corps du texte, la citation devient alors une appropriation illégitime qui tombe sous le droit d'auteur [7].

Soulignons que Levine a clairement mentionné l'origine de ses clichés dans le titre même de son exposition, et qu'il ne s'agit donc pas d'une appropriation illicite. D'ailleurs, nous pouvons remettre en question la prétendue indiscernabilité entre ses photographies et celles d'Evans. Si l'on admet que ces photographies sont effectivement qualitativement indiscernables (au moins du point de vue de la perception directe), il reste toutefois qu'en tant que copies, elles n'ont pas le même modèle. Remarquons d'abord que les photographies d'Evans prétendaient représenter des modèles ayant une consistance ontologique évidente, à savoir, elles représentaient des choses réelles. Celles de Levine ne représentent au contraire que des photographies. Une objection à ceci serait que celle de Levine représente bien les mêmes modèles que ceux d'Evans par transitivité de la représentation. Mais le caractère intentionnel des photographies, et de la notion de représentation en général,

[7] Il convient aussi de mentionner l'œuvre d'Elaine Sturtevant dont les productions consistent en des reproductions « de mémoire » d'œuvres d'artistes, questionnant ainsi la notion d'originalité.

ne permet pas cette transitivité : Levine n'a pas eu l'intention de représenter les mêmes modèles qu'Evans, elle a voulu représenter les photographies d'Evans.

Les photographies de Levine indiquent donc, dans son contexte esthétique propre, un problème très proche de celui soulevé en littérature par Borges avec son *Pierre Ménard* : des objets esthétiques peuvent s'avérer indiscernables tout en étant numériquement distincts. Nous pouvons avoir des concepts nominatifs différents au sujet des œuvres respectives de Levine et d'Evans, sans que le contenu de ces concepts s'avèrent différents. Nos représentations sont bien différentes, en dépit de leur ressemblance quasi-parfaite. L'authenticité qu'on peut prêter aux performances respectives et successives d'Evans et de Levine, ne doit donc pas être rapportée ni à l'identité numérique, ni à la ressemblance qualitative.

6.2.3 Le cas des fausses reproductions

La reproduction et son emplacement spatio-temporel

Dans le cas de l'artisanat, produire la copie d'un pot ou d'une chaise consiste à reproduire sa fonction de pot ou de chaise, avec d'éventuelles dissimilarités qualitatives (formes, couleurs, matières, etc.). Reproduire une œuvre est en revanche un exercice plus compliqué, voire, si l'on en croit Goodman, impossible lorsque les traits saillants sont d'une densité bien trop importante, comme par exemple les marques d'un pinceau sur une toile. La reproduction d'une toile nécessite en effet d'intégrer le processus même de production de l'original, ce qui peut demander un long apprentissage afin de reproduire avec exactitude les gestes mêmes de l'artiste copié, artiste qui est généralement l'inventeur de ses propres règles productives, « opérant par là même un acte d'instauration de code » (Eco, 1978, p. 145).

Cependant, il arrive que l'opération de copie mène à un échec, alors même que la copie semble exemplifier une qualité d'exécution supérieure à celle de l'original. Pourquoi le Parthénon de Najville, qui reproduit les canons de la beauté grecque, paraît esthétiquement douteux, contrairement aux ruines de son modèle athénien ? Si l'on suit la distinction que fait Goodman entre les arts autographiques et les arts allographiques, l'architecture figure dans le second groupe. En outre, il s'agit d'un art à deux phases, la première relevant du projet, abstrait, la seconde de sa réalisation matérielle. En conséquence, le Parthénon de Najville, qui n'a pas subi les outrages du temps, pourrait sans doute, de ce point de vue, être considéré comme une meilleure réalisation du Parthénon que son prédécesseur d'Athènes.

Rappelons, à cet égard, la prévention d'Eco de bien séparer l'authenticité d'une œuvre de sa valeur marchande :

> Une copie parfaite de la Pietà de Michel-Ange qui serait en mesure de restituer jusqu'à la plus infime veinure du marbre aurait les mêmes propriétés sémiotiques que l'original. Le fait que la société accorde une valeur fétichiste à l'original concerne une théorie des marchandises qui, si elle contribue à perturber le jugement esthétique, donne matière aux critiques de mœurs et aux censeurs des aberrations sociales. Le goût pour l'authenticité à tout prix est le produit idéologique d'une société mercantile, et quand la reproduction d'une sculpture est absolument parfaite, le privilège accordé à l'original est semblable au privilège accordé à la première édi-

> tion numérotée d'un livre plutôt qu'à sa deuxième : il "intéresse" l'antiquaire, non le critique littéraire. (ECO, 1978, p. 143)

Mais, plus que le sentiment d'une valeur marchande amoindrie, nous éprouvons à l'encontre de l'exemplaire américain une sorte de mépris, voire de dégoût. Pire que cela, nous ne classons généralement pas cet édifice comme une nouvelle réalisation de l'original, mais comme une reproduction d'un goût douteux. Le Parthénon de Najville n'est pas une occurrence nouvelle du Parthénon, il n'est même pas une œuvre d'art, même dans l'hypothèse où il s'avérerait indiscernable de son homologue grec à l'époque de sa construction. Comment expliquer alors que, pour ces deux indiscernables, l'un figure dans le domaine de l'art et pas l'autre ?

De toute évidence, si l'on accepte l'argument précédent, il vient que les propriétés esthétiques de l'édifice d'Athènes ne sont pas restreintes aux seules propriétés intrinsèques de l'objet matériel. Reste à savoir si l'ensemble des propriétés du projet architectural sont également indiscernables. Notre attention est maintenant portée sur le fait que produire le double d'un objet esthétique consiste essentiellement à en reproduire les conditions de création. De sorte que la production d'un double exige non seulement de connaître les propriétés matérielles de l'objet à copier, mais également les propriétés historiques qui ont présidé à la production de l'original. Comme le remarque Umberto Eco, « produire un double, ce n'est ni représenter ni imiter (dans le sens de "faire une image de"), mais produire selon des procédés semblables, de semblables conditions. » (ECO, 1978, p. 144) Le Parthénon de Najville n'est pas une réalisation du Parthénon original, parce qu'un projet architectural doit prendre en compte la situation géographique, temporelle et fonctionnelle de l'édifice à bâtir. Or les contextes respectifs des deux édifices grec et américain sont complètement différents. Lorsqu'un objet d'art comme le Parthénon est copié sans égard aux propriétés historiques et contextuelles de son projet, la copie tombe inévitablement hors du domaine de l'art.

Imaginons alors qu'on en vienne non à produire une duplique architecturale du Parthénon, mais à démonter celui d'Athènes, et à le remonter, à la perfection, à Najville. De fait, l'exemplaire réalisé à Athènes possède les qualités intrinsèques et extrinsèques requises pour l'authentifier au projet architectural initial. Mais déplacé, ses propriétés extrinsèques seraient corrompues par le déménagement. Parvenu à Najville, il resterait donc le même objet, mais nous ne serions plus enclins à lui conférer la même authenticité qu'avant son déplacement. Cela pose, plus généralement la question du rapport entre l'authenticité et la position spatio-temporelle d'un objet. Le déplacement des objets esthétiques : sont-ils indépendants ou partiellement indépendants de leur lieu de réalisation ? Dans le cas des projets architecturaux, la réponse est manifestement non, comme nous l'avons remarqué. Mais qu'en est-il des autres objets d'art ?

Martin Heidegger a insisté sur le fait que les œuvres d'art classiques sont inextricablement liées au contexte artisanal de leur création, contexte qui leur confère directement leur authenticité. L'authenticité est ainsi une composante spatio-temporelle relevant de l'historicité de l'œuvre et de sa fonction pour et dans son contexte social, culturel et cultuel. Ce ne sont donc pas seulement les reproductions architecturales qui s'avèrent esthétiquement défaillantes, mais toute œuvre sortie de son contexte de production et de réception. En conséquence, Heidegger considère avec un certain mépris la muséification, comparant ce phénomène d'arrachement des œuvres à leur monde à l'extraction de matières premières : « On expédie les œuvres comme le charbon de la Ruhr ou

les troncs d'arbres de la Forêt Noire. Les quatuors de Beethoven s'accumulent dans les réserves des maisons d'édition comme les pommes de terre dans la cave. » (HEIDEGGER, 1962, p. 15).

Cette conception liant l'authenticité à l'esthétique, par le biais de la proximité physique est également partagée par Benjamin, qui, dans « De l'art à l'époque de sa reproductibilité technique », soutient que les œuvres sont mobilisées par deux principes : le principe rituel et le principe d'exposition. Ces deux principes ont toujours coexisté, mais la mécanisation des reproductions a essentiellement favorisé le développement du second principe au détriment du premier. La reproduction photographique de l'œuvre engendrerait moins sa connaissance qu'une sorte de mise à distance. Tout comme Heidegger, Benjamin déplore l'arrachement de l'objet d'art à son milieu de création, parce qu'il induirait une perte esthétique : « à la plus parfaite reproduction il manquera toujours une chose, le *hic et nunc* de l'œuvre d'art qui constitue ce qu'on appelle son authenticité. » (BENJAMIN, 2005, p. 13) C'est que pour Benjamin, l'authentique est prévalent sur l'original, non pas pour des raisons marchandes, mais ontologiques. Les regrets exprimés par Benjamin à l'égard des reproductions d'une œuvre originale par les photographies semblent similaires aux reproches que Platon adressait, dans *La République* à l'imitation d'une réalité par l'œuvre artistique. Le fait est que l'authenticité ne peut être dupliquée :

> Tout ce qui relève de l'authenticité échappe à la reproduction - et bien entendu pas seulement à la reproductibilité technique. (BENJAMIN, 2005, p. 14)

Benjamin affirme en effet que, contrairement au processus artisanal qui laisse sauf l'original, dans la production de masse, le modèle est lui-même affecté par la série de ses reproductions parce qu'elles peuvent faire saillir des aspects du modèle qui n'auraient pas pu être aperçus autrement, et occuper des situations qui n'auraient pas été possibles pour l'original. Benjamin déplore que ces possibilités aient pour effet d'altérer l'autorité des œuvres, en les désolidarisant d'une tradition :

> On pourrait dire, de façon générale, que la technique de reproduction détache l'objet reproduit du domaine de la tradition. En multipliant les exemplaires, elle substitue à son occurrence unique son existence en série. Et en permettant à la reproductibilité de s'offrir au récepteur dans la situation où il se trouve, elle actualise l'objet reproduit. (BENJAMIN, 2005, p. 15)

Ce qui est perdu à travers le phénomène de reproduction, cet *ici et maintenant* délaissé, c'est ce que Benjamin désigne par le terme d'*aura* de l'œuvre : « On pourrait réunir tous ces indices dans la notion d'*aura* et dire : ce qui, dans l'œuvre d'art, à l'époque de la reproduction mécanisée, dépérit, c'est son aura. » (BENJAMIN, 2005) Cette notion d'aura est toutefois vague, et, disons-le, assez mystique, puisque Benjamin n'en donne jamais une définition précise, mais la décrit par des métaphores : « Qu'est-ce en somme que l'aura ? Une singulière trame de temps et d'espace : apparition unique d'un lointain, si proche soit-il ». (BENJAMIN, 2005) Il y a donc encore, chez Benjamin, une forme de sacralisation de l'œuvre, une part afférente non détachée de son origine religieuse. L'authenticité n'est pas celle rapportée par une expertise visant à assurer une identification, mais une sorte d'identité métaphysique relevant de l'unicité de l'œuvre d'art, dont l'appréhension la plus complète ne peut être procurée par une reproduction médiatique. Mais il semble que Benjamin ne soit jamais vraiment satisfait de l'estimation de distance à laquelle nous devons nous tenir face à l'œuvre. Si la proximité avec l'œuvre est, dit-il, nécessaire pour atteindre

son authenticité, en revanche, lorsque que cette distance s'annule presque, Benjamin déplore que l'objet perde son statut d'œuvre d'art pour devenir autre chose, un objet *kitsch* : « ce que nous appelions l'art ne commence qu'à deux mètres du corps. Mais voilà qu'avec le *kitsch*, le monde des objets se rapproche de l'homme ; il se laisse toucher, et dessine finalement ses figures dans l'intériorité humaine » (BENJAMIN, 2000, p. 11)

L'industrialisation, en particulier la sériation des objets, est contemporaine de la diffusion du terme « *kitsch* », mot servant à désigner un certain usage des produits manufacturés, dû à la massification de la culture et à sa transformation en mode de consommation. L'objet *kitsch* a pour fin l'imitation d'un objet authentique, mais sans prétendre à son identité numérique. D'innombrables objets *kitsch* représentent *La Joconde* et *La tour Eiffel*. Le processus de sériation *kitsch* peut être caractérisé par ses effets sur son modèle, à savoir, la démultiplication des références au modèle à travers une simplification réductive, une déformation, une caricature. L'origine étymologique du terme a été analysée par Ludwig Giesz dans son article *Phänomenologie des* kitsch*es* (1960) (GIESZ, 1960). L'auteur y propose trois hypothèses pour l'origine du terme. La première est que le terme proviendrait d'une déformation du mot anglais *sketch*, « esquisse », autrement dit, il aurait pour source la référence aux reproductions bon marché que les touristes américains achetaient à Munich en souvenir des monuments architecturaux qu'ils visitaient. La deuxième hypothèse se rapporte au verbe kitsch*en*, issu d'un dialecte allemand et signifiant « déguiser un meuble en le patinant artificiellement ». Enfin, la troisième hypothèse associe le terme « *kitsch* » au verbe allemand *verkitschen* qui signifie « vendre à bas prix », autrement dit, qui associe l'objet *kitsch* à une marchandise peu coûteuse, de mauvaise qualité. Ces trois sources résument bien les diverses connotations que recouvre l'appellation *kitsch* : un produit de bon marché résultant d'une imitation de mauvaise qualité, à l'inauthenticité patentée. Dans *La Guerre du Faux* (ECO, 1985), Umberto Eco reprend à son compte cette définition du *kitsch*, corrélée à celle de mauvais-goût. Il la complète en considérant le *kitsch* comme un ersatz, un succédané, une forme prédigérée de culture exigeant un moindre effort d'intégration dans sa consommation.

Nous ne poursuivrons pas ici la recherche d'une juste métrique pour la contemplation des objets d'art, et nous ne retiendrons pas non plus les reproches adressés par Benjamin aux reproductions, qui, comme nous l'avons déjà fait remarquer plus haut, semblent plutôt faire partie intégrante de la vie d'une œuvre. Ce que nous retiendrons cependant, c'est que ces différentes remarques contribuent encore à distinguer l'identité numérique de l'authenticité. Car, comme nous l'avons vu dans la première partie, nous admettons qu'un objet en déplacement continu reste le même, et que sa trace spatio-temporelle est d'ailleurs un très puissant critère d'identification. Or les constats précédents nous auront éclairés sur le fait qu'il n'en est pas de même concernant la relation d'authentification.

Les contrefaçons et les faux

Quels types de relations la copie entretient-elle avec l'authenticité ? Une copie est-elle analytiquement dépendante d'un authentique ? L'examen des contrefaçons et des faux devrait nous aider à déterminer ces liens. Constatons d'abord que contrefaçons et faux sont des cas extrêmes parmi des possibilités relativement diverses :

> When a work of art of the importance of the Supper at Emmaüs is suddenly revealed to be

> a modern fake, the story naturally becomes a sensational news-item; particularly if the artist responsible has confessed. But what exactly is a fake? Between a painting's being an original by a major master and its being an absolute fake made with the intent to deceive, there are many intermediate degrees. It may turn out to be a school piece, or by a follower rather than the artist himself. It may have been left unfinished and completed by another hand, possibly very much later; or it may have been substantially revised or retouched or prettied up at some later point, perhaps because of its poor condition or damage to it, or to make it more superficially attractive. Again, works by other artists of the same period may get under the wrong name, and may simply be of less exalted authorship. Then there are copies and imitations which were not originally made with any intent to deceive, but which may be passed off as originals subsequently on the art market, either by unscrupulous persons or simply by mistake. (ROSKILL, 1989)

Sous le régime de la contrefaçon, le double se rapporte à l'usurpation d'une chose originale et singulière. La copie de contrefaçon est le double illégitime d'un objet dont est revendiquée frauduleusement l'identité numérique. La contrefaçon ressortit ainsi à la fois du problème de l'originalité, c'est-à-dire de l'antécédence temporelle, de l'authenticité et de l'identité numérique. Remarquons d'abord que la contrefaçon est une pratique qui remonte à loin, et qui n'est évidemment pas la prérogative des objets esthétiques. Dans l'Antiquité déjà, des marchands commercialisaient des copies de statues imitant celles de sculpteurs prestigieux. Des commerçants peu scrupuleux ornaient les amphores de marques usurpant la réputation d'huiles recherchées, et des faussaires faisaient circuler de la fausse monnaie [8].

Il y a deux régimes de contrefaçons, l'un qui s'applique à un objet singulier donné, comme une toile de maître, et l'autre à des espèces comme la monnaie. Commençons par étudier ce deuxième type.

En économie et en droit, on utilise la notion de *fongibilité* pour décrire le caractère interchangeable des choses de même espèce, comme par exemple des céréales ou de la monnaie courante. Deux pièces de monnaie sont données pour identiques en ce que l'échange de l'une par l'autre n'occasionne, dans les circonstances habituelles, aucune différence dans l'usage qui peut en être fait. Ainsi, la fongibilité est une relation de duplicabilité par la valeur : deux objets fongibles sont substituables l'un à l'autre sous leur valeur d'échange [9]. Le caractère fongible de la monnaie motive évidemment sa contrefaçon. L'historienne Corinne Maitte souligne que la circulation des contrefaçons est d'ailleurs solidaire de l'économie, et ne doit pas nécessairement être vue comme une sorte de maladie due à son développement :

> de nombreuses recherches dans le champ de l'histoire économique ont [...] montré que l'imitation et la copie ne sont pas une pathologie du marché, mais un mode normal de son fonctionnement, voire une condition de son développement, et ceci que ce soit dans le domaine du luxe ou de l'ordinaire. (MAITTE, 2015, p.14)

[8] Pour une histoire de la contrefaçon de l'antiquité à nos jours, conférer (DUCHÊNE, 2006) et (COLLECTIF, 1988)

[9] Il arrive cependant que ces mêmes objets perdent leur fongibilité. Un numismate peut discerner deux pièces de même valeur faciale (valeur légale) par une différence de valeur intrinsèque et une différence de valeur de négoce liée à leur mode de production ou leur année de diffusion. Un souvenir affectif attaché par un individu à une certaine pièce de monnaie, pièce qui aurait été obtenue dans une occasion particulière, attribuera une valeur non fongible à l'objet en question.

La contrefaçon implique toujours qu'il existe une entité authentifiée, qu'elle soit connue ou non. C'est ce que souligne notamment Gilbert Ryle dans le cas de monnaies contrefaites :

> In a country where there is a coinage, false coins can be manufactured and passed; and the counterfeiting might be so efficient that an ordinary citizen, unable to tell which were false and which were genuine coins, might become suspicious of the genuineness of any particular coin that he received. But however general his suspicions might be, there remains one proposition which he cannot entertain, the proposition, namely, that it is possible that all coins are counterfeits. For there must be an answer to the question, "Counterfeits of what?"! (RYLE, 1954, p. 94-95)

Si la contrefaçon peut engendrer des soupçons de fausseté au sujet de certaines pièces de monnaie, ce soupçon ne peut s'étendre de manière générale à *toute la monnaie*.

Ryle souligne que sans original, nous ne saurions répondre à la question de savoir de quoi ces pièces seraient des contrefaçons. Le contexte de communication implique de considérer qu'il y a quelque chose dont les pièces sont des contrefaçons. Cette thèse implique qu'il n'y a pas de monde possible dans lequel *tous* les artefacts d'un certain type seraient des contrefaçons. C'est ce que considère Jay Rosenberg quand il écrit que :

> [...] a forged painting is a copy of some original painting, and it could not be the case that all paintings were copies. If all paintings were copies, no paintings would be originals, but if no paintings were originals there would be nothing for the supposed copies to be copies of. (ROSENBERG, 1978, p. 18-19)

Rosenberg, passant des pièces de monnaie aux toiles peintes, admet ainsi la thèse, d'impossibilité des contrefaçons universelles :

Il n'est pas possible que *toutes* les œuvres soient des contrefaçons.

Des objections s'imposent. Il est possible de produire plus généralement des expériences de pensée de mondes dans lesquels des contrefaçons n'auraient pas d'authentique. Michael Wreen propose par exemple celle-ci :

> Suppose that there were no paintings, but that an artistic genius, a well-known sculptor, say, one day announced to the world that the next day he would reveal a new art form, painting. After describing what paintings were, he let's call him Mr. X retires to his studio to execute the first paintings. As luck would have it, though, he's kidnapped, and someone else, Mr. Y, executes the first paintings, making sure to attribute them to Mr. X, and even to sign Mr. Xs name. The new paintings, the only ones in existence, would all be forgeries. (WREEN, 2002, p. 148)

Mais constatons que cet exemple relève sans doute moins de la contrefaçon que du faux. Il ne s'agit en effet plus de la copie directe d'un objet d'art original, mais d'une similarité comparative à d'autres objets dont il est prétendu qu'ils appartiennent à la même signature. Les contrefaçons

sont donc bien des copies, plus ou moins réussies, d'un objet (qu'il soit singulier, comme une toile peinte à l'huile, ou qu'il soit objet de répliques, comme des flacons de parfum). Elles dépendent, analytiquement parlant, d'un original. Les faux, quant à eux, sont des objets qui s'approprient une signature reconnue sans pour cela dupliquer un objet, mais en imitant un style associé à une signature. Il me semble ainsi que la notion de contrefaçon doivent bien être distinguée de celle de faux [10]. Développons cette idée. Bien que les exactions des faussaires soient encouragées par la promesse de bénéfices, remarquons que les faux n'ont pas seulement des effets économiques. Il arrive en effet qu'ils produisent un certaine confusion au sein des catégories établies en histoire, impliquant d'importantes réévaluations esthétiques et critiques, mettant en cause les paradigmes organisant l'histoire de l'art. Citons à ce propos l'affaire des faux Max Ernst, dont Jean Clair atteste que l'« habileté (du faussaire) est telle que les [faux] Max Ernst ont été immédiatement reconnus comme vrais par le spécialiste du peintre, auteur d'un monumental catalogue raisonné de l'œuvre, toujours en cours d'achèvement et dans lequel il propose, enthousiasmé, de les inclure, dans l'édition définitive » (CLAIR, 2012). Et il arrive que de tels réajustements attributifs ait des conséquences théoriques notables, comme le soulige l'historien d'art Philippe Sénéchal, qui affirme qu' « en rendant à Brunelleschi la paternité du Saint Pierre d'Orsanmichele, Luciano Bellosi a permis de repenser totalement la naissance de la sculpture de la Renaissance en Toscane » (ELSIG, GUICHARD, PARSHALL, SÉNÉCHAL & PHILIPPE, 2009). Comme le rappelle Michel Pastoureau, le concept de faux, en art, est culturel :

> Il a ses raisons, ses enjeux, son histoire. Il se définit par rapport à des systèmes de valeurs qui s'articulent dans le temps et dans l'espace, qui se transforment au fil des siècles, qui se positionnent et se repositionnent diversement selon les époques, les régions et les civilisations. (PASTOUREAU, 1988, p. 17)

Je vais m'attarder sur l'idée que, en dépit de l'habitude générale, l'authenticité ne devrait pas être attachée à une relation métaphysique. La fausseté attribuée à un objet se rapporte au lien entre la chose et un signe de cette chose, comme lorsque nous disons par exemple, « c'est une vraie pipe », ou lorsque nous jugeons au contraire qu'il ne s'agit que de la représentation d'une pipe, non d'une pipe *réelle*, autrement dit pour dénier qu'il s'agit d'un particulier concret étendu dans l'espace, possédant une certaine forme de persistance temporelle et remplissant comme artefact une fonction nettement définie. Lorsque nous disons que la fenêtre peinte en trompe-l'œil sur la façade d'un immeuble lyonnais est *fausse*, il ne s'agit pas de dire par là qu'il n'y a rien, mais plutôt que ce qui est perçu n'existe pas en tant que fenêtre. En revanche, en tant que représentation d'une fenêtre, la chose existe réellement. Nous avons ici affaire à un double par représentation [11]. En passant dans le régime du faux, un tableau n'est donc plus vu comme un signe mais comme un signe de signe, un signe représentant (faussement) un artiste. La fausseté n'est pas directement attribuable au tableau (régime ontologique), mais au discours qui est tenu à propos du tableau (régime sémantique), comme la valeur sémantique d'une proposition concernant le tableau. Il n'y a

[10] Pour un point de vue contraire, voir (CARRARA & SOAVI, 2010) qui rejettent l'idée que la notion de copie rentre, analytiquement parlant, dans celle de contrefaçon. Les deux auteurs soutiennent qu'il est seulement contingent que les contrefaçons soient aussi, souvent, des copies.

[11] Cette distinction ne va pas forcément de soi pour ceux qui ne démarquent pas le réel de ses apparences, comme c'est le cas de Pyrrhon qui a enseigné un pur phénoménisme. Selon les sceptiques antiques, la représentation a bien une assise ontologique, qualifiée d'*hypokeimenon*, mais son accès demeure impossible. Précisons qu'on retrouve des traces de ce phénoménisme dans les premières lignes de *L'être et le Néant* de Jean-Paul Sartre (SARTRE, 1976).

pas d'authenticité ou d'inauthenticité sans une déclaration attributive. Mais par ailleurs, cette déclaration ne suffit pas à assurer l'authenticité réelle. Encore faut-il qu'un diagnostic d'authenticité vienne confirmer ou infirmer l'attribution. Authentifier, consiste ainsi à déterminer un pédigrée, à assurer une reconnaissance historique. Avec l'authentification se distinguent alors deux sens du terme « reconnaissance ». Le premier signifie parvenir à établir une identification, un lien de persistance temporelle entre deux entités existantes : je reconnais ainsi la voiture que l'avais laissée sur le parking. Il s'agit là d'une question d'identification numérique. Le deuxième sens est institutionnel et sociétal : il se rapporte à la reconnaissance par les tiers, le fait que certains accordent à ce qui est reconnu un certain nombre de mérites et de propriétés qu'il ne pourrait avoir de lui-même sans ces rapports sociaux à des entités différentes. En cela, la reconnaissance d'une œuvre n'est pas seulement le fait de la retrouver, mais se rapporte également au phénomène de sa consécration en tant qu'œuvre.

L'authenticité n'est pas non plus une relation absolue, prenant ses valeurs entre « vrai » et « faux », en particulier quand il s'agit d'objets esthétiques, comme je vais m'employer à le montrer maintenant.

Considérons un exemple récent. En mai 2016, un résidant de Villeurbanne, amateur de brocante, a acquis une toile par l'entremise d'un site de revente entre particuliers. L'acquéreur affirmait qu'il s'agissait d'une toile de Renoir intitulée *Soirée d'été*. De fait, un tableau portant ce nom avait bien été exposé par le peintre au Salon de 1865. Mais le tableau est néanmoins absent du catalogue de la galerie Bernheim-Jeune, qui dénombre les 4654 œuvres connues de Renoir. L'Institut National d'Histoire de l'Art a donc classé *Soirée d'été* parmi les œuvres perdues de Renoir. Le Villeurbannais prétendait donc avoir retrouvé le tableau. Le cas est intéressant, parce qu'il témoigne d'une hésitation entre différents régimes : s'il ne s'agit pas d'un authentique Renoir, le tableau de Villeurbanne pourrait être ou une contrefaçon d'un vrai tableau de Renoir aujourd'hui disparu, ou un faux Renoir, autrement dit, un tableau sans original imitant la touche de Renoir.

Pour éclairer cette affaire, considérons les divers moyens qui permettent de démarquer une contrefaçon d'un authentique.

Une manière évidente d'accuser une contrefaçon relève bien sûr de l'identité numérique. Dans ce cas, la distinction numérique fondée sur le critère de dispersion spatiale s'avère particulièrement efficace. Une chose matérielle existe à un moment donné à un seul endroit. S'il était possible d'avoir sous les yeux, au même moment, à la fois le tableau de Renoir et la toile de Villeurbanne, nous pourrions conclure qu'au moins l'un des deux ne serait pas l'original. Mais nous n'en saurions pas plus sur l'authenticité de chacun. De toute façon, le plus souvent, un tel rapprochement est impossible. Dans le cas de la toile de Villeurbanne, cette impossibilité est bien sûr due à la disparition du tableau de Renoir. En outre, un tel procédé exige que soit confirmée l'authenticité d'un des deux tableaux afin de conclure en l'inauthenticité de l'autre.

Un autre critère d'identification des contrefaçons relève de la discernabilité. Comme l'a bien souligné Goodman, les contrefaçons réussies posent la question suivante : deux œuvres peuvent-elles être esthétiquement indiscernables ? Goodman pose ce problème de la manière suivante :

> Supposez que nous ayons devant nous, à gauche, le tableau original de Rembrandt, *Lucrèce*, et, à droite, une excellente imitation. Nous savons, grâce à un historique amplement documenté que la peinture de gauche est l'original ; et nous savons par des photographies aux

> rayons X, par l'examen microscopique et l'analyse chimique, que le tableau de droite est un faux récent. Bien qu'il y ait entre les deux œuvres de nombreuses différences – par exemple de paternité, d'âge, de caractéristiques physiques et chimiques, et de valeur de marché – nous ne pouvons voir de différences entre elles ; si on les déplaçait pendant notre sommeil, un simple regard ne suffirait pas à les distinguer. L'on nous demande alors instamment s'il peut se trouver une différence esthétique entre les deux images ; et le ton du questionneur donne souvent à entendre que la réponse est franchement *non*, que les seules différences ici sont esthétiquement non pertinentes. (GOODMAN, 1976, p. 136)

Pour Goodman, la densité graphique d'une toile peinte à l'huile empêche sa reproductibilité parfaite. Les différences qualitatives, si elles peuvent s'avérer indiscernables à première vue, seront inévitablement révélées par l'examen scrupuleux, voire par l'usage de prothèses technologiques (microscope, laser, etc.). Si l'on admet ainsi que le principe d'identité des indiscernables s'applique aux artefacts, il vient alors qu'il n'y a pas de contrefaçons parfaites, autrement dit, qu'une copie d'un original possède toujours quelque qualité qui la singularise de son modèle. L'authentification peut ainsi rappeler le jeu des sept différences, consistant à déceler, entre deux illustrations données, de légères différences. Mais le tableau de Villeurbanne ne peut être comparé à celui de Renoir, même grossièrement, car les informations concernant le tableau disparu s'avèrent extrêmement sommaires. On n'est assuré que de son nom, inscrit dans le registre de l'exposition. Au plus existe-t-il quelques brèves descriptions sujettes à caution [12].

Une autre façon encore, de révéler l'inauthenticité d'un objet d'art passe par l'évaluation de ses propriétés historiques. L'analyse scientifique contribue à obtenir de précieuses indications sur l'origine géographique de l'objet et sur sa datation. Le rôle dévolu à l'expertise scientifique consiste à faire l'hypothèse que le tableau est effectivement de Renoir pour la confronter ensuite à l'expérimentation, quitte à révéler l'aspect erroné de l'hypothèse émise. De telles expertises sont appuyées aujourd'hui par des technologies permettant de dater un objet par coupes stratigraphiques [13]. Dans le cas de la toile de Vileurbanne, une première analyse du tableau, effectuée par le CIRAM, a conclu que la toile a été peinte dans les années 1860, autrement dit, dans les années où Renoir a lui-même exposé son tableau. En outre certains pigments employés sur la toile de Villeurbanne sont également présents sur des toiles de Renoir. Le cadre utilisé est de la marque Dubourg, marque également utilisée par Renoir.

[12] L'institut Wildenstein, qui conserve les archives Renoir, affirme posséder une reproduction du tableau de Renoir, mais refuse de la révéler. L'historien d'art Marc Le Cœur, auteur de *Renoir au temps de la bohème* affirme toutefois avoir découvert une chronique de *Soirée d'été* où la toile est décrite comme représentant « un soleil couchant à travers les feuillages » avec « deux baigneuses ». En revanche, la toile de Villeurbanne représente deux hommes pêchant dans une barque.

[13] Cette méthode consiste à prélever un échantillon de l'œuvre afin de révéler les strates de matières déposées sur le support. Elle permet de mettre en évidence les surpeints, les accidents et les craquelures. Elle s'appuie sur différentes techniques comme l'analyse physique et microchimique de la matière qui permet une identification des différents matériaux en présence ou l'identification des pigments et des liants qui composent la couche picturale. Il existe bien d'autres techniques, comme celle utilisant la fluorescence X, la micro-analyse par microscope électronique à balayage qui fournit à la fois la topographie, l'image chimique et l'analyse spectrale de l'échantillon, ainsi que la radiographie et la réflectographie infrarouge, méthode d'investigation utilisant la lumière infrarouge pour révéler des traces de carbone dissimulées sous les pigments. Cette dernière technique permet de déceler les dessins préparatoires et les repentirs cachés sous l'œuvre exécutée. Citons enfin la dendrochronologie, analyse des coupes de bois par l'étude de la morphologie des cernes qui permettent la datation des bois.

L'authenticité est généralement attribuée par un méta-discours, comme l'est par exemple l'étiquette apposée à l'objet d'art au musée. Il peut s'agir aussi d'une signature, ou plus généralement d'un acte quelconque d'attribution. Le discours d'authenticité prétend que l'œuvre visée réponde bien à une histoire qui doit effectivement être la sienne pour coïncider numériquement avec l'objet visé, affirmant alors une coïncidence totale de toutes ses parties spatiales et temporelles. La principale source d'authentification passe par la signature, autrement dit, l'auto-attribution d'un objet. La signature provient des traditions corporatives des artisans comme les joailliers, ébénistes qui marquent d'un poinçon leurs productions afin d'en assurer l'origine. Avant de devenir le signe qui réfère à une personnalité, la signature renvoie d'abord à une collectivité comme un atelier de production. La signature est d'abord un label d'origine. La tradition de la signature des toiles par leurs auteurs remonte quant à elle à la Renaissance, lorsque l'œuvre passe du rôle d'icône sacralisée ne devant laisser apparaître aucune marque de personnification, au rôle d'ouvrage d'art personnifié révélateur d'un artiste. En outre, André Chastel remarque que :

> Le nombre des tableaux anonymes du xixe et du xxe siècles est incroyablement plus faible que pour les périodes antérieures. C'est à partir des Romantiques qu'il est habituel de trouver des tableaux régulièrement signés. (CHASTEL, 1974, p. 10.)

En commentant l'évolution de la signature sur les toiles, Charlotte Guichard écrit que :

> Cette évolution a été souvent interprétée dans une perspective d'histoire socio-culturelle comme un des signes de la lente individualisation de l'artiste singulier, saturnien voire maudit. Elle mène à la période romantique, qui voit triompher à la fois la signature et le régime de la singularité en art. (GUICHARD, 2008)

De fait, il appert que l'histoire de l'art est, plus qu'une histoire d'objets, une histoire de noms. La collection de musée est un univers de noms propres. En ce sens, la proposition d'une histoire de l'art sans noms propres qu'Heinrich Wölfflin imagine dans sa préface aux *Principes fondamentaux d'histoire de l'art* (WÖLFFLIN, 2016) demeure marginale. L'identité des artefacts d'art dépend de l'identité des personnes qui les ont créés. On dit par exemple « un Rembrandt » pour spécifier un type particulier d'œuvre. Dans ce contexte, le nom propre sert autant à désigner l'origine véritable d'un artefact que son origine présumée. Le nom propre peut même servir à désigner une origine erronée, lorsqu'on dit par exemple de telle gravure qu'il s'agit d'un « faux Rembrandt ». De fait, le faux révèle parfois le vrai, et comme l'affirme Frédéric Elsig :

> Les faussaires, comme Van Meegeren, sont d'excellents connaisseurs et le faux n'est que l'envers de l'authentique, de l'original : tous deux sont des éléments constitutifs du régime de singularité, d'où leur fascination partagée. (ELSIG et al., 2009, p. 351)

La signature peut évidemment être elle-même contrefaite. Un faux contient donc presque toujours un élément de contrefaçon, à savoir, un paraphe ou une signature. A la demande de son acquéreur, la société Lumière Technology a entrepris, en mai 2016, une expertise du tableau, par numérisation multispectrale, d'où fut finalement conclue l'inauthenticité de la toile de Villeurbanne. Contrairement à ce que prétendait le résident de Villeurbanne, le tableau examiné ne présente aucune signature permettant d'authentifier un Renoir. En conséquence, dans un article

consacré au sujet, le quotidien *Le Dauphiné* du 27 mai 2016 titra : « Le Renoir du grenier de Villeurbanne est un faux » (TRAN, 2016). Remarquons qu'avant les déclarations de son acquéreur, la toile de Villeurbanne ne pouvait pas être qualifiée de faux. En effet, la toile n'a pas été vendue comme « un Renoir », mais comme une peinture quelconque réalisée par un artiste mineur du XIXe siècle. En cela, la fausseté du tableau de Villeurbanne ne ressortit pas de celle d'une contrefaçon de signature (faux), ni de l'usurpation d'un artefact par la copie (contrefaçon). En réalité, le statut de fausseté attribué à la toile est la conséquence des déclarations erronées de son acquéreur. Ce n'est pas en vertu de son mode propre d'existence que le tableau est considéré comme un faux, c'est la proposition d'attribution émise par son acquéreur qui est « fausse ». Remarquons qu'avant son expertise, la toile pouvait être dite « fausse » en vertu d'une attribution erronée. Mais dès lors que la méprise est reconnue, l'objet reprend aussitôt sa place de vraie toile, bien que n'ayant pas pour auteur Renoir. Ainsi, le terme de faux attribué au tableau de Villeurbanne n'est sans doute pas approprié.

On voit à travers cet exemple que l'inauthenticité s'épuise ici dans la reconnaissance de l'erreur. Sous ce point de vue, affirmer d'un objet qu'il est un « faux » est seulement un raccourci grammatical pour dire que l'attribution qui lui est prédiquée n'est pas valide. Or cette reconnaissance ne peut s'affirmer qu'après avoir reconnu l'erreur. Dès l'instant où l'expertise révèle cette erreur, la toile de Villeurbanne cesse aussitôt d'apparaître comme un Renoir. Or, cela n'est généralement pas le cas des faux artistiques. Ce qui est particulier au régime des faux, c'est que la découverte d'une attribution fausse n'aboutit pas complètement à rendre une authenticité redirigée. Concernant le cas des faux Vermeer peints par van Meegeren, Lucien Stéphan remarque ainsi que « préciser qu'il s'agit d'un faux Vermeer de la main de van Meegeren n'en fait pas non plus un authentique van Meegeren. » (STÉPHAN, 1991, p. 8) Il y a donc une certaine résistance du faux même dans la découverte de la supercherie qui ne fait pas disparaître pour autant le caractère faux de l'œuvre. C'est sans doute ce qui fait aussi dire à Thierry Lenain que le :

> « Faux » ne se réduit en aucun cas à un prédicat privatif, dénotant l'absence d'authenticité. Bien au contraire, tout faux est un objet qui, en un sens positif, se donne faussement pour authentique : il fonctionne *ipso facto* comme un simulacre d'authenticité. (LENAIN, 2013)

Les faux relèveraient ainsi d'une catégorie médiane entre les inauthentiques patentés et authentiques reconnus, celles des simulacres, qui font perdurer le simulacre par delà le verdict d'inauthenticité.

6.2.4 La question de l'identité spécifique des objets esthétiques

Comme nous l'avons vu, l'objet *kitsch* représente une forme de double, à la fois par son caractère répétitif lié à son mode de production en série, mais aussi par le fait qu'il s'agit de la reproduction d'un objet original. Il est reproché à l'objet *kitsch* de ne pas conserver la signification de ce qu'il copie. Dans sa reproduction *kitsch*, la signification de l'objet esthétique est déformée, sinon, effacée. Le phénomène *kitsch* ressortit donc de la reproduction en série d'un objet autographique original, qui semble y perdre son authenticité pour tomber dans le régime des objets banals. A

partir des expérimentations esthétiques de Marcel Duchamp [14], puis plus généralement du Pop Art américain, on assiste au phénomène inverse, consistant à intégrer, au sein du parnasse des objets dits esthétiques, l'une des occurrences ordinaires d'une série d'objets tout aussi ordinaires. Le but visé par les sérigraphies d'Andy Warhol, reproduisant des bustes stylisés de Mona Lisa ou des conserves de soupe Campbell, et l'empilement de ses boites de savon Brillo ne consiste pas à copier une chose singulière, mais à copier leur caractère reproductible.

On raconte qu'Ingres s'était fâché de ce qu'un portefaix venu transporter son portrait de Cherubini chez l'encadreur, avait fait son office sans un mot. Sans formuler la moindre remarque à l'égard du remarquable travail du peintre, le portefaix avait emporté l'œuvre comme un quelconque objet fait de planches et de toile, ne semblant prêter aucun intérêt esthétique au tableau d'Ingres. De même, dans *Peinture et Réalité*, Étienne Gilson rappelle qu'en juillet 1954, une toile de Tintoret fut retrouvée dans le sous-sol de la cathédrale de Milan où elle servait alors de bâche pour protéger des objets hors d'usage. Un objet auquel on ne porte pas attention, peut-il être une œuvre d'art ? Ingres et Gilson s'offusquent que cette nature ne soit pas reconnue dans l'objet. Mais on peut aussi faire l'hypothèse que la nature d'œuvre qu'on attribue à un objet matériel dépend de l'attention qu'on lui porte. Sans cette attention, l'œuvre n'existe tout simplement pas. Comme le dit Goodman, il faut que l'œuvre soit « activée » pour exister. Beaucoup d'auteurs ont insisté sur le caractère fantomatique et intermittent des œuvres d'art : une œuvre n'existe que lorsqu'elle provoque, chez une personne, une attention adéquate, qui la fait vivre comme phénomène purement mental. Ne devrions-nous pas admettre alors que les œuvres n'existent que dans l'esprit de ceux qui y pensent ?

Gabriel Danto, stimulé par les expériences esthétiques de Duchamp et de Warhol, a repris l'expérience de pensée des indiscernables posée par Goodman là où ce dernier l'avait laissée, constatant que :

> Goodman, curieusement, rejette une des conditions de la question, à savoir la relation d'indiscernabilité. Il semble être d'avis qu'elle ne saurait être que momentanée et que tôt ou tard des différences se feront jour. (…) Goodman poursuit en soutenant qu'on ne saurait jamais prouver qu'une différence perceptuelle est par principe impossible : des objets qui nous paraissent indiscernables aujourd'hui pourront demain se révéler si différents l'un de l'autre que rétrospectivement on se demandera avec étonnement comment on a pu les confondre. Et pour donner plus de poids à sa position, il souligne notre extrême acuité visuelle et auditive qui nous rend capables de percevoir des différences très nettes à partir de variations minimes. A la limite, il s'agit pour lui d'un problème de psychophysique plutôt que d'ontologie. (A. DANTO, 1981, p. 87)

Danto concède à Goodman la validité du principe d'identité des indiscernables lorsqu'il est appliqué aux œuvres autographiques :

[14] Rappelons qu'avec *Fountain*, Duchamp avait proposé un urinoir aux organisateurs de l'exposition de la Société des artistes indépendants de New York de 1917, objet signé au nom de Richard Mutt et que Duchamp s'était procuré dans un magasin de New York. L'objet ayant été refusé par la galerie, il fut ensuite perdu. Il fut finalement plébiscité à partir de 1945, mais à cette époque, l'urinoir n'était plus fabriqué. Aussi, pour satisfaire à la demande des musées, Duchamp fit construire ses propres modèles, copiant les originaux. L'œuvre exposée au musée national d'art moderne du Centre Pompidou de Paris est donc non seulement une contrefaçon, créée de toute pièce par Duchamp, mais plus encore un faux *ready-made*.

> L'idée qu'il doit bien toujours exister une différence entre deux objets qui ne sont pas identiques est pratiquement de l'ordre d'une certitude logique. (A. DANTO, 1981, p. 88)

Bien sûr, toute la difficulté de la question réside dans ce *pratiquement* qui tend à affaiblir la force logique du principe. Car un principe est ou n'est pas logique.

Mais, de son côté, Danto complète l'approche de Goodman en posant la question supplémentaire : est-il possible que, pour deux objets qualitativement indiscernables, l'un appartienne au domaine esthétique et l'autre pas ? Pourquoi la boite de Brillo présentée par Warhol fait-elle figure d'œuvre quand ce n'est pas le cas la boite de savon originale ? Se pose ici la question de l'identité spécifique de l'œuvre, de son authenticité en tant qu'objet esthétique, de sa nature ontologique : qu'est-ce qui fait d'un carré noir sur fond blanc une œuvre d'art plutôt qu'un simple croquis d'élève dans un cours de géométrie ? La question de Danto n'est pas de déterminer le type de propriétés qui distingue deux œuvres entre elles, mais celle qui permet de discerner une œuvre de ce qui n'en n'est pas une. Alors que Goodman tentait justement d'évacuer la question ontologique de l'œuvre, Danto la réinjecte au centre du débat. En effet, la question suscitée par les indiscernables relève, selon Danto, de déterminer en quoi l'interprétation transfigure l'objet en élevant sa nature au rang esthétique :

> Apprendre qu'un objet est une œuvre d'art, c'est savoir qu'il faut être attentif à des qualités qui font défaut à sa réplique non transfigurée et que donc il provoquera des réactions esthétiques différentes. Et ceci n'est pas une question institutionnelle, mais ontologique. Il s'agit en effet de deux types d'identités totalement différentes. (A. DANTO, 1981, p. 168)

La question que pose Danto est la suivante : que reste-t-il d'un objet esthétique quand lui sont retirées toutes ses propriétés qualitatives ? Y a-t-il encore quelque chose, une propriété surnuméraire qui viendrait à témoigner du statut particulier, de l'idiosyncrasie de l'œuvre en tant qu'œuvre, une essence de l'œuvre ? L'expérience de pensée de toiles qualitativement indiscernables montre que ce qui fait la différence entre une œuvre d'art et un objet banal n'est pas localisable au seul niveau des propriétés qualitatives, puisqu'une toile accidentellement enduite de peinture rouge n'aura pas la même valeur esthétique qu'une toile qualitativement identique mais peinte avec l'intention de produire une œuvre [15]. Les propriétés esthétiques ne sont pas exclusivement des propriétés qualitatives. Il est certain que ce qui fait de *Fountain* une œuvre d'art n'est pas localisable au niveau d'une propriété qualitative, mettons, une blancheur qu'exemplifierait l'émail particulier de l'objet signé par Duchamp. De même, ce qui fait de *la* boite Brillo d'Andy Warhol tout autre chose qu'une boite *de* Brillo ne relève pas des propriétés matérielles de l'objet, sinon l'on devrait rendre à son créateur initial, Steve Harvey, le mérite qui est, en fait, dévolu à Warhol. La boite n'aurait certainement pas attendu le vernissage du 21 avril 1964 à la galerie Stable de la 74e Rue Est de Manhatan pour gagner son statut d'œuvre d'art. L'effort de Danto vise donc à montrer que la nature d'œuvre n'est pas intrinsèque à l'objet esthétique. Elle est au contraire extrinsèque :

[15] Avant même les provocations esthétiques de Duchamp, et en réponse aux premières œuvres sabordant la figuration sans prétendre néanmoins s'en détourner, Alphonse Allais avait introduit en peinture l'idée des monochromes. Allais avait ainsi proposé une toile constituée d'un aplat rouge, et étiquetée *Récolte de la tomate par des cardinaux apoplectiques au bord de la Mer Rouge* (1882). Un autre tableau, complètement bleu, est appelé *Stupeur de jeunes recrues en apercevant pour la première fois ton azur, ô Méditerranée*. Enfin, n'oublions pas aussi le magnifique monochrome blanc intitulé *Première communion de jeunes filles chlorotiques par un temps de neige*.

> Depuis le début de ce livre, je soutiens qu'on ne saurait réduire une œuvre d'art à son support matériel, ni l'identifier à lui, car dans ce cas elle serait ce qu'est le simple objet - un carré de toile peint en rouge, quelques feuilles de papier de riz sale, etc. J'ai proposé au contraire de soustraire cet objet matériel, quel qu'il soit, de l'œuvre, afin de voir ce qui subsisterait, supposant que c'est là qu'on pourrait trouver l'essence de l'art. (A. DANTO, 1981)

Ainsi, conclut Danto, concevoir le caractère esthétique d'un objet nécessite quelque chose qui échappe à la seule perception :

> To see something as art requires something the eye cannot decry–an atmosphere of artistic theory, a knowledge of the history of art : an artworld. (G. DANTO, 1964, p. 580)

Cette problématique amène Danto à répondre d'abord à la question de Goodman « quand y a-t-il art ? », même si cette question n'est pas encore suffisante pour l'approche essentialiste qu'il mène. Mais elle permet déjà de déterminer à quelle condition une œuvre est reconnue comme telle, condition qui est certes nécessaire, mais qui n'est pas, à ses yeux, suffisante. Cette condition, c'est l'ensemble du discours des raisons dominantes qui fait le jeu des artistes, des critiques, des galleristes, etc. Danto recherche donc une définition essentialiste de l'art, mais il est conduit à proposer les conditions de réception des œuvres par des propriétés relationnelles liées au contexte historico-culturel développé par et dans *le monde de l'art*. Mais comme le précise Danto lui-même, contrairement à la théorie de George Dickie pour qui le monde de l'art est une structure de rôles au sein de laquelle des artistes créent de l'art, il ne s'agit pas d'une théorie institutionnelle de l'art :

> Je voulais en fait parler d'un monde consistant en œuvres d'art, une communauté d'objets ontologiquement complexes, souvent inter-référentiels (ou, suivant l'expression qui devait plus tard s'imposer dans l'usage, "intertextuels"), s'enrichissant elle-même, et qui surtout était orientée historiquement, de sorte qu'il devenait possible à un moment donné que quelque chose fasse partie de ce monde alors que ça ne l'était pas auparavant. (G. DANTO, 2016, p. 130)

Danto prévient que ce n'est pas parce qu'une œuvre est vue, de manière pragmatique, comme de l'art qu'elle est, de manière ontologique, de l'art. Cela ne répond donc pas, conclut Danto, à la question de savoir « qu'est-ce que l'art ? » :

> Je demeure convaincu que même si les facteurs institutionnels doivent être pris en compte à un certain moment par la philosophie de l'art, la définition de l'art doit transcender les différences institutionnelles pour être véritablement philosophique. (G. DANTO, 2016, p. 133)

La reconnaissance d'une œuvre d'art, son authenticité, est donc générée par des propriétés relationnelles liée au contexte, mais, n'en déplaise à Heidegger et Benjamin, ce contexte n'est pas nécessairement le lieu de sa production, comme l'indique le cas de *Fontaine*. Il ne s'agit donc pas d'une authenticité ayant un rapport ontologique à un emplacement, mais plus justement un rapport à un ensemble de rôles joués par les membres du « monde de l'art ».

6.3 De l'authenticité esthétique à l'esthétique de l'authenticité

Les critères d'authenticité répondent traditionnellement aux questions suivantes :

1. L'objet est-il un original ou une copie ?
2. L'objet visé provient-il de la main de l'auteur qui se l'attribue ou à qui il est attribué ?
3. Quel est le degré de sincérité du concepteur de l'œuvre ?
4. Quel est le statut professionnel de son concepteur ? Quelle reconnaissance sociale lui concède-t-on ?

Avec l'art moderne, puis avec l'art contemporain, chacun des critères précédents s'est vu rejeté, déplaçant ainsi la frontière de l'art toujours un peu plus loin et la notion d'authenticité s'est vue progressivement glisser de l'objet d'art en direction de son concepteur. Je vais analyser, dans ce qui suit, les différentes étapes de cette transaction.

Jean Clair souligne que toute œuvre d'art devrait exemplifier la qualité d'authenticité, caractérisée par l'originalité temporelle et l'unicité ontologique. Mais celle-ci est immédiatement déportée de l'objet d'art à celle de son créateur :

> Une œuvre d'art serait nécessairement, en tant qu'œuvre, originale, unique. Faite d'une seule main. C'est un *unicum* tirant sa valeur d'être le fruit inimitable, irreproductible, de cet individu devenu tout aussi unique et irreproductible que serait un artiste crédité du fait d'être un génie, et par conséquent d'une originalité et d'une origine sans égales. Dans un monde voué à la réplication technique illimitée, où il devient le plus souvent impossible de discerner l'original de la copie ou de la répétition, une telle qualité de l'œuvre d'art, quasiment métaphysique d'unicité, et la seule au demeurant à la manifester, lui assure d'atteindre à la vente des sommets financiers sans limite, alors que tous les autres biens, puisque reproductibles et d'une banalité plus ou moins évidente, sont au contraire mis sur le marché à des prix régulièrement en baisse, des voitures aux ordinateurs. (CLAIR, 2012, p. 150)

L'authenticité est ainsi circonscrite par une sorte de cercle vicieux : si l'œuvre mérite d'être dite « authentique », elle le doit à l'authenticité de son créateur vu comme un génie, sans quoi tout barbouillage ferait office d'œuvre d'art. Mais d'où vient ce caractère génial ? Il le doit à sa propre unicité. C'est le geste de l'artiste qui fait l'œuvre. Le prestige d'une œuvre est associé à son unicité d'une manière quasi-superstitieuse :

> On préférera *Les Noces de Cana* du Louvre à la version réalisée à Venise, quand bien même ce tableau serait endommagé, mal éclairé et mis dans un contexte qui le dénature, parce que l'on sait, ou que l'on croit savoir, qu'il a été fait de la main même de l'artiste. Et ce qu'on cherche dans l'œuvre et devant quoi on s'incline, plus que devant la perfection de sa composition et de sa facture, c'est la présence de cette main, ou plutôt son fantôme. [...] Ce que la relique était à l'œuvre ancienne, l'ombre de la main de l'artiste l'est à l'œuvre moderne. (CLAIR, 2012, p. 167-168)

Ce passage est exemplaire du glissement d'attention opéré entre l'œuvre esthétique authentique et l'authenticité de son créateur vu comme un phénomène esthétique. De ce point de vue, l'objet d'art est essentiellement un signe dénotant le caractère singulier de son créateur.

6.3.1 Connoisseurship et identité personnelle

Depuis le XIXe siècle, on désigne du terme de connoisseurship la pratique de juger de l'authenticité d'une œuvre d'art. Parmi les critères d'authentification qui restent axées sur l'évaluation des qualités intrinsèques de l'objet matériel, il vaut la peine de citer la méthode développée par le critique d'art Giovanni Morelli (1816-1891). Elle est essentiellement comparative, basée sur l'analyse les relations de similitude (au sens géométrique du terme) entre l'œuvre en instance d'attribution et des œuvres déjà authentifiées. L'originalité du procédé de Morelli est qu'il ne se veut pas synthétique mais analytique, recherchant à travers *le détail* à confirmer ou infirmer le diagnostique d'authenticité. La méthode de Morelli a en effet pour principe premier de négliger les traits les plus prééminents, les plus visibles, comme le dessin des mains et des sourires, et de ne s'intéresser qu'aux traits les plus anedoctiques, comme le dessin des ongles ou du galbe des oreilles, détails dont l'exécution n'est généralement pas enseignée dans les écoles d'art et qui sont donc fortement personnalisés. Ainsi, selon Morelli, le détail révèle toujours la pâte, la « manière », le trait caractéristique d'un artiste. La méthode exige donc de prendre en considération des données biographiques de l'artiste lui-même par l'examen de détails caractéristiques qui doivent se retrouver de tableau en tableau. On n'a d'ailleurs pas manqué de critiquer les livres de Morelli de ressembler à une vaste galerie anatomique, tant ils sont parsemés d'illustrations de doigts, d'orteils et d'oreilles [16].

Par la prise en compte de l'anecdotique, Morelli anticipe les thèses de Freud au sujet des contenus latents et manifestes des rêves [17]. Selon Freud, la méthode de Morelli permet de relever les contenus latents des intentions. Les petits détails sont, comme les lapsus, des gestes échappés sans intention consciente, ils sont révélateurs de la personnalité de leur auteur. Freud a d'ailleurs lui-même mentionné les travaux de Morelli comme une source d'inspiration pour l'élaboration de la méthode psychanalytique. Dans *Le Moïse de Michel-Ange* (1914), il écrit :

> Longtemps avant que j'aie pu entendre parler de psychanalyse, j'avais entendu dire qu'un connaisseur d'art, Ivan Lermolieff, dont les premiers essais furent publiés en langue allemande de 1874 à 1876, avait opéré une révolution dans les musées d'Europe, en révisant l'attribution de beaucoup de tableaux, en enseignant comment distinguer avec certitude les copies des originaux, et en reconstruisant, avec les œuvres ainsi libérées de leurs attributions primitives, de nouvelles individualités artistiques. Il obtint ce résultat en faisant abstraction de l'effet d'ensemble et des grands traits d'un tableau et en relevant la signification caractéristique de détails secondaires, minuties telles que la conformation des ongles, des bouts d'oreille, des auréoles et autres choses inobservées que le copiste néglige, mais néanmoins exécutées par chaque artiste d'une manière qui le caractérise. J'appris ensuite que sous ce pseudonyme russe se dissimulait un médecin italien du nom de Morelli. Il mourut en 1891, sénateur du Royaume d'Italie. Je crois sa méthode apparentée de très près à la technique médicale de la

[16] Pour une analyse de l'insolite de la méthode de Morelli, voir Wind (WIND, 1986, p. 63-74)

[17] Pour une analyse des rapports entre les travaux de Freud et de ceux de Morelli, voir (GINZBURG, 1980)

> psychanalyse. Elle aussi a coutume de deviner par des traits dédaignés ou inobservés, par le refus de l'observation, les choses secrètes ou cachées [18]. (FREUD, 1985b)

Les indices iconographiques que relève Morelli sont traités par Freud comme des symptômes médicaux. L'objet esthétique est alors le signe de la personnalité, de l'identité personnelle de son créateur.

6.3.2 L'esthétisation de l'artiste

Les pratiques de l'art contemporain semblent montrer que la spécificité esthétique est sans doute moins attachée aux objets qu'aux personnes. Le problème, comme je souhaite le montrer, est que l'authentification d'une œuvre relève d'un discours fondé sur l'authentification d'une personnalité du monde de l'art. Ce qui importe, ce n'est pas la question d'intégrer un artefact au rang des objets d'art, mais d'intégrer un artiste dans le monde produisant le discours des raisons dominantes du marché de l'art.

De son point de vue de sociologue de l'art, Nathalie Heinlich montre que l'épreuve d'authentification « ne porte pas seulement sur les objets, par le contrôle de leur origine : elle porte aussi sur les personnes des artistes, à travers l'interprétation de leurs propriétés et de leurs intentions » (HEINICH, 1999). Les œuvres d'art ne sont évidemment pas les seuls objets signés, mais on peut toutefois se demander si un objet qui n'est pas signé conserve l'intégralité de son statut d'objet esthétique : « n'y a-t-il artiste que s'il y a signataire d'une œuvre, et œuvre d'art que si l'objet en question est signé ? » (HEINICH, 2008, p. 97) Heinich conclut que :

> Qu'il s'agisse [...] des arts autographiques ou des arts allographiques, on observe bien, soit par l'enquête historique, soit par l'observation du temps présent, une constante : la signature va de pair avec l'accession d'une activité au rang d'art, et d'un producteur au rang d'artiste ou d'auteur. On est donc en droit de la considérer comme un indicateur majeur d'artification. (HEINICH, 2008, p. 106)

Selon la sociologue, l'authenticité donne consistance à la singularité artistique :

> L'authenticité est ce qui fait « tenir » la singularité, ce qui la garantit, ce qui permet de jouer positivement de son ambivalence en la constituant en atout et non pas en stigmate. (HEINICH, 1998, p. 123)

Cette condition d'authenticité est la continuité de lien entre le créateur et l'œuvre créée.

Chez Heinich, l'authenticité esthétique relève de trois grands critères : l'intériorité, l'originalité et l'universalité *de l'artiste*. La cible du crible d'authenticité est donc, on le voit, complètement déplacée de l'œuvre en direction de son créateur. C'est à l'auteur de l'œuvre, et non à l'œuvre elle-même, qu'on en vient aujourd'hui à demander des gages d'authenticité. Les productions bénéficient ensuite des retombées de cette attribution, devenant authentiques *en un second temps*. Le

[18] À ce sujet, on peut souligner également que l'essai de Freud, tout comme celui de Morelli, ont été publiés sous un pseudonyme. Il est notable que deux des principales contributions à la méthode d'authenticité et à l'analyse de la personnalité aient été signées sous de faux noms.

problème de la signature est donc celui d'une esthétisation de l'identité personnelle. Ce n'est finalement plus seulement l'objet, mais l'artiste lui-même qui est mis en demeure d'authentification esthétique.

Notons que l'accent mis sur le créateur plutôt que sur sa production a pour effet, parfois, d'éclipser la réalisation de tout objet. De fait, l'authenticité attachée à l'identité numérique de l'objet demande encore au moins qu'il y ait *une* chose, mais l'authenticité dégagée de l'identité numérique de l'œuvre autorise la disparition des objets. Un créateur peut donc maintenant se prétendre sans œuvre, tout en participant encore au domaine de l'art. Robert Rauschenberg demanda par exemple à Willem De Kooning un dessin, qu'il se proposa de gommer entièrement, mobilisant pour cela 40 gommes. Sa performance fut élevée au rang d'œuvre. De son côté, en 1963, Ben Vautier rédigea le certificat suivant : « Je souligné Ben Vautier déclare authentique œuvre d'art : l'absence d'art ». Avec l'art conceptuel, l'objet s'efface devant l'œuvre : c'est que, des deux, il y en a un de trop. Pour que l'œuvre advienne, il a fallu passer par la suppression de la chose matérielle. La quête d'authenticité esthétique aboutit ainsi, de manière assez paradoxale, à la disparition de tout objet esthétique. Et quand il y a encore un objet produit, c'est pour en vider toute la substance esthétique. En 1963, Robert Morris a par exemple produit une déclaration de nullité artistique : « Le souligné Robert Morris, ayant fabriqué la construction en métal intitulée LITANIES, décrite dans l'annexe A, retire par la présente de ladite construction toute qualité et contenu esthétiques, et déclare qu'à partir de la date d'aujourd'hui, ladite construction ne possède aucune qualité et contenu de cet ordre. Fait le 5 novembre 1963 ». De sorte que l'objet gagne le statut d'œuvre d'art par son statut vide d'*objet sans qualité*.

Aussi, des plasticiens qui s'affirment sans œuvre (entendu sans production d'objets matériels ou immatériels), prétendent aujourd'hui être, par eux-mêmes, leur propre production esthétique. Et parfois, l'authenticité est recherchée alors même que les intentions afférentes sont fictives. Sophie Calle, artiste s'inscrivant dans le courant de l'art narratif, a commencé sa carrière en suivant, dans la rue, des personnes inconnues. Leur arrachant quelques données biographiques lacunaires, elle les a ensuite mises en scène pas le biais de textes et de photographies. Inspirée par le roman photographique et le roman policier, Calle mélange et confond éléments réels et éléments fictionnels. Elle a ensuite appliqué ce principe à sa propre personnalité, se faisant suivre par un détective privé. Les procédés mis en place par Calle tendent ainsi à effacer la distinction entre vie privé et vie publique, entre le moi « intime » et le moi social, mais également, à rompre la démarcation entre réalité et fiction. Puisque le réel n'est finalement qu'entre-aperçu par quelques fragments discrets prenant la forme de photographies éparses commentées de quelques notes, si les œuvres de Calle proposent des récits du monde, elles ramènent aussi, et inversement, le monde à la dimension du récit. Pour sa rétrospective intitulée *Doubles-jeux*, Calle a collaboré avec l'écrivain Paul Auster, à qui elle demanda de prolonger, par le récit, la vie d'un personnage, Marie Turner, tiré du roman *Léviathan* (AUSTER, 1992). Dans le roman, Auster décrit ainsi son personnage :

> [elle] avait demandé à une amie de venir ce soir-là prendre des photos d'elle pendant son exhibition - sans intention de les montrer à qui que ce fût, juste pour elle-même, afin de satisfaire sa propre curiosité quant à son apparence. Elle se muait délibérément en objet, en image anonyme du désir, et il lui paraissait capital de comprendre avec précision en quoi consistait cet objet. (AUSTER, 1992, p. 84)

De connivence avec Auster, Calle s'appropria cet alter-ego, dont elle imita les activités, perturbant ainsi les démarcations avec son double fictionnel. L'identification de Calle croise les références au point que, selon l'artiste, Marie Turner devient une partie de Sophie Calle, voire l'inverse. Le régime autobiographique cède ainsi la place à une sorte de biographie imaginaire, qui en retour, devient l'autobiographie de Calle. Le pacte référentiel qui noue la relation de l'auteur à l'exercice de l'autobiographie n'est pas rompu, puisque les éléments fictionnels ou imaginaires sont mis en scène dans le réel et y participent ainsi. Notons aussi que le rapport à l'obscénité qu'interroge Calle n'est pas tant celui d'un voyeurisme ou d'un exibitionnisme refluant la vie intime, mais celui d'une quête de l'identité d'un moi intime, qui bien sûr, n'aboutit jamais que dans la fiction. Cette quête impossible est alors déjouée par une mise en scène qui, seule, rend à Calle une sorte d'identité personnelle, celle-ci, entièrement fictive.

6.3.3 L'inauthenticité personnelle ou l'homme *kitsch*

Si l'authenticité de l'objet esthétique est ramenée, *in fine*, à l'esthétisation de l'identité personnelle de son créateur, on peut alors, en retour, poser, toujours dans le champ esthétique, la question des personnes « inauthentiques ». Or cette question a bien été posée, notamment par Hermann Broch. C'est d'ailleurs vraisemblablement Broch qui, dans une conférence donnée en 1950 à l'université de Yale, a introduit le terme de *kitsch* dans le discours critique. Broch fait remonter l'origine du terme *kitsch* au XIXe siècle, critiquant au passage les thèses marxistes, comme celle de Benjamin, qui le corrèle au mode industriel de production en série, hypothétiquement responsable d'une aliénation des œuvres d'art. Selon Broch, cette hypothèse est contredite par le fait même que la vague du *kitsch* s'est développée même dans les pays socialistes. Pour Broch, le *kitsch* concerne moins les objets eux-mêmes que les attentions dont ils sont sujets. Plutôt que d'en faire un épiphénomène du mode de production capitaliste, Broch associe le *kitsch* à une forme de revendication identitaire. Broch parle notamment de « *kitsch-mensh* », de l'« homme du *kitsch* » :

> [...] je ne parle pas véritablement d'art, mais d'une attitude de vie déterminée. Car l'art *kitsch* ne saurait naître ni subsister s'il n'existait pas l'homme du *kitsch*, qui aime celui-ci, qui comme producteur veut en fabriquer et comme consommateur est prêt à en acheter et même à le payer bon prix. (BROCH, 2001)

Broch associe ainsi le *kitsch* à une attitude, celle que se donnait la classe bourgeoise naissante. Cette attitude, à savoir le romantisme, provient selon Broch de la résolution d'une double contrainte érotico-religieuse à laquelle était soumise la bourgeoisie. Dans son mode de vie, l'aristocratie s'était arrogée le libertinage ainsi qu'un « décor de vie de grand style, d'une beauté exubérante ». Étant en demeure de prendre la place de l'aristocratie après la révolution, la classe bourgeoise revendiquait l'appropriation des plaisirs qui étaient précédemment l'apanage de l'aristocratie. Mais dans le même temps, elle tâchait de rester fidèle aux principes d'éducation que lui conférait une tradition ascétique issue de la Réforme. En occupant la place de l'aristocratie, la bourgeoisie naissante était donc dans la nécessité d'harmoniser son désir de libertinage à son éducation ascétique puritaine. La vogue *kitsch* qui a particulièrement sévi lors de la période romantique, serait ainsi due à une exaspération provoquée par la double injonction de jouissance esthétique et de respect des vertus calvinistes :

> La nouvelle époque, précisément celle de la bourgeoisie, veut la monogamie, mais elle veut jouir par là de toutes les joies du libertinage, sous une forme encore plus concentrée si possible. L'acte sexuel du couple monogame, non seulement est porté jusqu'au astres, mais on impose aux astres, y compris toutes les éternités, la peine de descendre sur la Terre afin qu'il y naisse la plus haute volupté accessible. Le moyen d'y parvenir, ce sont les hauteurs d'imagination propres à l'exubérance. (BROCH, 2001)

Le *kitsch* serait ainsi le produit de doubles contraintes contradictoires, d'un dualisme psychologique. La contrainte à laquelle était soumise la bourgeoisie se résolvait ainsi dans une sorte d'exubérance esthétique. Selon Broch, la Réforme joua un grand rôle dans ce phénomène, en faisant passer le sacré au cœur même du profane. L'intériorisation du divin au sein de la vie profane aboutit à la vision romantique du monde encourageant « à exalter les misérables événements de la vie de tous les jours jusqu'aux sphères de l'absolu ». Or c'est par et à travers l'art que s'affranchit de manière paroxystique cette vision réformiste du divin. Elle s'achève alors en *kitsch* :

> L'ombre redoutable de la beauté divine descendue dans l'œuvre d'art, hante le XIXe siècle tout entier. (...) La déesse de la beauté dans l'art est la déesse du *kitsch*. (BROCH, 2001, p. 25-26)

L'incapacité de l'artiste à accéder sans cesse aux cimes du génie donna naissance à la catégorie du *kitsch*. Lorsque l'artiste ne parvient pas à combler cette demande inflationniste, son œuvre tombe alors dans une catégorie intermédiaire. Ainsi, pour Broch, le *kitsch* provient de la tension romantique exaspérant le modèle de l'œuvre sacralisée, transmise par l'esprit réformiste à la classe bourgeoise.

Si, selon Broch, la notion de *kitsch* doit être reliée à la recherche que la bourgeoisie fait de sa propre identité de classe et de sa représentation sociale, Kundera stigmatise l'attitude de l'homme *kitsch* qui, d'après lui, culminerait dans le double-psychologique par le rejet des choses vulgaires, c'est-dire communes, triviales. L'homme *kitsch* est un homme qui renonce aux vérités du commun, à commencer par la réalité de la merde :

> Le désaccord avec la merde est métaphysique. L'instant de la défécation est la preuve quotidienne du caractère inacceptable de la Création. De deux choses l'une : ou bien la merde est acceptable (alors ne vous enfermez pas à clé dans les waters !), ou bien la manière dont on nous a créés est inadmissible. Il s'ensuit que l'accord catégorique avec l'être a pour idéal esthétique un monde où la merde est niée et où chacun se comporte comme si elle n'existait pas. Cet idéal esthétique s'appelle le *kitsch*. C'est un mot allemand qui est apparu au milieu du XIXe siècle sentimental et qui s'est ensuite répandu dans toutes les langues. Mais l'utilisation fréquente qui en est faite a gommé sa valeur métaphysique originelle, à savoir : le *kitsch*, par essence, est la négation absolue de la merde ; au sens littéral comme au sens figuré : le *kitsch* exclut de son champ de vision tout ce que l'existence humaine a d'essentiellement inacceptable. (KUNDERA, 1987, p. 311–312)

La dégradation morale de l'individu *kitsch* fait également l'enjeu d'une critique acerbe chez Philip K. Dick, anticipant la notion de post-vérité par le jeux du simulacre qui en vient à contaminer la spécificité même des êtres humains :

> Le bombardement de pseudo-réalités commence à produire très vite des êtres humains non authentiques, des simulacres d'êtres humains - aussi faux que les données qui les assaillent de toutes parts. [...] De fausses réalités engendrent de faux êtres humains. Ou de faux êtres humains créent de fausses réalités qu'ils vendent ensuite à d'autres humains, les transformant, au bout du compte, en contre-façon d'eux-mêmes. Et nous nous retrouvons donc avec de faux humains inventant de fausses réalités, puis les refilant à d'autres faux êtres humains. C'est juste une immense variante de Disneyland. On peut faire un tour avec les pirates, voir le simulacre de Lincoln ou monter dans le Mr Toad's Wild Ride - on peut *tout* faire, mais rien n'est vrai. (DICK, 1987, p. 17)

On sait d'ailleurs l'importance que revêt chez Dick la question d'une possible indiscernabilité entre une machine et un être humain, problématique qu'il exploite longuement dans *Les androïdes rêvent-ils de moutons éléctriques ?*

6.4 L'œuvre d'art *Gesamtkunstwerk* : l'automate parfait

Œuvre parmi les œuvres, la quête ultime de l'artiste n'est-elle pas la création d'un être vivant ? En témoignent les mythes grecs de Prométhée et de Galatée, la légende talmudique du Golem, le roman *Frankenstein* (1818) de Mary Shelley (SHELLEY, 2008), le *Pinocchio* (1881) de Carlo Collodi (COLLODI, 2003), *L'Ève Future* (1886) de Villiers de L'Isle-Adam (L'ISLE-ADAM, 1993), parmi bien d'autres fictions encore. Copier la créature humaine, la reproduire par des voies différentes de la seule reproduction sexuée est sans doute le rêve de tout artiste, elle figure une œuvre totale, une *Gesamtkunstwerk*, du moins si l'on considère avec les Grecs antiques que le mythe de Pygmalion serait à l'origine de toute production artistique [19].

Nous pourrions illustrer la problématique que je souhaite développer maintenant par une comparaison, à quelques siècles de distance, entre l'expérience de pensée des automates que propose Descartes dans *Les médiations métaphysiques*, et celle d'Alan Turing dans *Les ordinateurs et l'intelligence*. Chez Descartes, les automates parfaits sont des imitations mécaniques d'organismes vivants, des répliques matérielles d'êtres vivants. Chez Turing, l'automate se rapporte à la simulation de la pensée dans des jeux de dissimulation.

Le problème des indiscernables esthétiques doit ainsi, à mon sens, être aussi corrélé à la question plus générale des indiscernables spécifiques, notamment au cas de deux entités indiscernables dont l'une serait qualifiée d'humain et l'autre pas. Descartes imagine le cas d'un être qui serait un automate reproduisant parfaitement le corps humain :

> (...) je me contentai de supposer que Dieu formât le corps d'un homme entièrement semblable à l'un des nôtres, (...) sans mettre en lui au commencement aucune âme raisonnable, ni aucune autre chose pour y servir d'âme végétante ou animale. (...) Examinant les fonctions qui en suite de cela pouvaient être en ce corps, j'y trouvais exactement toutes celles qui peuvent être en nous sans que nous y pensions. (DESCARTES, 1953a, p. V)

[19] Parfois même, à travers l'art, c'est la transgression de la spécificité humaine qui est recherchée, comme l'indiquent les quêtes transhumanistes de Stelarc et d'Orlan.

Un tel corps serait, dit Descartes, comparable à une machine. Mais qu'est-ce qu'une machine ? Selon Descartes, c'est ce qui est capable de communiquer du mouvement par déformation des parties constitutives. Les mâchoires sont des tenailles, l'estomac est une cornue, les veines sont des tubes hydrauliques, le cœur est un ressort, les poumons sont des soufflets, les muscles sont des cordes, etc. L'analyse cartésienne du corps humain correspond à la mécanique d'une machine dont les parties sont animées par des mouvements de rotation et de translation. C'est donc la géométrie qui doit guider la description des corps, et l'*Essai sur la dioptrique* témoigne d'ailleurs d'une telle analyse en faisant correspondre la perception interne du sujet à l'externalité de la substance étendue par la médiation d'une correspondance mathématique.

Descartes propose donc de faire l'expérience de pensée d'une machine qui serait une réplique parfaite d'un corps organique, selon le modèle d'un automate idéal dont la sophistication mécanique atteindrait une perfection identique à celle de la complexité du corps humain. Se pose alors la question de la possibilité de distinguer une telle machine d'un organisme véritable. Pour ceci, il recourt à la médiation des animaux qui, étant dépourvus d'âme rationnelle, sont conséquemment de simples corps-machines organiques comparables à des automates mécaniques :

> Et je m'étais ici particulièrement arrêté à faire voir que, s'il y avait de telles machines qui eussent les organes et la figure d'un singe ou de quelque autre animal sans raison, nous n'aurions aucun moyen pour reconnaître qu'elles ne seraient pas en tout de même nature que ces animaux (...) (DESCARTES, 1953a, p.)

En revanche, Descartes détermine deux moyens pour discriminer l'automate mimant l'humain de son modèle authentique. Le premier moyen relève de l'usage du langage sensé qui est prérogative humaine. La limite de l'imitation chez Descartes correspond à celle du langage : car l'automate sera peut-être capable, par le développement d'une technologie idoine, d'imiter la physionomie humaine et ses mouvements propres, et même de produire des sons voire des mots et des phrases, mais il buterait néanmoins sur l'incapacité à exprimer une pensée. Car c'est moins la possibilité du langage qui est interdite à la machine cartésienne que la possibilité d'exprimer une pensée propre. Ainsi, de telles machines :

> (...) ne pourraient user de paroles ni d'autres signes les composant, comme nous faisons pour déclarer aux autres nos pensées. (...) (DESCARTES, 1953a)

La seconde manière dont Descartes discrimine la machine de l'humain consiste à admettre que des répliques mécaniques d'humains n'agiraient pas par connaissance consciente, mais par la seule disposition de leur complexion mécanique. Les imitations, corps-machines sans pensées, seraient certes compétentes dans l'accomplissement de certaines tâches. Elles les exécuteraient même mieux que tout homme grâce à leur degré de spécialisation, ce qui contribuerait déjà à les distinguer. Par ailleurs, ce gain d'efficacité aurait pour effet inévitable la perte de plasticité humaine dans l'exécution des tâches. Cette faculté d'adaptation proprement humaine est donc ce qui, a priori, différencie l'humain de la machine : celle-ci est bien capable de produire de meilleurs effets, mais uniquement sur un nombre restreint d'actions. La raison, alliée au corps, permet quant à elle une malléabilité plus importante [20] :

[20] C'est un argument similaire qu'Edgar Poe utilise afin de révéler la supercherie du joueur d'échec de Maëlzel. Selon

> Car au lieu que la raison est un instrument universel qui peut servir en toutes sortes de rencontres, ces organes ont besoin de quelques particulières disposition pour chaque action particulière ; d'où vient qu'il est moralement impossible qu'il y en ait assez de divers en une machine, pour la faire agir en toutes les occurrences de la vie de même façon que notre raison nous fait agir. (DESCARTES, 1953a)

Si les automates cartésiens peuvent imiter les corps-machines des animaux en répliquant leurs mouvements, ils ne sont néanmoins pas capable de créer leur mouvement premier. Car la production de ce mouvement exige l'action extérieure d'un moteur apporté par l'ingénieur-concepteur pour les artefacts de la machine (ou par le démiurge pour les animaux-machines). Insistons ici sur le fait que Descartes n'est pas à l'origine de la comparaison entre les artefacts et les êtres vivants. Platon compare, dans le Timée, les mouvements des vertèbres à ceux des charnières, et Aristote constate des analogies entre la complexion musculaire et les mécanismes des machines de siège comme les catapultes, et définit l'esclave comme une machine animée. L'originalité de Descartes trouve sa source ailleurs, précisément dans le paradigme du mécanisme qui émerge au XVIe siècle, accompagnant l'invention des automates. A côté des théories contemporaines mécanicistes, c'est la fascination qu'éprouve Descartes pour une technique qui lui préexiste, celle des horloges et des jeux d'eau, qui est responsable de l'émergence de sa théorie mécaniste. Alors que, dans l'Antiquité, l'apport de l'énergie indispensable à l'action des outils nécessitait encore la proximité directe avec leur source motrice vivante, au siècle de Descartes, l'utilisation du ressort permet d'atermoyer la libération de cette énergie. Cette dernière est d'abord stockée provisoirement puis restituée ultérieurement à la machine. Autrement dit, elle ne passe plus directement du muscle à l'artefact. Cela a pour effet de simuler une forme d'autonomie de la machine par rapport à sa source motrice. C'est ici qu'intervient la notion de simulation chez Descartes. L'imitation ne consiste de fait qu'à mimer l'apparence, la physiologie de l'organisation externe et interne du corps humain. La simulation, quant à elle, vise à mettre sous le regard de l'observateur une capacité feinte : celle d'être autonome, d'être responsable de son propre mouvement.

C'est donc à la fois la question du moteur premier et l'enjeu des conditions initiales qui sont posés ici. En effet, l'automate n'est capable de s'animer et de développer un mouvement qu'en fonction de données mécaniques initiales. L'automate n'est pas encore conçu comme une machine cybernétique, c'est-à-dire une machine capable de réajuster par rétroaction ses actions à la variabilité des conditions de son environnement. L'automate cartésien est par conséquent dépendant des conditions initiales de sa fabrication. Si les corps-machines sont des analogues d'automates, ils doivent eux-mêmes être tributaires de données initiales qui les contraignent : données initiales qui sont celles du constructeur, à savoir, pour Descartes, un démiurge. La pensée cartésienne du vivant est donc une pensée statique tributaire de conditions initiales, la cause finale étant déplacée vers l'origine du mouvement. C'est ce en quoi échoue la simulation.

On remarquera que l'expérience de pensée proposée par Alan Turing dans l'introduction de son article *Les ordinateurs et l'intelligence*, et dont l'objet est de répondre à la question de savoir si

---

Poe, le caractère statique des automates, dont le fonctionnement est entièrement déterminé par des conditions initiales, est absolument contraire à la dynamique d'une partie d'échec dont les coups successifs n'obéissent pas à de strictes lois déterministes mais relèvent au contraire d'une imprévisibilité liée aux aléas de la partie. Poe en conclut que la machine de Maëlzel contient en réalité un véritable homme, seul capable de réagir avec une plasticité suffisante aux obligations de la partie. Les machines de Descartes ou de Poe ne connaissent pas l'effet retro-actif qui caractérise la cybernétique.

les machines peuvent penser, suit le mécanisme inverse de celle de Descartes. Le cas présenté par Turing consiste à demander si une machine serait éventuellement capable de tromper un homme en lui communiquant des informations visant à simuler l'appartenance à l'espèce humaine. Le test de Turing, ou jeu de l'imitation, est décrit ainsi par son auteur : un homme A et une femme B sont isolés d'un interrogateur C. Ce dernier soumet les individus A et B à des questions visant à identifier leur appartenance sexuelle. L'homme essaie de se faire passer pour une femme, tandis que la femme essaie de se faire reconnaître comme femme. Il est supposé par ailleurs que C obtienne les réponses de A et B par la médiation d'un téléimprimeur. Turing demande alors ce qui se passerait si l'on substituait à la personne A une machine capable de répondre aux questions posées par C. Une machine conçue à cet effet serait-elle capable de tromper l'interrogateur ?

Contrairement à celle de Descartes, l'expérience de pensée de Turing contribue à refouler les corps organiques hors du champ d'étude de la pensée de deux façons distinctes. Premièrement par le fait que les protagonistes du test sont isolés les uns des autres. Deuxièmement par le fait qu'une machine vient remplacer l'organisme humain. Le protocole du test de Turing vise donc à faire disparaître les corps, là où, inversement, la célèbre analyse de Descartes reposait sur l'investigation d'un corps-machine atrophié de sa faculté de penser. La stratégie de Turing consiste à extraire la pensée hors des corps, sans tenir compte de l'organicité spécifique du corps humain. Ce n'est pas l'aspect de la complexité mécanique simulant l'apparence humaine qui doit être objet de fascination.

Dans la liste des objections qui peuvent être adressées au test de Turing, celui-ci relève le problème de la pensée consciente. La seule manière dont on pourrait s'assurer qu'une machine pense serait d'être la machine et de ressentir qu'on pense. On pourrait alors décrire ces sentiments au monde, mais bien sûr personne n'aurait de raisons d'en tenir compte. De même, suivant ce point de vue, la seule manière de savoir si un homme pense est d'être cet homme lui-même. Nous ne pourrons jamais être capables de savoir si une machine sait qu'elle pense, puisque pour cela, il faudrait être « à la place » de cette machine. Pour éviter les conséquences d'une philosophie solipsiste aboutissant à des apories inéluctables, Turing propose de réduire méthodologiquement la pensée à sa manifestation (TURING, 1995). Il rejette donc le problème de la conscience. Cela constitue évidemment une différence majeure par rapport à la méthode du cogito cartésien, et c'est au fond, peut-être déjà ce que Hobbes objectait à Descartes : la pensée n'est-elle pas, au fond, seulement du calcul ? Or, qu'est-ce que calculer ? L'observation d'un individu en train de calculer révèle qu'il prend certains états successifs, qu'il lit, qu'il déplace sa main, qu'il écrit, ce qui correspond, souligne Turing, au fonctionnement même de la machine de Turing. La complexité organique du corps qui est l'hôte cartésien de la substance raisonnante est donc remplacée, chez Turing, par la simplicité mécanique de la machine effectuant des opérations.

Parmi les objections qui peuvent être retenues contre l'idée de machine pensante, Turing relève aussi le problème de la liberté de volonté indiqué par Ada Lovelace : les machines ne sont capables de produire que ce qu'on leur a demandé de faire, elles n'ont pas de capacité à l'initiative. Pour Turing, la question est en fait de savoir si les machines sont ou non capables de surprendre, et y répond par l'affirmative. Car la surprise, dit Turing en substance, procède de l'oubli des causes.

Nous avons vu que Descartes relevait trois causes impliquant l'impossibilité de machines pensantes : leur incapacité à l'expression par le langage, leur dépendance à un mouvement princeps et enfin les conditions initiales de leur mise en fonctionnement, qui régissent rigidement leur com-

portement. Comparons ceci aux problèmes soulevés par Turing. Premièrement, constatons que le test de Turing procède d'un leurre au moyen du langage. Les machines de Turing sont donc des machines qui communiquent, et qui communiquent de façon à réagir aux sollicitations de l'intervenant, en s'y adaptant afin de produire des réponses adaptées. Deuxièmement, la problématique du mouvement premier est identifiable à l'hypothèse du ruban infini sur lequel écrit la machine. Infini, c'est-à-dire sans commencement. On peut donc deviner que la question du commencement est pour Turing un pseudo-problème, et qu'il est possible de le tenir à l'écart de la réflexion sur ce qui constitue la simulation. Enfin, troisièmement, le problème des conditions initiales est résolu par la distinction entre matériel et logiciel. La machine n'est plus astreinte à un programme singulier déterminant entièrement son fonctionnement, elle est dotée au contraire d'une habilité à la programmation qui permet de réajuster à loisir son rayon d'action, et qui, *in fine*, la rend *universelle*. Turing libère le programme de son joug à la machine. Le logiciel n'est plus gravé dans la mécanique à la manière dont la substance pensante est complètement liée au au corps physique.

Les entreprises respectives de Descartes et Turing mobilisent la même sorte de méthodologie heuristique : tous deux commencent par dissocier l'esprit du corps et médiatisent leurs comparaisons par le tiers de la machine. Mais là où la pensée cartésienne n'est pas une production due au corps-machine, chez Turing, c'est au contraire bien la machine qui produit le calcul.



Nous avons rapproché, via notre discussion sur le *kitsch*, la notion d'authenticité ontologique des œuvres d'art à celle des personnes humaines. Comme nous l'avons souligné, le discours sur l'authenticité a progressivement migré, dans la critique des arts plastiques, de l'objet d'art singulier à son producteur, de la toile à l'artiste. La notion romantique d'œuvre d'art totale conduit à l'idée que l'artiste est l'œuvre, et qu'il est donc directement le sujet du jugement d'authenticité : la question de la reconnaissance artistique est celui de la manière dont un individu possède ou gagne le statut d'artiste. Avec l'œuvre idéale, on passe donc d'une question portant sur l'identité numérique d'un particulier concret (est-ce le même objet ?) à une question portant sur l'identité spécifique (est-ce un artiste authentique ?).

Or si l'on peut juger un artiste d'être plus ou moins authentique, le soupçon d'inauthenticité peut alors être porté, de manière plus générale, aux espèces, et entre autres, à l'espèce humaine. L'automate parfait mime l'apparence humaine (Descartes) ou en simule la pensée (Turing). Les automates cartésiens, répliques matérielles d'êtres humains, tombent nécessairement dans le domaine des contrefaçons : des êtres qui, sous des apparences humaines, ne possèdent pas l'essence proprement humaine du « moi ». Œuvre d'art idéale, l'automate parfait reste chez Descartes une contrefaçon discernable de son modèle, cela même en dépit de ses similarités qualitatives, parce que la machine ne sera jamais capable de produire des pensées. En revanche, chez Turing, l'authenticité ne tient plus à un discours métaphysique sur la nature essentielle de la pensée, mais sur un acte de *reconnaissance*, lié à l'incapacité de l'expérimentateur du jeu de la simulation à distinguer les réponses de la machine de celles de l'homme. Les automates parfaits cartésiens sont des répliques matérielles, mais des contrefaçons d'être humains. Les machines pensantes de Turing sont, quant à elles, des dupliques d'humains et ce que, du point de vue des réponses qu'elles fournissent aux questions, elles sont fonctionnellement indiscernables d'êtres humains.

6.5 Conclusion : authenticité et reconnaissance

Pour résumer ce chapitre, je rappellerai la distinction qui a été faite plus haut entre originalité et authenticité. L'originalité est toujours factuelle, elle désigne une primeur dans le temps, le fait qu'une chose advienne *avant* une autre. L'authenticité est en revanche une notion axiologique, donc attribuée de manière conventionnelle, un statut particulier accordé aux entités pour des raisons qui peuvent s'avérer diverses et plus ou moins justifiées. Il est vrai que l'originalité peut servir parfois de critère d'authenticité, notamment dans les beaux-arts où une entité originale est souvent considérée comme authentique. Mais cela n'est pas du tout une nécessité. Une occurrence peut s'avérer être originale sans être authentique ou authentique sans être originale, voire ni originale, ni authentique. Contrairement à l'originalité qui peut s'avérer relative au sens où une réplique peut être plus ou moins séparée de son modèle dans le temps, l'authenticité est généralement considérée comme absolue. On dit en effet d'un tableau qu'il est un « faux » ou un « vrai », en revanche, on ne lui accorde généralement pas de statut intermédiaire. On ne dit pas qu'il est « plus ou moins faux ». On parlera peut-être de tableau « attribué à » pour désigner une œuvre dont l'auteur ne fait pas consensus et qui reste en attente d'identification du fait d'une connaissance imparfaite au sujet d'une chose dont l'authenticité reste, du point de vue absolu, toujours déterminable.

Ainsi, ce que tendent à montrer les affirmations précédentes est que l'authenticité ne relève ni de l'identité numérique, ni de la ressemblance parfaite. Elle est un acte de reconnaissance sociale basé sur des contextes culturels et historiques. Elle relève d'une identification intentionnelle, et non d'une identification physique, qui ressortit, quant à elle, à l'originalité. La reconnaissance est le fruit d'une discussion collective, d'un consensus social médiatisé par un ensemble de conventions et un enracinement contextuel dans le lieu et l'histoire.

L'interrogation sur l'authenticité des œuvres nous a mené à l'interrogation sur l'identité spécifique. De là, c'est la question de l'identité personnelle qui se présente maintenant. Dans le prochain chapitre, je souhaite montrer que la notion d'identité personnelle est fortement ancrée sur la reconnaissance, qu'elle relève d'un enjeu basé sur le modèle de la simulation : l'identité personnelle est une identité fictionnelle. Le travail de Sophie Calle illustre une tentative visant à faire coïncider la personne de l'artiste avec celle d'un « automate » fictionnel, d'un personnage fictif. L'identité personnelle peut ainsi être considérée comme une sorte de création *sui generis* de soi par soi, accréditée (ou non) par une authentification, une reconnaissance sociale. Elle retourne donc d'une métalepse entre personne réelle et personnage [21]. Cette association entre personne fictive et personne réelle interroge la question de la présumée réalité du soi.

[21] C'est aussi le geste d'Elaine Sturtevant, qui, dans une performance filmée, interprète le rôle de Joseph Beuys performant lui-même une évasion du gangster américain John Dillinger.

Chapitre 7

Fiction de l'identité personnelle et identité fictionnelle de la personne

Why Can't I Be You ?

The Cure, « Why Can't I Be You », in *Kiss me, Kiss me, Kiss me*, 1987

I am just a copy of a copy of a copy
Everything I say has come before
Assembled into something into something into something
I don't know for certain anymore
I am just a shadow of a shadow of a shadow
Always trying to catch up with myself
I am just an echo of an echo of an echo
Listening to someone's cry for help

Nine Inch Nails, « Copy of A », *Hesitation Marks*, Columbia Records, 2013

Soyez vous-mêmes, tous les autres sont déjà pris.

Oscar Wilde

Si les personnages d'une fiction peuvent être lecteurs ou spectateurs, nous, leurs lecteurs ou spectateurs, pouvons être des personnages fictifs.

Jorge Luis Borges, *Magie partielle du* Quichotte

L'imaginaire du double fictionnel doit maintenant être complété par l'imaginaire de l'identité personnelle, moment où les notions de personnage et de personne viennent à se brouiller. L'étymologie de « personne » devrait d'ailleurs aiguiller notre attention sur son origine fictionnelle : « persona », en latin, désigne le masque que l'acteur de théâtre revêt pour emprunter la

personnalité de son personnage en dissimulant ses traits. Dans l'introduction à son essai *Personne, Personnage*, Christian Dours écrit qu' « on peut attendre de la question philosophique de l'identité personnelle qu'elle interroge la notion de personnage » (DOURS, 2003, p. 9). Dans ce qui suit, mon objectif sera contraire à cette proposition. Car il me semble que c'est la notion de personnage qui doit interroger celle d'identité personnelle. Précisément, ma thèse est que l'identité personnelle est une construction logique projetée sur un corps matériel. Nous pouvons dire « je » parce que nous nous identifions, à la première personne, comme des personnages. Comme le dit Montaigne dans un passage bien connu des *Essais* : « Notre fait, ce ne sont que des pièces rapportées. » (MONTAIGNE, 1928a, livre II, chap.1).

Ce n'est pas seulement dans des fictions fantastiques que des individus en viennent à revendiquer la même identité. Rappelons le cas du procès de Martin Guerre, dont Montaigne fut témoin dans sa jeunesse et qu'il évoque dans les *Essais* : « Je vis en mon enfance, un procès que Corras Conseiller de Toulouse fit imprimer, d'un accident étrange ; de deux hommes, qui se présentaient l'un pour l'autre. » (MONTAIGNE, 1928b, livre III, ch. 11) En voici les circonstances. En 1548, un habitant de l'Ariège nom de Martin Guerre, alors âgé de 24 ans, abandonna femme et enfant. Bertrande, son épouse, vécut durant des années sans son mari, demeurant dans l'impossibilité d'en prendre un autre parce qu'elle ne pouvait démontrer la mort du précédent. Cependant, huit années plus tard, en 1556, un homme vint, prétendant être le fugitif, bien qu'il parût plus trapu et moins grand que lui. L'individu écarta les soupçons d'usurpation en prouvant qu'il avait souvenir de sa vie maritale. En outre, il était plus besogneux et se montrait plus affectueux que son prédécesseur, ce qui acheva de convaincre celle qu'il prétendait être sa femme. Certains habitants accusèrent toutefois le revenant d'usurpation d'identité, affirmant reconnaître en lui un certain Arnaud du Thil. Un procès fut mené, durant lequel l'accusé maintint être l'authentique Martin Guerre. L'affaire n'était pas claire, et il apparut encore un autre prétendant revendiquant la même identité durant le procès. Le jury eut bien du mal à démêler l'écheveau des personnalités. Un verdict fut toutefois prononcé. On reconnut le revenant coupable d'usurpation d'identité, d'imposture et d'adultère, et il fut condamné à la pendaison. Dans les *Essais*, Montaigne déplore l'arrêt de la Cour en raison que, pour que le doute ne subsista pas, il fallut que le juge prêta à l'accusé une extraordinaire capacité à l'imitation : « [Corras] me sembla avoir rendu l'imposture de celui qu'il jugea coupable, si merveilleuse et excédant de si loin notre connaissance, et la sienne, qui était juge, que je trouvai beaucoup de hardiesse en l'arrêt qui l'avait condamné à être pendu. Recevons quelque forme d'arrêt qui dise : La Cour n'y entend rien » (MONTAIGNE, 1928b, livre III, chap. 11)

La réponse à donner à certaines questions concernant l'identité d'un individu dépend effectivement, parfois, d'un arrêt, c'est-à-dire, d'une décision. Nous avons par exemple vu que la théorie du meilleur candidat permet de déterminer de manière contextuelle si une chose comme un bateau a persisté entre deux instants. Et nous avons vu aussi que l'identité d'un personnage de fiction dépend des décisions collectives des récepteurs de la fiction. Cependant, lorsque la question « est-ce le même ? » est posée à propos d'individus biologiques, et en particulier à propos de *personnes*, il peut sembler naturel que la réponse ne soit pas seulement contextuelle. Certains philosophes présument ainsi qu'il doit exister une réponse absolue à la question de savoir si la carrière d'un même individu humain s'étend bien entre deux instants donnés. Derek Parfit a donné une interprétation de cette croyance selon le principe suivant : « quoi qu'il se passe dans le laps de temps qui me sépare maintenant d'un instant futur, soit j'existerai encore, soit je n'existerai plus. Soit

une expérience future quelconque sera *mon* expérience, soit elle ne le sera pas. » (PARFIT, 2015) L'identité des personnes à travers le temps peut-elle trouver une réponse absolue ? Nous avons ici un problème lié à l'individuation des personnes.

Lorsque nous comptons le nombre de personnes présentes autour d'une table, nous recourons au dénombrement des corps physiques qui sont attablés. Une personne est identifiable par la présence d'un corps, tout comme une fourchette est repérable par ses limites matérielles dans l'espace. Les corps sont des particuliers faits de matière. Lorsqu'un corps humain s'alimente ou lorsqu'il inspire de l'oxygène, certaines molécules qui entraient dans la constitution d'autres individus deviennent partie intégrante du corps. Ces molécules finiront ensuite par être rejetées par déjection, transpiration, desquamation, ou expiration. En ce sens, le corps d'une personne peut donc être vu comme une sorte de bateau de Thésée : il est fort probable qu'à la vieillesse, il ne contienne plus aucune des molécules qui le constituaient à la naissance.

Les corps humains sont constitués de la matière qui compose leur corps animal qui s'étend jusqu'aux limites de l'épiderme. Mais certains philosophes affirment que la frontière des personnes doit être avancée, ramenée aux limites du seul cerveau. Pour d'autres, elle doit au contraire être reculée, car elle doit encore inclure les substances matérielles rejetées, comme les excréments et les phéromones. D'autres considèrent que le corps s'étend encore plus loin, et que la répartition des frontières doit aller aussi loin que l'autorisent les manifestations de la personne par le biais des technologies de la communication.

Quoi qu'il en soit, ces remarques indiquent que pour le métaphysicien, la question de l'identité personnelle doit ainsi être articulée à une seconde, à savoir de déterminer ce qui constitue la nature d'une personne. On peut lister quatre conceptions défendues :

1. La première conception affirme qu'une personne est une substance immatérielle, un *ego* cartésien, un moi (au sens de substance), ou une âme.

2. La seconde conception, qui identifie la personne à une substance matérielle peut être déclinée en deux versions :

   (a) une personne est un organisme vivant, un animal humain ;

   (b) une personne est seulement une petite partie d'un animal humain, à savoir, son cerveau.

3. La troisième conception, mixte, soutient qu'une personne est l'union indivisible d'une entité immatérielle et d'un corps.

4. Enfin, la dernière conception, anti-substantialiste, rejette toutes ces possibilités en soutenant qu'une personne n'est rien d'autre qu'une construction logique, qu'un point de vue sur le monde.

Une personne est-elle un corps ou occupe-t-elle un corps ? Coïncide-t-elle ou non avec le corps ? Pour tenter de répondre à ces questions, les philosophes ont, depuis Descartes, développé des expériences de pensée bâties sur des amputations plus ou moins sévères, sur des greffes de cerveaux et sur des téléportations dignes de scénarios de science-fiction. Le rapport à l'identité personnelle peut y être caractérisé par le rôle joué par deux sortes de doubles. Les uns, les doubles

ontologiques, apparaissent essentiellement dans les théories dualistes esprit-matière qui prennent au sérieux la notion d'identité personnelle. Les autres, les doubles épistémiques, apparaissent essentiellement dans les théories qui considèrent l'identité personnelle comme une fiction.

Parmi les options précédentes, c'est la quatrième possibilité que je défendrai. La question centrale du moi n'est sans doute rien d'autre que celle de savoir ce qui détermine à se considérer, à se voir, à se raconter, à s'imaginer comme une personne. On peut se voir comme ayant un tempérament égal ou comme colérique. On peut se voir comme conciliateur ou comme intrigant. Dans ces cas, nous nous attribuons, à tort ou à raison, des qualités qui contribuent à façonner l'imaginaire de notre moi. Il s'agit ainsi de souscrire à ce que Clément Rosset écrit dans *Loin de Moi* :

> [...] j'ai toujours tenu l'identité sociale pour la seule identité réelle ; et l'autre, la prétendue identité personnelle, pour une illusion totale autant que tenace, puisqu'elle est tenue par le plus grand nombre pour être au contraire la seule identité réelle, suivant ici plutôt le sentiment de Rousseau dont la raison a achevé de se perdre dans la recherche éperdue de cette identité fantomatique. (ROSSET, 1999, p. 11)

Comme Rosset, je souscris ainsi à l'idée qu'une personne n'a d'autre identité que sociale, que représentative. Comme le constate encore Rosset, si l'on dénomme par « identitaire » ce qui relève justement de l'identité sociale, le Moi-vrai, le Moi-réel, bref, tout ce qui évoquerait une nature « profonde » de l'identité personnelle opposée au caractère conventionnel de ce qui peut être décrit de manière purement sociale devrait alors être considéré comme une mystérieuse instance « pré-identitaire » (ROSSET, 1999, p. 12). Or l'hypothèse d'une telle duplication ne semble pas nécessaire. Je pars donc du principe que l'identité personnelle, au sens métaphysique du terme, est une illusion, que l'identité est toujours fictionnelle en ce qu'elle relève d'une construction logique après-coup, là où les corps matériels sont toujours donnés d'emblée.

Aussi, dans ce chapitre, il s'agira en particulier de critiquer les théories substantialistes du moi et de fournir des arguments pour montrer que ce dernier est une instance imaginaire construite de manière sociale. Le plan du chapitre est le suivant : dans la section 1, nous exposerons et critiquerons les théories substantialistes du moi. La section 2 est dévolue aux théories psychologiques du moi. Dans la section 3, nous examinerons les théories sociologiques du moi qui tendent à montrer que le moi consiste en l'occupation d'un rôle par la médiation d'un alter-ego. Dans les deux dernières sections, il s'agira de considérer plus précisément certains actes mentaux afférents aux doubles épistémiques : la section 4 est consacrée à l'étude des phénomènes d'autoscopie et du reflet spéculaire, et la section 5 est consacrée aux pensées *de se* de duplications imaginaires.

7.1 Les théories de l'identité personnelle ontologiquement engagée

Nous commencerons donc par évoquer les théories substantialistes du moi. La critique que nous porterons visera à montrer que ces théories, confrontées aux paradoxes de l'identité, nécessitent de stipuler des doubles ontologiques.

7.1.1 Le dualisme cartésien

On peut sans doute considérer que, parmi les systèmes métaphysiques, le dualisme confère d'emblée une place de choix au double, et qu'il relève d'une sorte de généralisation de la duplication. En effet, ce type de métaphysique tend à dissocier ce qui existe selon deux natures incompatibles, qu'il s'agisse, chez Platon, d'une désolidarisation entre le monde des idées intelligibles et le monde sensible, ou qu'il s'agisse, chez Descartes, d'un désancrage de la substance pensante à de la substance étendue, d'une coupure ontologique entre animal-machine et soi immatériel. Dans cette partie, je souhaite m'attarder plus particulièrement sur les raisons philosophiques qui poussent certaines théories de l'identité personnelle à souscrire au dualisme métaphysique. Il me semble que la raison principale est l'adhésion à un raisonnement que je nommerai *argument de désolidarisation* (AD) et qu'il s'agira d'expliciter afin, ensuite, de le contredire.

L'argument de désolidarisation

Chez Descartes, la théorie du substantialisme immatérialiste de l'identité personnelle a pour point de départ l'expérience de pensée consistant à douter de l'existence de tout ce qui nous entoure. Ce doute quant à l'existence des choses, Descartes l'aggrave encore en stipulant la possibilité d'un génie trompeur. L'auteur des *Méditations* en tire deux résultats bien connus, qu'on commentera dans ce qui suit.

Le premier résultat est que le sujet du doute ne peut être un corps matériel. En effet dans la première des *Méditations métaphysiques*, alors qu'il se demande de quoi il peut être absolument certain, Descartes commence par douter de l'existence de son propre corps, comme s'il s'agissait d'une hallucination sans fondement réel. Le cas clinique des membres fantômes, connu par Descartes, l'incite à envisager l'hypothèse inouïe que l'existence du corps pourrait au fond n'être qu'une illusion. Pour s'en convaincre, il entame une série d'auto-amputations imaginaires de plus en plus conséquentes. « Je me considérais moi-même comme n'ayant point de mains, point d'yeux, point de chair, point de sang, comme n'ayant aucun sens, mais croyant faussement avoir toutes ces choses » (DESCARTES, 1953a, p. 272) À quel stade de l'opération finirais-je par disparaître ? demande Descartes. La mise en morceau imaginaire du corps ne s'achève qu'avec sa disparition totale. Le sujet cartésien survit donc à la suppression de l'intégralité de son corps. Le « je », conclut Descartes, n'est donc pas le corps : « [...] examinant avec attention ce que j'étais, [...] je pouvais feindre que je n'avais aucun corps, et qu'il n'y avait aucun monde, ni aucun lieu où je fusse [...] » (DESCARTES, 1953a, p. 148). Le « je » cartésien n'est ni d'ici ni d'ailleurs, il n'est en aucun monde, il est hors-monde, c'est un « horla ».

Le second résultat notoire que Descartes tire de son expérience de pensée est qu'un soupçon universel ne pourrait rien contre la certitude que chacun peut avoir de sa propre existence à travers l'acte de penser. L'étendue des exactions du malin génie s'échinant à constamment tromper nos jugements et nos perceptions est bornée par la limite du sentiment de soi. L'ultime vérité cartésienne est donc une connaissance réflexive : « qu'il me trompe tant qu'il voudra, il ne saurait jamais faire que je ne sois rien, tant que je penserai être quelque chose » (DESCARTES, 1953c). Alors que tout ce qui est donné aux sens peut être tenu pour suspect, en revanche, pendant que *je* fais l'expérience du doute m'apparaît la certitude apodictique qu'un *je* doute : « Il faut conclure et tenir pour constant que cette proposition : *Je suis, j'existe*, est nécessairement vraie, toutes les fois

que je la prononce, ou que je la conçois en mon esprit » (DESCARTES, 1953c, p. 275). Puisque nous connaissons une substance par ses propriétés, celle qui nous est le mieux connue, en déduit Descartes, est le moi, car elle nous est donnée directement par la pensée réfléchie.

Reste alors à déterminer la nature véritable de ce « je » qui résiste au doute hyperbolique. Deux arguments viennent étayer l'idée que ce « je » est une substance immatérielle. Le premier argument relève de la limite de dissection : je peux penser exister sans corps, mais je ne peux pas exister sans penser. Même sans corps, dit Descartes,

> [...] je ne pouvais pas feindre, pour cela, que je n'étais point ; et qu'au contraire, de cela même que je pensais à douter de la vérité des autres choses, il suivait très évidemment et très certainement que j'étais ; au lieu que, si j'eusse seulement cessé de penser, encore que tout le reste de ce que j'avais jamais imaginé eût été vrai, je n'avais aucune raison de croire que j'eusse été : je connus de là que j'étais une substance dont toute l'essence ou la nature n'est que de penser, et qui, pour être, n'a besoin d'aucun lieu, ni ne dépend d'aucune chose matérielle. (DESCARTES, 1953a, p. 148)

L'hypothèse d'un substantiel résidu qui demeurerait nonobstant la destruction totale du corps, conduit Descartes à l'introduction d'un nouvel usage du terme « moi », usage qui fait du pronom personnel (« moi, Descartes, ») un substantif (« Le moi de Descartes »). Ainsi, Descartes nomme *moi* ce quelque chose qui sert de frontière au doute, donc au monde. Quand je remets tout en question, dit Descartes, il reste la seule certitude que c'est bien moi qui fait cette expérience.

Outre ce premier argument, un second vient soutenir l'immatérialité du « je ». Ce « je » qui pense doit demeurer temporellement identique à lui-même. Or l'analyse cartésienne de l'identité des objets physiques, dont le célèbre passage sur la transformation de la bougie en cire chaude est symptomatique, indique que pour Descartes, les choses matérielles ne demeurent jamais identiques à elles-mêmes. Il est donc entendu que ce « je », qui doit demeurer le même, ne peut relever de la substance étendue. Par conséquent, conclut Descartes, le sujet est une substance immatérielle.

Que reste-t-il du sujet qui fait l'expérience du doute hyperbolique ? Qu'il est le sujet de cette expérience, autrement dit, qu'il peut témoigner à la première personne de cette expérience de pensée.

> Par exemple, si je me persuade qu'il y a une terre à cause que je la touche ou que je la vois : de cela même, par une raison encore plus forte, je dois être persuadé que ma pensée est ou existe, à cause qu'il se peut faire que je pense toucher la terre, encore qu'il n'y ait peut-être aucune terre au monde ; et qu'il n'est pas possible que moi, c'est-à-dire mon âme, ne soit rien pendant qu'elle a cette pensée ; nous pouvons conclure le même de toutes les autres choses qui nous viennent en la pensée, à savoir, que nous, qui les pensons, existons, encore qu'elles soient peut-être fausses ou qu'elle n'aient aucune existence. (DESCARTES, 1953b, p. 576)

Faire l'expérience de penser, c'est justement faire l'expérience des limites du monde par l'expérience d'une immédiateté avec soi-même. En outre, il s'agit d'une révocation de la question de l'identification, qui n'a plus à advenir, parce que le moi est toujours donné d'avance. Le syllogisme cartésien impose l'évidence apodictique du *moi*, « je suis, j'existe » comme celle d'une instance insécable et individualisante. Comme l'écrit Edmund Husserl, « L'axiome de la certitude absolue

du moi et de ses principes axiomatiques innés joue chez Descartes, par rapport à la science universelle, un rôle analogue à celui des axiomes géométriques en géométrie » (HUSSERL, 1996, p. 26). Aussi le doute prend-t-il fin, chez Descartes, lorsque le sujet du doute parvient à la certitude de cette proximité à lui-même. Le *moi* devance les pensées et en autorise l'effet. La réflexivité de la pensée consciente d'elle-même est la garantie d'un point archimédien susceptible de rendre à nouveau le monde. Le moi est le point de départ de la reconstruction du réel. De fait, le doute cartésien s'appliquant à tout ce qui est, la seule évidence est celle du rapport du moi au moi dans la proximité immanente de la mêmeté réfléchie. C'est dans le retour réflexif de la pensée que lorsque je doute, c'est « je » qui doute. C'est par la pensée que s'effectue l'évidence apodictique d'une certitude révélée par l'existence du moi se donnant de lui-même à lui-même. Il ne faut pas attendre d'autrui la révélation que moi, j'existe. Le moi est trouvé dans l'intimité d'une pensée. À travers l'acte de penser, il se confère l'existence *par lui-même* à la manière dont le Baron de Münchhausen s'extrait des eaux du lac en tirant ses propres cheveux.

Mais s'il n'est de ce monde, où loger alors ce moi ? Plutôt que d'invoquer, de manière platonicienne, un monde supraterrestre garant de la stabilité de formes pérennes, Descartes duplique le monde physique en superposant deux réalités distinctes. Le moi est un supplément qui vient coller à la matière mais ne s'y résout pas. Sensations et pensées ressortissent à cette instance qui joint le corps par le biais de la glande pinéale, mais qui ne se réduit en aucun cas à la pure mécanique de la matière étendue. Chose extrêmement versatile et subtile, comme « un vent, une flamme ou un air très délié » qui s'épand dans l'hospitalité de l'arche corps-machine. La pensée dualiste est d'emblée une pensée du double : un monde physique aux correspondances synchronisées avec un monde entièrement spirituel. La géométrie est la science de cette correspondance qui articule le corps aux pensées. L'intelligibilité du monde répond aux sollicitations géométriques, correspondance que Descartes développe à travers la science optique.

La thèse immatérialiste du moi a connu quelques variantes historiques qui ont toutes en commun l'idée que le corps n'est pas une propriété nécessaire de la personne, autrement dit l'idée que la personne pourrait encore exister sans son corps, ou plus exactement, sans aucun corps. L'argument procède de l'expérience de pensée d'une amputation complète, que celle-ci soit progressive ou non, en stipulant ensuite qu'il ne serait pas logiquement contradictoire de considérer que la personne qui en est victime puisse encore exister sans son corps. De cette putative concevabilité découlerait alors la possibilité métaphysique d'exister de manière immatérielle.

Richard Swinburne propose le raisonnement suivant, que j'appellerai argument de désolidarisation (AD) :

> Étant donné qu'il est logiquement possible qu'une personne consciente continue d'exister sans son corps quelles que soient actuellement les autres vérités sur cette personne, il s'ensuit que cette dernière doit actuellement comprendre une partie, non corporelle, qui puisse continuer d'exister sans son corps. [...] La conscience d'une personne doit donc être analysée comme son noyau immatériel, son âme consciente ». (SWINBURNE, 2015, p. 113)

De fait, l'argument de Swinburne est similaire à celui présenté par Descartes dans la sixième des *Méditations métaphysiques* :

> Et quoique de peut-être (ou plutôt certainement, comme je le dirais tantôt) j'aie un corps

> auquel je suis très étroitement conjoint ; néanmoins, parce que d'un côté j'ai une claire et distincte idée de moi-même, en tant que je suis seulement une chose qui pense et non étendue, et que d'un autre j'ai une idée distincte du corps, en tant qu'il est seulement une chose étendue et qui ne pense point, il est certain que ce moi, c'est-à-dire mon âme, par laquelle je suis ce que je suis, est entièrement et véritablement distincte de mon corps, et qu'elle peut être ou exister sans lui. (DESCARTES, 1953c, p. 324)

L'objection de circularité

La théorie substantialiste de l'identité personnelle se heurte à diverses difficultés bien connues. Comment expliquer les interactions entre la matière et la substance pensante ? Comment expliquer la coïncidence spatio-temporelle de l'âme avec le corps ? En particulier, comment expliquer leur sychronicité motrice, pour quelle raison la conscience reste-t-elle « collée » au corps dans le mouvement ? Et comment expliquer que nous puissions avoir des pertes de conscience causées par des traumatismes physiques si la conscience n'est pas matérielle ? Mais c'est une autre objection que je souhaite d'abord développer : le dualisme présuppose d'emblée le sujet dans la définition de l'identité personnelle. Le « sujet » cartésien est une pétition de principe menant à une définition circulaire du Moi.

Commençons par remarquer que, d'un strict point de vue logique, l'inférence tirée dans AD, est très discutable [1]. En effet, elle repose sur ce principe modal que la proposition $\Diamond(\exists x)\neg \mathrm{C}x$ implique la proposition $(\exists x)\Diamond\neg \mathrm{C}x$. Or cette implication ne tient qu'à condition d'interdire, entre transactions intermondaines, les domaines croissants. Mais cela n'a pourtant rien d'évident.

À première lecture, l'argument cartésien paraît d'emblée circulaire puisqu'il consiste à présupposer qu'une personne puisse exister sans matière pour conclure ensuite qu'une personne possède une partie non matérielle. Le mouvement général de l'argument donne l'impression d'une argutie comique comme celles qui sont débitées par les médecins de Molière, constatant que si l'opium fait dormir, c'est par une certaine vertu « dormitive ». Afin de rendre complètement justice à Swinburne, notons que dans une heureuse et bienvenue note de bas de page, il reconnaît que l'on pourrait suspecter son argument – mais c'est sans doute peu dire ! – d'être un « paralogisme modal » (SWINBURNE, 2015, p. 113). Aussi, Swinburne prévoit-il d'en proposer une démonstration passablement formelle, mais confuse, mobilisant au passage une prémisse incluant l'âme et l'argument des zombis. On est vite pris sous l'impression qu'il s'agit de noyer le poisson dans une eau trouble. Aussi, il me semble propice de présenter le raisonnement de Swinburne de la manière suivante :

[1] Je remercie Manuel Rebuschi d'avoir attiré mon attention sur ce point logique.

(1)	Les expressions « je » et « mon corps » sont des désignateurs rigides	hypothèse
(2)	S'il est actuellement le cas que je suis identique à mon corps, alors nécessairement je suis identique à mon corps ;	nécessité de l'identité appliquée à (1)
(3)	Or il aurait pu être le cas que je ne sois pas identique à mon corps ;	hypothèse
	———	
(4)	Donc il n'est actuellement pas le cas je suis identique à mon corps ;	*modus tolens* appliqué à (2) et (3)

La forme la plus générale de AD, l'argument de désolidarisation entre les désignateurs a et b, est basée sur le raisonnement par l'absurde suivant :

(1)	Les expressions « a » et « b » sont des désignateurs rigides	hypothèse
(2)	S'il est actuellement le cas que $a = b$, alors nécessairement $a = b$;	nécessité de l'identité appliquée à (1)
(3)	Or il aurait pu être le cas que $a \neq b$;	hypothèse
	———	
(4)	Donc il n'est actuellement pas le cas que $a = b$;	*modus tolens* appliqué à (2) et (3)

Nous reconnaissons ici la même forme de démonstration que celle employée par les partisans du sortalisme pour montrer qu'une statue n'est pas identique à son bloc de matière. Quelles voies de recours avons-nous contre la conclusion de cet argument ? Elles sont multiples. On peut :

1. rejeter le principe de nécessité de l'identité et donc, rejeter la théorie de la désignation directe ;
2. rejeter la prémisse que les expressions concernées sont des désignateurs rigides ;
3. rejeter la prémisse que l'identité aurait pu ne pas être réalisée.

Observons le raisonnement de désolidarisation de plus près. Commençons par admettre la théorie de la désignation rigide et l'hypothèse que les deux termes employés sont bien des désignateurs rigides. Examinons donc la prémisse que l'identité aurait pu ne pas être réalisée. Swinburne constate avec raison que « si je suis seulement de la matière et que cette dernière est détruite, alors la continuation de mon existence n'est même pas une possibilité logique » (SWINBURNE, 2015, p. 113). Cette remarque pourrait achever la démonstration, mais à l'encontre de la conclusion de Swinburne. Il n'est donc pas possible, conclurions-nous, d'exister sans corps matériel. Mais c'est justement l'hypothèse d'être constitué seulement de matière qui fait l'objet de l'attaque de Swinburne (et donc aussi de celle de Descartes). Il semble pourtant contradictoire qu'une personne consciente continue d'exister sans son corps, à moins de conférer à la notion de personne et au verbe « exister » une toute autre signification que celle d'usage consistant à voir dans l'existence

une présence matérielle, causale. Mais la putative possibilité d'exister sans corps n'est évidemment pas une possibilité physique : elle se veut métaphysique.

Il y a donc un « saut » ontologique mobilisé dans cette prémisse, qui consiste à passer, sans autre précaution, de la possibilité de concevoir que P à la possibilité métaphysique que P, à savoir, le principe de Hume que rien de ce qui est concevable n'est absolument impossible. Tout ce qui est épistémiquement possible est aussi métaphysiquement possible. On trouve une version de cette thèse dans la sixième méditation métaphysique de Descartes :

> Il ne me reste plus maintenant qu'à examiner s'il y a des choses matérielles : et certes, au moins sais-déjà qu'il y en peut avoir, en tant qu'on les considère comme l'objet des démonstrations en géométrie, vu que de cette façon je les conçois fort clairement et fort distinctement. Car il n'y a pas de doute que Dieu n'ait la puissance de produire toutes les choses que je suis capable de concevoir avec distinction [...]. (DESCARTES, 1953c, p. 318)

Ce que Descartes dit en substance, c'est que ce que je peux concevoir en pensée est métaphysiquement réalisable par Dieu, c'est-à-dire, possible.

Dans notre critique de l'argument de désolidarisation, nous avons commencé par admettre la théorie de la désignation rigide, pour montrer que, même avec cette concession, la démonstration n'était pas convaincante. Mais plus généralement, qu'est-ce qui ne va pas au juste dans ce raisonnement ? Lorsque nous parlons de personnes, nous parlons de personnes physiques, de corps repérables dans l'espace, quoi que nous puissions penser au sujet d'entités immatérielles surajoutées à ces personnes. Si nous affirmons qu'une personne pourrait être P sans qu'elle le soit, notre pensée contrefactuelle n'est pas complètement désengagée du concept de personne physique, sans quoi nous pourrions tout aussi bien affirmer qu'il n'est pas logiquement contradictoire de dire qu'elle aurait pu être un œuf-poché, qu'elle aurait pu n'avoir ni corps ni âme, voire qu'elle pourrait être l'une des racines cubiques de 1. Lorsque nous produisons une pensée contrefactuelle, nous le faisons en vertu d'un monde déjà donné dans lequel nous affirmons cette modalité *ceteris paribus*. Et de fait, nous absorbons les concepts qui soutiennent notre manière de voir ce monde. Autrement dit, une pensée contrefactuelle à propos d'une personne l'est toujours dans un certain contexte, associée à un certain corps présent physiquement dans l'espace et dans le temps et à un certain mode de présentation de ce concept.

L'objection de la réplique pure

Poussée dans ses derniers retranchements, le dualisme cartésien mène certainement à une théorie très étrange et difficilement compréhensible. La thèse cartésienne qui relie la possibilité métaphysique à la concevabilité permet d'audacieuses hypothèses contribuant à formuler de nouvelles objections contre le dualisme. En particulier, le second argument que je souhaite développer concerne l'hypothèse d'une personne dont on aurait produit une réplique pure matérielle. A quelle nature appartiendrait cette réplique ? S'agirait-il d'une personne ?

Remarquons que pour déduire un « je », il ne suffit pas de dire que « c'est moi qui pense », comme l'on dirait encore que « c'est lui qui pense ». Il s'agit non seulement de constater que c'est moi qui pense, mais encore que c'est moi-même qui constate que c'est moi qui pense, autrement

dit, de le constater à la première personne. Il y a donc une double réflexivité : celle de constater que c'est moi qui pense, et celle de constater ce fait à la première personne. Mais comment savoir alors que les autres hommes ne sont pas de simples mécaniques sans « moi » ? Car si le moi ne peut être atteint que par un acte performatif auto-réflexif, il m'est alors impossible d'atteindre le « moi » d'autrui. Ces silhouettes qui s'animent dans la rue, comment savoir alors si elles ne sont pas « des chapeaux et des manteaux qui peuvent couvrir des spectres ou des hommes feints qui ne se remuent que par ressorts (DESCARTES, 1953c, p. 281) ? »

À ce sujet, mentionnons que depuis le dix-huitième siècle, une étrange histoire circule au sujet de Descartes, faisant de lui une sorte de prédécesseur de Frankenstein. L'auteur des *Méditations métaphysiques* aurait été inséparable, dit-on, d'une poupée de métal qu'il aurait lui-même fabriqué à l'image de sa défunte fille illégitime, Francine, dont il aurait prêté le nom. La légende raconte que la poupée paraissait vivante au point d'épouvanter le capitaine d'un bateau sur lequel voyageait Descartes, qui, horrifié par l'artifice de vie qu'elle suscitait, la jeta par dessus bord :

> Descartes and the doll were evidently inseparable, and he is said to have slept with her encased in a trunk at his side. Once, during a crossing over the Holland Sea some time in the early 1640s, while Descartes was sleeping, the captain of the ship, suspicious about the contents of the trunk, stole into the cabin and opened it. To his horror, he discovered the mechanical monstrosity, dragged her from the trunk and across the decks, and finally managed to throw her into the water. We are not told whether she put up a struggle. (GRAUKROGER, 1995)

Le double métallique de Francine est sans doute une légende stimulée par la théorie des animaux-machines. Mais elle a le mérite de poser la question de la différence entre deux Francine physiquement indiscernables. Imaginons que Descartes eut réussi la prouesse de façonner un simulacre entièrement indiscernable de son modèle vivant, une poupée en chair et en os, et non en métal, répondant aux lois de la biologie plutôt qu'aux lois de la mécanique. S'agirait-il là d'un double parfait, d'une réplique pure de Francine ? Et sinon, qu'est-ce qui les distinguerait encore, si ce n'est une qualité physique ? Nous pourrions bien envisager une réplique pure du corps de Francine, mais cette copie serait-elle, en plus, une réplique pure de la personne de Francine ? Non, répondrait Descartes, car il y aurait bien, entre les deux, une dissemblance radicale, à savoir, la différence résiduelle de l'âme. Contrairement à Francine, sa réplique matérielle ne serait qu'un mécanisme complexe incapable de penser et d'agir volontairement.

L'éventualité de mécaniques humaines sans esprit est reprise dans « Zombies versus Materialists », article qui est à l'origine des discussions en philosophie de l'esprit sur la possibilité logique de corps sans conscience. Robert Kirk, l'auteur, suppose l'existence de deux individus physiquement indiscernables. De telles entités seraient-elles des répliques pures ? Du point de vue cartésien, cela ne serait le cas qu'à condition qu'elles possédassent, chacune, une âme.

> if Descartes'view of the nature of man were correct, an exact physical replica of a given man would not necessarily be an exact replica of that man - unless by natural necessity a soul were automatically assigned to it. (KIRK & ROGER, 1974, p. 135)

Admettons par exemple que la partie matérielle d'un individu humain soit entièrement dupliquée par une espèce de machine à répliquer précise à la particule quantique près, de sorte que les cerveaux des deux individus seraient aussi fonctionnels l'un que l'autre. Nous aurions donc affaire

à des individus dotés d'attitudes intentionnelles, de connaissances, de désirs. Il resterait pourtant qu'il ne serait attribué un authentique moi cartésien qu'au seul original, le seul doté d'une âme, substance qui, par définition, échapperait à la réplication matérielle. Pourtant, nous serions certainement enclins à penser que la réplique aurait exactement les mêmes dispositions mentales que l'original, et qu'elle mériterait ainsi le même attribut de personne. Et si l'on accepte ceci, il faut alors reconnaître que l'âme est une entité tout à fait superfétatoire pour la notion de personne.

Le partisan de l'immatérialisme pourrait alors proposer une réponse émergentiste. La conscience, dira-t-il, quoique immatérielle, est néanmoins conditionnée par l'exercice adéquat de la substance matérielle du cerveau. La substance immatérielle du moi serait ainsi dépendante de sa substance matérielle. Mais dans le cas d'un individu dupliqué, le modèle et sa copie auraient, nous l'avons dit, des cerveaux parfaitement similaires. Dans ce cas, la machine à dupliquer reproduirait, en même temps que la matière, les conditions fonctionnelles d'émergence des substances immatérielles, donc, les individus en question partageraient la même âme. On pourrait éventuellement se replier sur l'idée d'un individu bifide, fissionné dans des corps séparés. Mais de toute façon, dans une telle théorie ontologiquement hiérarchisée, il faudrait rejeter la possibilité d'exister sans corps. Quel intérêt resterait-il alors à ne pas céder au matérialisme ?

L'objection de la cruche pensante

Le sentiment d'exister implique-t-il nécessairement l'existence d'un moi ? Pourrions-nous avoir un tel sentiment sans exister ? On peut accorder à Descartes qu'une telle éventualité ressemble *a priori* à une sorte d'auto-réfutation. Car il faut bien que quelque chose existe pour avoir, par soi, le sentiment d'exister. Si je crois que j'existe, cette croyance semble ne pouvoir être entretenue qu'à la condition, justement, que j'existe. Un tel raisonnement pourrait en outre servir de réfutation au nihilisme des personnes. Pourtant, à bien y réfléchir, l'argument semble circulaire. Certes, si quelqu'un croit qu'il existe, il découle qu'il existe de fait. Mais en revanche, il n'est pas certain qu'il y ait quelqu'un croyant quoi que ce soit. L'argument d'existence que Descartes emploi présuppose d'emblée l'existence d'un « moi » qui croit. Il ne s'agit donc pas d'une conclusion tirée de la croyance d'exister. Or il est vraisemblable qu'aucun argument concernant l'existence d'un « moi » ne devrait partir de l'hypothèse que ce moi pense. C'est plutôt à partir de l'hypothèse qu'il y a de la pensée, qu'il faudrait établir ensuite la conclusion du moi qui en est l'auteur.

On trouve cependant chez Nietzsche une éclairante critique de la présupposition cartésienne du sujet, ou plus précisément, de cette *superstition* d'un moi posé comme substance et qui serait le sujet émancipé d'une énonciation à la première personne. Le point d'attaque est grammatical, et Nietzsche délivre ici une philosophie qui anticipe la critique de Wittgenstein. En effet, la généalogie critique du « moi » passe chez Nietzsche par une analyse du langage. Ce dernier, par ses usages aux raisons parfois oubliées, nous fait miroiter des entités qui n'ont, en fait, aucune sorte de réalité. La conclusion du *Cogito* cartésien qui émancipe un sujet n'est en réalité qu'un miroir aux alouettes :

> Pour ce qui est de la superstition des logiciens, je ne me lasserai jamais de souligner un petit fait que ces esprits superstitieux ne reconnaissent pas volontiers, à savoir qu'une pensée se présente quand « elle » veut, et non pas quand « je » veux ; de sorte que c'est falsifier la réalité que de dire : le sujet « je » est la condition du prédicat « pense ». Quelque chose pense,

> mais que ce quelque chose soit justement l'antique et fameux « je », voilà, pour nous exprimer avec modération, une simple hypothèse, une assertion, et en tout cas pas une « certitude immédiate ». En définitive, ce « quelque chose pense » affirme déjà trop ; ce « quelque chose » contient déjà une interprétation du processus et n'appartient pas au processus lui-même. En cette matière, nous raisonnons d'après la routine grammaticale : « Penser est une action, toute action suppose un sujet qui l'accomplit, par conséquent... » C'est en se conformant à peu près au même schéma que l'atomisme ancien s'efforça de rattacher à l'« énergie » qui agit une particule de matière qu'elle tenait pour son siège et son origine, l'atome. Des esprits plus rigoureux nous ont enfin appris à nous passer de ce reliquat de matière, et peut-être un jour les logiciens s'habitueront-ils eux aussi à se passer de ce « quelque chose », auquel s'est réduit le respectable « je » du passé. (NIETZSCHE, 1968, p. 35, aphorisme 17)

Cette critique peut d'ailleurs être formulée dans le cadre même des hypothèses cartésiennes, en faisant appel au malin génie. Car comme le concède également Eric Olson, celui-ci serait sans doute incapable de faire croire à quelqu'un qu'il n'existe pas, mais il pourrait « provoquer une instance de la pensée *j'existe* dans des circonstances où la chose est fausse - à savoir, lorsque ce ne serait pas une pensée *de* quiconque ni *de* quoi que ce soit. De même, il se peut que quiconque affirme qu'il n'existe pas soit destiné à être dans l'erreur ; mais du fait qu'une telle assertion soit faite, il ne suit pas que quiconque la fasse » (OLSON, 2017, p.230).

Ces remarques mènent directement à un nouveau problème : sur quoi faire reposer la différence entre deux moi ? Si « je » est réductible à un « quelque chose », il est menacé par une sorte d'indifférenciation schizophrénique. Car comment « je » pourrait-il vraiment parvenir à la conviction que c'est bien « lui » qui pense, et non un autre ?

Descartes a-t-il poussé le doute à ses limites ultimes ? Est-il vraiment, aux frontières du moi, parvenu à une « indubitabilité absolue » ? Par le processus d'*époché*, Husserl a tenté de radicaliser encore la mise en doute cartésienne. Selon l'auteur des *Méditations cartésiennes*, la recherche de bases apodictiques reste nécessaire pour servir de socle à la philosophie. Or, contrairement à Descartes, Husserl prétend qu'il n'y a guère « d'évidence de l'être », que l'être peut toujours faire l'objet des assauts du doute puisqu'il « peut toujours se révéler une simple apparence » (HUSSERL, 1996, p. 38). C'est donc ailleurs qu'il faut chercher l'apodicité. Des faits apodictiques ne doivent pas se contenter d'une « certitude de l'existence des choses ou des "faits" évidents » (HUSSERL, 1996, p. 38). Douter de l'existence de tels faits devrait relever d'une « inconcevabilité absolue », et exclure « d'avance tout doute imaginable comme dépourvu de sens » (HUSSERL, 1996, p. 38–39).

La critique de Husserl au sujet des méditations cartésiennes repose sur la perspective d'une philosophie qui resterait de part en part phénoménologique, indépendamment de toute thèse concernant la nature objective des choses. Cette ambition, que Paul Ricœur désigne par le terme d' « égologie », n'aurait finalement été qu'ébauchée par Descartes, mais pas réellement développée. Aussi, la philosophie transcendantale d'Husserl se devrait d'affronter les divers problèmes

> d'une philosophie qui ne serait qu'une égologie et jamais une ontologie, d'une philosophie où l'Être ne serait jamais ce qui non seulement donne un poids de réalité à l'objet, mais surtout fonde la réalité de l'Ego lui-même. Une égologie, c'est-à-dire un *Cogito* sans *res cogitans*, un *cogito* sans la mesure absolue de l'idée d'infini, sans ce *cogitatum* singulier qui serait la marque sur lui d'une toute autre subjectivité fondatrice. (RICŒUR, 1954, p. 77)

Précisons que Husserl distingue deux formes du « moi », d'une part celle, subjective, du moi psychologique et d'autre part celle, ontologique, du moi transcendantal :

> "moi", ce n'est plus l'homme qui se saisit dans l'intuition naturelle de soi en tant qu'homme naturel, ni encore l'homme qui, délimité par abstraction aux données pures de l'expérience "interne" et purement psychologique, saisit son propre *mens sive animus sive intellectus*, ni même l'âme prise séparément. Par l'*époché* phénoménologique, je réduis mon moi humain naturel et ma vie psychique [...] à mon moi transcendantal et phénoménologique, domaine de l'*expérience interne transcendantale et phénoménologique*. (HUSSERL, 1996, p. 54)

Mais la définition résiduelle du moi implique d'emblée une série de problèmes concernant les limites internes et externes de ce qui constitue le moi. Où s'arrête le moi ? Combien de corps peuvent constituer un même moi ? Comment savoir qu'un corps loge ou non un moi ? Le moi peut-il être dupliqué, fusionné, fissionné ? Puisque l'âme relève de l'immatériel, et donc de l'invisible, en disant « je », ne prend-t-on pas le risque de s'attribuer indûment quelque chose qui n'est pas soi ? De la question « qu'est-ce qui doute ? » on passe ainsi à la question « combien doutent quand « je » doute ? ». Combien sommes-nous à être *moi* » ? Sommes-nous chacun des *moi* distincts ou sommes-nous un seul et même *moi* ? Au fond, à travers la mise en doute du monde, Descartes pourrait tout aussi bien mettre en doute sa propre individualité, et comme l'écrit Vincent Descombes, se demander : « Peut-être ne suis-je pas Descartes ? » (DESCOMBES, 2014, p. 65).

Le moi cartésien, comme conscience nue, pose un véritable problème d'individuation. Car Descartes pourrait certainement aller jusqu'à douter que c'est lui, Descartes, qui pense, et, comme le souligne Nietzsche, se contenter de dire qu'il y a quelque chose qui pense. Si la substance pensante est hétéromère, autrement dit, s'il existe plusieurs « moi », comment éviter une confusion entre eux ? Descartes pourrait prendre pour *soi* quelque chose d'autre par une forme d'affranchissement de l'individualité. L'auteur des *Méditations* reconnaît d'ailleurs de manière très explicite cette difficulté :

> désormais, il faut que je prenne soigneusement garde de ne pas prendre imprudemment autre chose pour moi [...] (DESCARTES, 1953c, p. 275)

Il appert bien, dans la seconde des *Méditations métaphysiques*, cette menace de prendre le *soi* pour autre chose qu'il n'est. Car le *cogito*, s'il autorise l'évidence de la saisie psychologique d'un *soi*, n'en garantit cependant pas les frontières. Il ne garantit pas encore un moi individué, puisqu'il renonce à l'identification en transposant la question « qui est-ce qui pense ? » (c'est moi, c'est vous, c'est lui) en la question « qu'est-ce qui pense ? » (c'est une substance pensante). C'est que le sentiment d'existence, s'il est, selon l'avis de Descartes, « la plus certaine et la plus évidente » (DESCARTES, 1953c, p. 275) des connaissances, est toutefois sujette à des méprises lorsqu'il s'agit de dénombrer les penseurs. Si chacun peut s'approprier un « je », comment en délimiter alors les circonscriptions ? Comment le moi ferait-il pour se donner comme un quelque chose détaché des autres choses pensantes ? Si le moi est une instance atomique, non sécable, qui ramasse toute la singularité de la personne qui en fait l'expérience, qu'est-ce qui fait du moi, un singulier ? Différents « moi » ont-ils différentes propriétés ? Mais le doute cartésien s'applique justement à élaguer toutes les propriétés qualitatives : dans l'expérience du *Cogito*, le moi n'est plus qu'une forme

pure. Comment considérer alors des propriétés de rapports permettant une différenciation avec le non-moi ? Si les substances sont réellement distinctes les unes des autres, qu'est-ce qui fait alors la singularité d'une chose ? Qu'est-ce qui assure que telle chose n'est pas telle autre ? Dans le cas de la matière, on pourra toujours rétorquer qu'elle est étendue, et diversement étendue, que certaines choses ne sont pas connexes, que les choses matérielles ne sont pas les mêmes parce qu'elles n'occupent pas le même emplacement. La différences matérielle est une différence topographique ressortissant à une géométrie. Mais en revanche, qu'est-ce qui permettrait d'individuer des âmes ?

Martial Gueroult a particulièrement insisté sur le fait qu'on peut difficilement considérer le sujet du *cogito* comme une singularité, comme un sujet individué :

> Le fait que le *cogito* trouve dans son caractère de nature dernière la plus simple et la plus générale, la justification profonde de la certitude que nous sommes contraints de lui accorder, prouve que la réalité qui l'enveloppe n'est nullement celle de mon moi personnel concret, mais celle du moi pensant en général comme condition universelle de toute connaissance possible. (GUEROULT, 1968, p. 54)

Sous cette perspective, le sujet du *cogito* ne semble être ni personnel ni individuel, mais tendre au contraire vers l'idée d'une uniforme substance de pensée. C'est ce que Gueroult semble expliquer, en rapprochant alors le modèle cartésien du monisme spinoziste, dans une intéressante note de bas de page qui mériterait sans doute une place bien plus grande :

> (...) Car que veut dire Descartes ? Ceci tout simplement : Que sais-je si je ne me trompe pas lorsque je crois être Descartes ? Des fous ne se croient-ils être une cruche, Louis XIII ou le Cardinal ? Le Malin Génie ne peut-il pas m'abuser sur ma propre individualité ? Mercure ne réussit-il pas, pour un instant, à faire douter Sosie de son identité personnelle ? Mais, pour que je sois trompé en l'espèce, tant faut-il à tout le moins que je pense ; et puisque je pense, je suis ; car pour penser il faut être. Bref, pour m'être trompé sur moi-même, il n'est pas nécessaire que je sois Sosie, Descartes, le Cardinal, ni tel individu, ni même un individu ou une personne en général, mais simplement que je sois « une chose pensante », c'est-à-dire un moi ou un sujet pensant en général : une essence pensante quelconque. (GUEROULT, 1968, p. 54)

Doutant, Descartes pourrait se tromper de personne, se prendre pour un autre, pour Louis XIII ou pour le Cardinal. Cela pourrait aussi être plus grave, Descartes pourrait peut-être même douter qu'il est une personne, douter qu'il n'est pas une cruche, puisqu'il ne peut pas *a priori* se fier à la réalité des espèces qui démarquent les existants. De fait, l'expérimentation du *Cogito* avoisine les frontières de la folie. Pourquoi ne pas supposer alors qu'il aurait pu y avoir un monde dans lequel une cruche penserait, si cela est concevable ? Et c'est bien concevable, comme le souligne par exemple, dans l'épisode 14 de la troisième saison de la série télévisée *Twin Peaks* réalisée par David Lynch, le fait que nous comprenions que le personnage de Philip Jeffries a été transformé en une sorte de grande théière parlante. Le doute hyperbolique consiste ainsi à mettre à égale distance tous les mondes concevables, sans sélectionner d'avance celui qui conviendrait le mieux pour constituer le monde actuel. Or, puisque ce qui est concevable, estime Descartes, est aussi possible, il se pourrait que celui qui s'adonne au *cogito* habite en vérité un monde très différent de ce qu'il paraît être. Dans ce monde, il pourrait ne pas avoir de corps humain et être trompé

là-dessus par un certain nombre d'illusions. Qu'est-ce qui empêcherait alors Descartes d'être une cruche, une grosse théière fumante ou un œuf poché ?

L'objection du Moi indifférencié

La question de savoir si Descartes aurait pu être une cruche, et celle de savoir si Descartes pourrait encore répondre à la question « suis-je ? », sont différentes. La première relève de la nature de Descartes - homme, cruche - et seule la seconde, relève de l'individuation proprement dite. Les problèmes qui relèvent de la caractérisation égologique du moi ne peuvent se rassembler tous sous la même catégorie. Et le plus important, comme le souligne Descombes, est finalement l'erreur de se prendre ni pour une autre personne (Louis XIII), ni pour une autre sorte de chose (cruche), mais pour une chose *en général* :

> [...] il reste plus difficile encore à comprendre sans cette échelle des degrés de l'erreur sur soi qu'a esquissé Guéroult : le passage de l'erreur de se prendre pour quelqu'un d'autre ou pour quelque chose d'autre à l'erreur de se prendre pour "un individu" ou "une personne en général". (DESCOMBES, 2014, p. 67)

C'est ici, constate Descombes, qu'apparaît, « par une sorte d'alchimie philosophique » (DESCOMBES, 2014, p. 67) la substantialisation cartésienne du moi. Mais comme nous l'avons montré, cette caractérisation substantielle résiste encore à toute individuation, elle ressortit à un sujet *généralisé*, un sujet qui ne serait ni Descartes, ni Louis XIII, ni même telle cruche particulière, ni aucun sujet particulier. Avec une telle caractérisation, il s'avère vain de poser la question « qui est-ce ? ». Ainsi, la putative identité égologique a cette surprenante et malvenue conséquence de rendre impossible l'identification. Ainsi, l'attention que la phénoménologie porte sur le moi pur, transcendant, pendant la phase du doute cartésien, implique de considérer que ce moi n'est pas individué, qu'il est un moi « général », ce qui constitue, reconnaissons-le, la perspective paradoxale d'un *ego* qui n'aurait aucune identité particulière.

Si nous doutons de tout, il nous faut cependant prendre garde aux reflets trompeurs et ne pas nous noyer de manière narcissique dans l'eau spéculaire, trompés par les psittacismes d'une nymphe jalouse. Car si je ne puis faire confiance ni en mes jugements, ni en mes perceptions, comment puis-je savoir alors s'il y a autre chose que moi dans le monde ? Il se pourrait que le sujet du *cogito* soit réellement seul dans le monde. L'univers entier pourrait aussi bien n'inclure qu'un seul sujet. Si *je* est bien quelque chose, il faut encore s'assurer aussi qu'il y a encore quelque chose en dehors de ce *je*. Il y a tout lieu de penser que l'idéalisme problématique pourrait conduire Descartes au solipsisme, ou à l'existence spinoziste d'un unique *moi*, dont l'accès serait rendu possible par tout ceux qui doutent. Il ne pourrait au fond y avoir qu'une seule substance pensante. Car ce n'est pas le *soi* qui se donne d'emblée comme altérité. Il ne pourrait au fond n'y ait qu'un seul *soi*. C'est la conséquence inévitable d'un idéalisme radical révoquant la réalité extérieure et intérieure, n'ayant plus aucune confiance dans l'altérité mondaine. Bien sûr, Descartes ne suit pas cette voie, mais il l'esquisse tout aussi sûrement. Comme le souligne Paul Ricœur, le « je » cartésien, détaché du sujet autobiographique n'est finalement qu'une vacuité, il est « tout aussi métaphysique et hyperbolique que le doute lui-même par rapport à tous les contenus. Il n'est à vrai dire personne » (RICŒUR, 1990, p. 16)

La théorie cartésienne bifurque du solipsisme en stipulant l'existence d'au moins une autre sorte d'altérité, celle qu'occupe la présence du malin génie. Car force est de reconnaître que ce dernier diffère du *moi* qu'il cherche à tromper. Sa supposée existence vient donc récuser, avant même qu'elle n'apparaisse, l'idée que le monde n'est finalement constitué que de la substance pensante homogène d'un seul et unique *moi* universel. L'altérité au *je* est posée d'emblée par l'hypothèse du malin génie, qui postule, à l'aurore du *cogito*, la présence d'un non-soi. Ainsi, l'altérité est-elle construite à partir d'un second « Horla », à savoir le malin génie. Mais il faut bien constater la fragilité de cette explication, et la menace du retour d'une théorie solipsiste qu'elle occasionne.

L'objection du véhicule de substitution

Toutes les théories dualistes de l'identité personnelle ont sans doute le même motif récurrent : placer l'homme au dessus des autres espèces naturelles, à mi-chemin entre la mécanique de la bête et la spiritualité de l'ange. Quand Descartes a l'audace de développer la théorie de l'animal-machine, c'est ensuite pour en extraire l'homme en partie. Si la science anthropologique a corrigé cette erreur, en rendant aux hommes leur généalogie animale, l'illusion est cependant largement maintenue en métaphysique, où l'on entend parler, dans certains séminaires contemporains, de la possibilité de purs esprits.

Nietzsche accuse la théorie substantialiste de l'identité personnelle, avec la tonicité féroce dont il est coutumier, d'être une illusion théologico-métaphysique. Il s'agit d'une erreur de perspective consistant à confondre espoir et vérité. Or, écrit Nietzsche,

> La vérité et la *foi* en la vérité de quelque chose : ce sont là deux sphères d'intérêt totalement différentes, presque deux mondes antithétiques – c'est par des voies opposées que l'on parvient à l'un et à l'autre. (NIETZSCHE, 1974a, p. 180)

Les thèses de l'identité personnelle instaurent ainsi une sorte de fongibilité entre la réalité et un monde *désiré*, un monde de remplacement, un double. En outre, poursuit Nietzsche, nous devrions admettre que « "le pur esprit" est une pure sottise : si dans nos calculs, nous faisons abstraction du système nerveux et des sens, bref, de l'"enveloppe mortelle", et bien, nous faisons un calcul faux – et un faux calcul, un point c'est tout ! » (NIETZSCHE, 1974a, p. 172).

Le dualisme serait donc, de ce point de vue, moins une théorie qu'une sorte de consolation à la destruction certaine des personnes. La poupée de Descartes, réplique exacte de Francine, est un exemple de double de substitution, double ontologique dont l'idée est suscitée par le désir de fournir à une personne un véhicule neuf lorsque son corps vient à défaillir. Avec les études séminales d'Otto Rank (RANK, 1973), d'Enrst Jentsch (JENTSCH, 1906) et de Sigmund Freud (FREUD, 1985c), la psychanalyse a essentiellement concentré son attention au double sur le rapport de substitution symbolique qu'il occasionne, en y voyant un faux-fuyant psychologique, un détour parfois inconscient permettant d'éluder l'idée de mort.

Dans son essai intitulé *Le Double*, Otto Rank propose l'étude systématique des croyances portant sur la duplication de personnes. Le thème est analysé à partir d'œuvres cinématographiques et littéraires auxquelles succèdent des études portant sur les notions d'ombre, de reflet, de dualisme gémellaire et d'immortalité du moi dans diverses traditions. Rank distingue essentiellement

deux types de doubles. La première sorte relève des ombres et des reflets qui se rapportent à des conceptions primitives de l'âme. Dans le *Rameau d'Or*, l'anthropologue James George Frazer avait souligné que le folklore de nombreuses civilisations associent intimement l'ombre et le reflet à l'âme (FRAZER, 1981), autrement dit à un double capable de subsister après la destruction matérielle de l'individu. Le rapprochement entre l'ombre et l'âme, qui représente l'hypothèse d'une immortalité du Moi, est ainsi expliqué par leur caractère commun d'immatérialité. L'ombre et le reflet fournissent à la personne une sorte de véhicule de remplacement lorsque le corps disparaît. Rank constate que de telles conceptions ressurgissent dans l'inconscient collectif à travers les fictions littéraires et cinématographiques du double. Prolongeant le constat de Frazer, Rank conclut que l'engouement pour la littérature du double à la fin du XIXe siècle peut ainsi s'expliquer par une forme de conjuration de la destruction matérielle du corps, offrant « un démenti à la mort par le dédoublement du Moi sous forme d'ombre ou de reflet » (RANK, 1973, p. 112).

La deuxième sorte de doubles décrite par Rank s'apparente aux jumeaux et aux sosies, et plus généralement, aux corps matériels *redondants*, aux *doppelgängers*. Cette fois encore, Rank procède en collectant des données anthropologiques qui indiquent que, pour diverses traditions anciennes, l'apparition de doubles gémellaires est due à l'affranchissement d'un tabou. Le sacrifice d'un des jumeaux est alors nécessaire afin de rétablir l'harmonie sociale. Ce sacrifice est d'ailleurs bien apparent dans les mythes fondateurs comme celui de Romus et Remulus. Ces mythes indiquent aussi l'association entre gémellité et ce que Rank désigne comme un principe auto-créateur :

> L'idée que les jumeaux se sont créés eux-mêmes me paraît se manifester avec évidence dans la croyance très répandue d'après laquelle les jumeaux d'un sexe différent peuvent accomplir l'acte sexuel déjà avant leur naissance, dans le corps de leur mère, et transgresser ainsi le tabou de l'exogamie. (RANK, 1973, pp. 97-98)

Cette capacité auto-productive est certainement responsable de l'évolution du statut des jumeaux, passant d'objets « tabous » à objets de culte. La fascination qui leur est témoignée indiquerait là encore l'espoir d'immortalité, en particulier lorsque le développement culturel et scientifique a rendu l'hypothèse de l'âme caduque ou du moins plus difficilement acceptable. Le couple de jumeaux représente un être dédoublé dont la destruction d'un membre peut être compensée par la survie de l'autre.

De manière plus générale, Rank constate que l'idée même du *doppelgänger* tient à une tentative pour surmonter l'angoisse de la mort par l'hypothèse d'un corps de substitution, un corps qui pourrait être fabriqué par l'homme lui-même [2]. Les traits caractéristiques du thème du double

[2] L'idée de la création d'un être vivant par des moyens artificiels est bien sûr ancienne, puisqu'on la retrouve dans les mythes grecs de Prométhée et de Pygmalion, ainsi que dans le mythe hébraïque du Golem. On peut aussi citer la fascination qu'exercent les automates depuis l'antiquité, et surtout depuis le siècle de Descartes avec l'invention du ressort qui permet un mouvement « retardé » conférant à la machine une sorte d'autonomie. Les récits mettant en scène des ingénieurs démiurges prolifèrent à partir du XIXe siècle, rappelons par exemple la poupée Olympia dans le conte d'Hoffmann (HOFFMANN, 1979), l'histoire de Pinocchio de Collodi (COLLODI, 2003) et bien sûr le Frankenstein de Mary Shelley (SHELLEY, 2008). De manière plus contemporaine, le double comme phénomène d'auto-création intervient régulièrement dans la science-fiction actuelle, par exemple dans la figure des androïdes du roman *Do Androïds Dream of Electric Sheep ?* de Philip K. Dick (DICK, 1968), dans la réplication génétique des clones de *The Island* (BAY, 2005) et dans l'hypothèse de l'auto-création de l'espèce humaine dans le film *Prometheus* de Ridley Scott (SCOTT, 2012).

sont donc à la fois l'angoisse de la mort et son affranchissement par la création d'un double, celui-ci pouvant s'avérer immatériel dans le cas des récits d'ombres et de reflets, ou matériel dans le cas des récits de jumeaux et des *doppelgängers*. Rank conclut ainsi que « l'idée de la mort a été rendue supportable par l'intervention du Double, dont le rôle est d'assurer une autre vie après celle-ci. » (RANK, 1973, p. 114). L'intérêt littéraire que rencontre le thème du double serait donc essentiellement dû, écrit Rank, au refus de vieillir, à un vœu sans espoir de jeunesse éternelle [3].

Mais il ne fait pas de doute, constate Rank, que de telles pulsions narcissiques sont inaptes à refouler continuellement l'angoisse de la mort, voire qu'elles la favorisent. Le double devient alors l'incarnation même de l'angoisse de destruction. C'est ce que rappelle d'ailleurs le proverbe qui affirme que celui qui rencontre son double doit bientôt périr. Car le remède s'avère finalement plus néfaste que curatif :

> Dans les mêmes phénomènes de défense revient aussi la menace contre laquelle l'individu a voulu se défendre et se garder ; ainsi s'explique que le double qui devrait personnifier l'amour narcissique devient précisément un rival dans l'amour sexuel ou, créé au début dans le désir d'éloigner l'anéantissement éternel tant redouté, revient dans la superstition comme un messager effrayant de la mort. (RANK, 1973, p. 115)

Alors qu'il est d'abord destiné à prolonger l'individu pour tromper la mort en s'interposant entre lui et elle, le double se rapproche, puis se confond bientôt avec elle, devenant un rival qu'il faut abattre sous peine de périr [4], destruction qui symbolise, selon Rank, le vœu d'anéantir la mort elle-même, par une sorte de suicide à vide. La mort est ainsi substantivée par le double, le double la personnifie, elle devient une menace actuelle et concrète plutôt que possible et abstraite.

L'analyse des profondeurs psychologiques du double montre ainsi, pour Rank, que le thème est fondamentalement lié au retour d'un conjuré : l'idée de mort. L'individu se reproche sa mortalité et la met à distance par la création d'un double. Cette mise à distance par le double est dans un premier temps consolatrice car elle fournit l'artifice d'une illusion d'éternité, par la possibilité d'un prolongement de soi. Mais elle devient ensuite progressivement menaçante parce qu'elle ne cesse de confronter le sujet à l'illusion qu'il s'est ainsi créée. Il doit alors détruire le double au risque de se détruire lui-même, c'est-à-dire de produire justement l'avènement de ce qu'il rejette, sa propre destruction.

Les analyses biographiques que Rank développe au sujet des écrivains partent du principe freudien que l'activité artistique a toujours pour origine une névrose. Pour Freud, le moi est en effet une instance jouant le rôle d'intermédiaire entre le monde extérieur et le ça, domaine des pulsions narcissiques. Le psychanalyste utilise ensuite cette tripartition de l'identité psychologique pour élaborer la distinction entre deux types de maladies mentales, névroses et psychoses. Le premier type relève d'un conflit entre le moi et le ça, le second d'un conflit entre le moi et le monde extérieur. Dans le cas des psychoses, écrit Freud, il y a production d'un double du monde :

> Le Moi se crée automatiquement un nouveau monde, extérieur et intérieur à la fois ; deux faits

[3] On retrouve d'ailleurs le thème de l'éternelle jeunesse explicitement corrélé à celui de l'ombre dissociée dans la figure de Peter Pan, personnage qui apparaît pour la première fois en 1902 dans le conte *The Little White Bird* de James Matthew Barrie (BARIE, 1902).

[4] Voir par exemple le *William Wilson* d'Edgar Allan Poe (E. A. POE, 1932).

> ne font aucun doute : ce nouveau monde est bâti suivant les désirs de ça, et le motif de cette rupture avec le monde extérieur, c'est que la réalité s'est refusée au désir d'une façon grave, apparue comme intolérable. (FREUD, 2010, p. 286)

La défense du sujet contre l'intolérable réalité le conduit ainsi à produire un dédoublement du Moi.

Le principal mérite de l'essai de Rank est qu'il est le premier à avoir planifié une étude diachronique et universelle des croyances relatives aux doubles en corrélant des analyses d'œuvres cinématographiques et littéraires à des études anthropologiques. En outre, il a mis en relief l'idée de véhicule de remplacement que constitue le double face à la menace de destruction de l'individu. Son approche, qui associe le double à un clivage entre l'âme et le corps, reste néanmoins confinée à une interprétation symbolique ancrée sur une métaphysique dualiste, et en particulier, sur des doubles ontologiques.

Après l'étude de Rank, Freud s'est lui-même intéressé à la littérature du double, en empruntant à Ernst Jentsch le concept d'*unheimlich* (JENTSCH, 1906) afin de qualifier un certain type d'angoisse, un sentiment « d'inquiétante étrangeté », une sorte de non-familiarité qui désoriente. Chez Jentsch, l'*unheimlich* est particulièrement attaché aux sentiment qui émerge lorsqu'un individu est confronté à la notion de réplique pure du corps d'une personne.

Jentsch utilise plus généralement le terme d'*unheimlich* pour qualifier l'hésitation devant la nouveauté, l'incertitude intellectuelle devant l'originalité, ou plus généralement encore, une forme de désorientation ressentie devant des phénomènes inhabituels. Mais selon Jentsch, ce sentiment résulte plus particulièrement de deux sortes d'hésitations à l'égard d'un objet. Le premier type d'hésitation se rapporte à une forme de doute à l'égard de ce que l'on doit penser d'une chose qu'on croyait non vivante et qui pourtant imite la vie, comme par exemple les poupées de cire ou les automates. La poupée Olympia qui, dans le conte d'Hoffmann *L'homme au sable*, provoque les sentiments amoureux du personnage principal, en est une illustration littéraire. Ce sentiment peut donc être convoqué de manière littéraire par certains procédés, dont Jentsch spécifie les mécanismes. Comme le rappelle Freud, :

> l'un des stratagèmes les plus sûrs pour provoquer aisément par des récits des effets d'inquiétante étrangeté, consiste donc à laisser le lecteur dans le flou quand à savoir s'il a affaire, à propos d'un personnage déterminé, à une personne ou par exemple à un automate, et ce de telle sorte que cette incertitude ne s'inscrive pas directement au foyer de son attention, afin qu'il ne soit pas amené à examiner et à tirer la chose aussitôt au clair, vu que (...) cela peut compromettre l'effet affectif spécifique. (JENTSCH, 1906, p. 224)

Le deuxième type d'hésitation qui définit le sentiment d'*unheimlich* selon Jentsch provient inversement du spectacle de comportements humains mécanisés, autrement dit de situations où des organismes agissent comme s'ils étaient des pantins ou comme s'ils étaient commandés par des rouages d'horlogerie, à la manière dont peuvent parfois se comporter des individus victimes de certaines manifestations de démence ou de crises d'épilepsie [5].

[5] Notons au passage que, contrairement à Jentsch, chez Bergson, l'idée de mécanique plaquée sur du vivant est responsable du sentiment de comique. (BERGSON, 1955).

Dans l'essai intitulé *Das Unheimliche* (1919), Freud associe directement ce terme au sentiment provoqué par les récits de doubles. Mais Freud cherche à préciser la portée du concept en lui conférant un sens plus psychanalytique que celui de Jentsch. Aussi, Freud commence par lier l'inquiétante étrangeté « à l'effrayant, à ce qui suscite l'angoisse et l'épouvante » (FREUD, 1992, p. 213). Une analyse étymologique lui permet ensuite de dégager la signification plurivoque du terme « unheimlich » :

> (Le) terme de *heimlich* n'est pas univoque, mais (…) appartient à deux ensembles de représentations qui, sans être opposés, n'en sont pas moins fortement étrangers, celui du familier, du confortable, et celui du caché, du dissimulé. (…) Serait *unheimlich* tout ce qui devait rester un secret, dans l'ombre, et qui en est sorti. (FREUD, 1985c, pp. 221–222)

L'*unheimlich* manifesterait ainsi une forme de réunion inattendue entre la familiarité et la révélation malvenue.

A travers l'étude de la littérature du double, l'impression d'*unheimlich* est finalement définie par Freud comme le sentiment ressenti au contact du « retour d'un refoulé ». D'un point de vue général, Freud considère que l'activité littéraire est liée à la pratique enfantine consistant à se créer des univers de jeux. L'enfant investit une grande quantité d'affects dans cette activité élaborée généralement avec beaucoup de sérieux.

> L'opposé du jeu n'est pas le sérieux, mais (…) la réalité. L'enfant distingue très bien son monde ludique, en dépit de tout son investissement affectif, de la réalité, et il aime étayer ses objets est ses situations imaginées sur des choses palpables et visibles du monde réel. Ce n'est rien d'autre que cet étayage qui distingue encore le « jeu » de l'enfant de la « fantaisie ». (FREUD, 1985a, p. 34)

Ainsi, le jeu précède la fantaisie. Celle-ci a pour fonction l'accomplissement illusoire de désirs insatisfaits. Or, selon le fondateur de la psychanalyse, la création littéraire n'est rien d'autre que l'exaspération de ces fantaisies.

Freud prend donc pour point de départ la même approche mimétique que celle élaborée par Aristote dans *La Poétique*. Comme Aristote, Freud constate le plaisir produit par les activités mimétiques, même quand celles-ci se rapportent à des choses qui devraient dans d'autres conditions inspirer le dégoût. Et comme Aristote, Freud relève la dimension cathartique de l'imitation, par l'assouvissement de pulsions refoulées. Mais contrairement à Aristote, la conception de Freud réduit l'activité littéraire toute entière à la canalisation des pulsions individuelles. L'auteur d'un récit de fiction est lui-même double en ce que son moi est éclaté, scindé en pulsions multiples et contradictoires qu'il canalise en les personnifiant :

> Le roman psychologique doit sans doute dans l'ensemble sa particularité à la tendance du créateur littéraire moderne à scinder son moi en moi partiels, par l'effet de l'observation de soi ; et par voix de conséquence, à personnifier les courant conflictuels de la vie psychique en plusieurs héros. (FREUD, 1985a, p. 43)

Ainsi, Freud traite la littérature, dans sa globalité, comme une masse uniforme de récits de complexes latents et d'actes de refoulements. L'écrivain est essentiellement un névrosé qui soigne

ou accentue sa maladie par l'écriture. Le Moi est vu comme un courant de positions contraires, de frustrations, de volontés, d'affects et de sentiments contradictoires qui ne cherchent qu'à assouvir des pulsions réprimées. Ne pouvant aboutir à leur satisfaction par un contexte direct, les pulsions supportent la frustration en revenant sous une forme sublimée, ce que Freud qualifie de « retour du refoulé ». Le personnage de fiction est ainsi le produit de pulsions qui sont projetées dans la fiction parce qu'elles ne peuvent être satisfaites ailleurs. Selon Freud, le roman psychologique le plus réaliste peut ainsi être interprété comme un équivalent aux fantaisies de la rêverie diurne [6]. La création d'une travail littéraire est ainsi rapporté à un sentiment de déjà-vu : l'œuvre n'est qu'un rappel latent d'un complexe issu de l'enfance :

> une expérience actuelle intense réveille chez l'écrivain le souvenir d'une expérience antérieure, appartenant la plupart du temps à l'enfance, dont émane maintenant le désir qui se crée son accomplissement dans l'œuvre littéraire ; l'œuvre littéraire elle-même permet de reconnaître aussi bien des éléments de l'occasion récente que des éléments du souvenir ancien. (FREUD, 1985a, p. 44)

Comme le regrette Cécile Kovacshazy (KOVACSHAZY, 2012), l'*unheimlich* ne permet pas à Freud de définir véritablement le rôle du double en rapport à l'identité personnelle, car sa définition s'avère sans cesse contredite et rendue caduque. En tentant de concilier les contraires, la définition de l'*unheimlich* se perd en multiples alternatives successives. Le double est d'abord conçu, avec Jentsch, comme un malaise provenant d'une incertitude intellectuelle, puis, avec Otto Rank, comme l'angoisse issue de la volonté d'immortalité, puis successivement comme fonction d'autocritique, comme sublimation de fantasmes non accomplis, puis comme « retour du refoulé ».

Les travaux de Rank, de Jentsch et de Freud ont le mérite de souligner le rôle de véhicule de remplacement que promet le double face à l'angoisse de la mort. Mais les interprétations psychanalytiques du double restent cloisonnées à des herméneutiques symboliques trop restrictives. En outre, si la théorie psychanalytique des profondeurs de la psyché éclate l'identité personnelle en différents courants pulsionnels contradictoires, il n'est pas du tout évident qu'elle renonce à un dualisme corps-esprit [7], et sans doute ne fait-elle que l'exacerber, en rapprochant le double du dual. Elle s'affaire de l'illusion produite par le double, insiste sur le caractère irréel des doubles ontologiques, mais reste tacite sur le rôle épistémique joué par les doubles dans notre rapport à l'identité personnelle.

---

[6] De manière plus contemporaine, Robert Rogers reprend l'analyse psychanalytique du thème littéraire du double, en lui appliquant la distinction freudienne entre double latent et double manifeste. La thèse de Rogers consiste essentiellement à prouver la validité des complexes freudiens par le biais d'exemplifications littéraires. Le double serait toujours le témoignage des désir refoulés d'une libido en bute avec le Sur-Moi. (ROGERS, 1970)

[7] En 1890, Freud proclame : « "Psyché" est un mot grec que l'on traduit par âme (*Seele*). "Traitement psychique" veut dire par conséquent "traitement d'âme" » (FREUD, 1984, p. 1). L'ambition du psychanalyste aura été de séparer nettement l'esprit de la psyché. Mais Freud ne s'est jamais prononcé de manière explicite sur la nature de l'âme. Dans certains passages, il parle de l'interaction de l'âme et du corps, en les distinguant manifestement. Mais dans un passage lapidaire, il écrit contre Descartes : « Psyche ist ausgedehnt, weiss nichts davon » (FREUD, 1938, p. 288). L'âme est étendue, mais n'en sait rien. Et contre les conditions *a priori* que Kant confère à l'intuition, Freud fait découler la spatialité d'une projection de l'appareil psychique. Enfin, Freud rejette la théorie de Groddeck (GRODDECK, 1976) qui assimile l'âme au corps. En conséquence, et quoi qu'il en soit de la nature de psyché, il ne fait pas de doute que la conception de Freud considère l'âme comme une entité objective.

7.1.2 Les théories du Moi absolu

Avant de passer à la conception matérialiste de l'identité personnelle, je souhaite développer quelques commentaires sur la conception du Moi dans l'idéalisme allemand, en particulier, dans ses rapports à la littérature romantique du double.

Nous l'avons vu, la menace du cartésianisme est le solipsisme absolu, l'existence d'un unique soi. Comme le constate Vincent Descombes, les doctrines de Fichte et de ses successeurs « sont justement le sommet de l'effort spéculatif de toute une tradition pour penser le sujet d'une conscience de soi *comme sujet rapporté à lui-même par un acte de réflexion sur lui-même* » (DESCOMBES & LARMORE, 2009, p. 70). Plus que nul autre système métaphysique, l'idéalisme post-kantien a maintenu une tension dialectique entre le sujet pensant et le monde, en débouchant sur un solipsisme absolu, où le Moi se donne comme fondamentalement premier, producteur de lui-même et du monde extérieur. La pression du moi occasionnée par l'idée qu'il est responsable du monde, a exercé une influence notable sur la littérature romantique. En particulier, certains auteurs ont reconnu son influence sur la prolifération des récits de doubles au XIXe siècle. C'est dire que l'idéalisme absolu entretient un évident rapport au double, que je me propose d'expliciter maintenant.

Rappelons d'abord que, dans l'idéalisme allemand, le monde n'est plus vu comme un fait objectif qui serait donné *à coté du moi*, mais qu'il est observé depuis l'intimité propre du moi, et avec son matériau de connaissance propre. Avec Kant, le temps et l'espace sont devenus des formes transcendantes invaginés dans le creux du sujet. Le monde empirique vient prendre consistance à travers le moule des schèmes *a priori* qui confèrent une structure aux connaissances empiriques venant s'y constituer comme représentations. Parmi les paradoxes de l'identité personnelle, les philosophes qui ont accepté l'héritage des doctrines kantiennes ont mis en exergue l'opposition contradictoire entre le « moi » posé à la fois comme sujet de la connaissance et à la fois comme objet de connaissance. Le moi, en tant que conscience de soi, ne peut ainsi être défini qu'en étant un objet pour soi. Mais en tant qu'instance devant laquelle se présentent justement tous les objets, il doit aussi être considéré comme sujet, c'est-à-dire ne pas être pris lui-même pour un objet. C'est la contradiction majeure à laquelle aboutit la conception dialectique du sujet, qui nécessite à la fois d'être un *objet-pour-soi* et à la fois d'être pour soi, autre chose qu'un objet. Comme l'écrit Cécile Kovacshazy à propos des récits de double, chez les romanciers,

> *Je* raconté comme un objet de la pensée s'écarte du *je* sujet : puisqu'il s'en écarte grammaticalement, il s'en écarte peu à peu dans la fiction. Entre la dualité du sujet et sa duplication il n'y a qu'un pas à franchir : l'attribution d'un statut du non-autre comme autre. La dualité intérieur devient une dualité extérieure, objective : elle devient duplication. (KOVACSHAZY, 2012)

La scission entre un moi vivant et un moi stagnant autorise alors une sorte de réification supplémentaire, rejetant l'objet par delà le sujet, à la manière dont l'image spéculaire pose le sujet par devant lui-même. Cette image étant toujours accidentelle, elle fait dire à Pierre Jourde et Paolo Tortonese que « cette découverte de la contingence correspond donc à une double perte : en tant qu'objet, le personnage ne se possède plus, il appartient aux autres, à leurs désirs, aux lois de la morale ou de l'esthétique ; en tant que sujet, il perd tout contenu et toute identité » (JOURDE &

TORTONESE, 1996, p. 92). Le double vient questionner la relation entre le sujet, dont la dynamique de pensée, l'être-pour-soi, échappe par nature à toute réification stabilisée en objet-pour-autrui, et son inscription dans le monde, devant autrui, qui le ramène justement à cette réification : l'objet qu'est le sujet pour les autres sujets pensants.

Dans la *Doctrine de la science*, Johann Fichte va plus loin encore dans l'association entre le moi et la représentation du monde. Fichte constate d'abord l'impasse dans laquelle s'est plongée la philosophie idéaliste. Le sujet, qui y est défini en terme de conscience de soi, nécessite aussi d'être pensé comme objet. Mais si le sujet peut penser à soi comme il peut penser à tout autre objet, le mode sous lequel il le fait doit alors s'en distinguer.

> C'est avec la même liberté que celle avec laquelle je pense le mur que je pense le moi ; ainsi dans la pensée du moi, il y a aussi quelque chose qui est pensé mais ce qui pense et ce qui est pensé ne sont pas distingués de la même manière que lorsque je pense le mur. (FICHTE, 2000, p. 90)

Qu'est-ce qui distingue alors la conscience de soi de la conscience du mur ? Fichte met en exergue l'aspect réflexif de la conscience de l'objet. Quand je fais l'expérience d'un objet, je fais l'expérience d'en faire l'expérience, autrement dit de faire moi-même l'expérience. Or, selon Fichte, cette caractéristique de mise en abîme de la conscience implique une nature primitive et inexplicable :

> Je suis conscient d'un objet quelconque B mais je ne puis cependant en être conscient sans être conscient de moi-même, car B n'est pas moi et je ne suis pas B. Or, je ne suis conscient de moi-même que parce que je suis conscient de la conscience. Je dois donc être conscient de l'acte de la conscience, de la conscience de la conscience. Comment puis-je en prendre conscience ? On est amené ici à redresser à l'infini, et de cette manière on ne peut expliquer la conscience. (FICHTE, 2000, p. 90)

Lorsque le Moi est pris comme objet de pensée, il est tendu devant lui-même, à bout de perche, comme une carotte est tendue devant un âne pour le faire avancer, et la trajectoire décrite est circulaire, renvoyant le sujet à l'objet et réciproquement.

Quelle solution à cette impasse du Moi comme sujet et comme objet ? Selon Fichte, c'est parce que l'on considère le Moi comme un objet qu'on est amené à cette régression infinie. Or suffit de supprimer l'idée que la conscience soit un *objet de conscience* pour trouver une issue à ce paradoxe. Pour cela, Fichte réalise un tour de force, il supprime tout simplement l'opposition sujet-objet rabattant complètement l'objet sur le sujet. Le moi n'est plus la condition nécessaire de la connaissance des objets, mais l'instance même qui fait d'eux des objets.

> Tout acte de représenter est un acte d'auto-position. Tout procède du moi. Le moi n'est pas un composant de la représentation mais toute représentation procède du moi. Toute conscience possible suppose la conscience immédiate et ne peut être conçue en dehors d'elle. L'identité du posant et du posé est absolue, elle ne peut être apprise ni issue de l'expérience, elle est ce qui rend possible l'apprentissage et l'expérience. Le moi n'est nullement sujet mais est : sujet = objet ; s'il n'était que sujet, alors la conscience serait inconcevable ; s'il n'était qu'objet, alors on rechercherait un sujet hors de lui que l'on ne trouverait jamais. (FICHTE, 2000, p. 92)

L'idéalisme transcendantal de Fichte renonce ainsi à laisser échapper quelque chose qui ne serait pas le moi, si bien que la continuité du moi au monde implique la non-démarcation entre chose en soi et phénomène. Le moi empirique, identifié au moi transcendantal, devient une instance totalitaire qui englobe tout, qui usurpe l'altérité du monde. Pour les tenants de la tradition romantique qui adhèrent à l'idéalisme transcendantal, le monde, dans son entièreté, n'est au fond qu'un reflet du moi. Contrairement au sujet individué, le moi transcendantal perd paradoxalement sa singularité. Le moi est toujours *ailleurs*, il ne peut être saisi qu'hors de lui parce que le sujet conscient se représente toujours quelque chose qui n'est pas lui-même. L'intérieur et l'extérieur qui démarquaient sujet et objet chez Descartes deviennent, dans l'idéalisme allemand, indémarquables, ambigus, intangibles. Le moi s'effrite pour se dissoudre dans le monde.

C'est là, sans doute, que réside tout l'enjeu du double romantique qui, fier d'un Moi absolu, ne peut se satisfaire d'une réduction matérielle. Wilhelmine Kraus (KRAUS, 1927) caractérise d'ailleurs le thème littéraire du double comme émanation directe de la conception de l'identité du romantisme allemand. Son analyse, adoptant un point de vue psycho-historique, se concentre sur les conflits de l'identité personnelle dans les œuvres de Jean-Paul et d'E.T.A Hoffmann. Kraus soutient que les théories de Fichte et de Schelling ont produit les conditions culturelles idoines pour ces littératures qui démultiplient les composantes subjectives du Moi. Selon Kraus, la littérature du double serait née du conflit entre une aspiration vers *l'absolu* et la finitude propre à l'existence humaine. La littérature du double mettrait ainsi en évidence, à sa manière propre, les tentatives désespérées du soi pour dépasser ses limites matérielles et atteindre à l'absolu. Cet idéal est vécu comme une sorte d'échec toujours renouvelé, qui engendre une tension au sein de l'individu, tension menant à un clivage du moi [8]. Comme l'écrit encore René Girard, qui voit dans l'idée d'une identité autonome le mensonge romantique par excellence, « Les grands élans lyriques détournent du réel sans libérer vraiment, car les ambitions qu'ils éveillent sont, en définitive, terriblement mondaines. La victime du romantisme devient toujours plus inapte à la vie, tout en exigeant d'elle des choses toujours plus exorbitantes. L'individualisme littéraire est une sorte de drogue dont il faut sans cesse augmenter les doses pour se procurer, au prix de souffrances toujours accrues, quelques douteuses extases. L'écartèlement entre « l'idéal » et la réalité sordide est aggravé. Après avoir fait l'ange, le héros souterrain fait la bête. Les dédoublements se multiplient. » (GIRARD, 1983a, pp. 67-68).

La littérature romantique du double questionne ainsi les conséquences d'une philosophie qui encourage la dissolution de l'identité personnelle dans la totalité du monde, en concevant le sujet romantique comme, selon les mots de Wladimir Troubetzkoy, « le vertige du moi penché sur son propre vide, sur un univers vidé de toute altérité » (TROUBETZKOY, 1996, p. 26). L'idéalisme allemand fournit ainsi à la littérature les conditions métaphysiques de la disjonction du « je-me-pense » cartésien. La connaissance d'un soi-diffusé ailleurs, la *Selbstgefühl*, est l'hymne de l'exploration romantique de la psyché. Il n'y a plus d'autorité suprême en dehors du moi. Il n'y a plus de vérité ailleurs qu'en soi. Le monde s'efface derrière sa représentation. Et ce dédoublement est répercuté sur le sujet qui devient son seul et unique objet de pensée, mais éparpillé. A travers son exploitation du thème du double, la littérature allemande du XIXe siècle s'est faite témoin de cette

[8] On retrouve chez Ralph Tymms (TYMMS, 1949) une analyse semblable, à cela près que l'auteur élargit la conception de Kraus à d'autres littératures. En outre, Tymms nuance la distinction de sa devancière en proposant un clivage entre le double allégorique et le double psychologique, le premier relevant de l'inconscient, et le deuxième de la conscience.

dissolution métaphysique de l'individu, manifestant une perte, pour le sujet pensant, des frontières avec le monde, à l'instar des peintures romantiques de Caspar David Friedrich où le spectateur voit le spectacle du monde *à travers un personnage représenté de dos*. Le but est théologique, ou plus exactement encore, téléologique, il a pour finalité la restitution d'une globalité qui échappe à l'individu romantique :

> Le romantisme et la philosophie germaniques, avec Fichte au premier chef, apparaîtront, dans cette perspective, comme une tentative de reglobalisation de la saisie de l'univers, comme un retour tenté vers ce que la théologie médiévale avait de complet, pour l'homme, et surtout de rassurant : l'homme se redonne un Dieu et un maître, de qui faire dépendre la création continuée du monde du soi, mais il n'y a plus personne au-dessus de l'empyrée, et ce Dieu garant et soutien de l'être, c'est lui-même, l'homme-Dieu, le moi transcendantal qui se produit en produisant le monde, en le connaissant et en l'investissant. (TROUBETZKOY, 1996, p. 7)

La réintégration, dans un système global, de l'individu atomisé passe donc paradoxalement par la dispersion du moi dans le monde.

En admettant un moi absolu réifié, l'idéalisme allemand passe à côté de l'idée que moi ne serait qu'un effet de perspective, qu'un point de vue sur le monde. L'idéalisme absolu n'est finalement qu'une forme de substantivation hypertrophiée du moi, un moi dont les limites coïncident avec la totalité du monde. La menace d'une telle phagocytose mutuelle du moi et du monde rejaillit sur la littérature romantique du double. Là où la conception classique du double, celle des *Ménechmes* de Plaute, n'était encore qu'un prétexte à une confusion explicable, à une usurpation qu'il est toujours possible de dévoiler, la conception romantique du double devient plus inquiétante car la relation entre l'original et son double s'avère maintenant inextricable. Car si l'on admet l'identité « objet = sujet », rien ne différencie plus alors le moi et l'autre. Le double n'est alors plus l'affaire d'une ressemblance troublante de qualités intrinsèques, il n'est plus seulement une affaire de répliques pures, mais devient l'objet de doubles purs, d'entités pas même discernables de manière extrinsèque par des rapports spatio-temporels. Le double s'insinue même dans le rapport à l'identité jusqu'à admettre comme dupliques des individus au demeurant parfaitement dissemblables, comme dans la *Méprise* de Vladimir Nabokov, où le personnage principal considère comme son alter-égo un individu qui lui est quasiment en tout point différent. C'est qu'il n'y a plus aucun repère pour tenir la ressemblance écartée de la représentation : tout objet est finalement capable de représenter n'importe quel autre. L'idéalisme du moi absolu qui rend équivalent le monde et sa représentation conduit ainsi à la conclusion que c'est toute altérité qui est finalement reniée, menant ainsi la métaphysique tout droit à la thèse de l'hénadisme, autrement dit, d'une substance unique et indivisible.

7.1.3 Le monisme matérialiste

La conception matérialiste de l'identité personnelle basée sur la biologie contemporaine et définissant la personne par la continuité spatio-temporelle d'un corps matériel partage avec la conception de l'âme immatérielle l'idée d'une unicité stricte de l'individu.

L'approche moniste de Lewis

David Lewis propose de soutenir la thèse d'identité entre la personne et son corps en la formulant de la façon suivante : « Nécessairement, une personne occupe un corps à un temps donné si et seulement si cette personne est identique à ce corps à ce moment. » (LEWIS, 1971, p. 203) Mais dans le même temps, Lewis constate que nous pouvons produire des assertions modales significatives en admettant que nous pourrions avoir eu un autre corps. Il n'y a pas de contradiction logique à une permutation de corps.

Nous avons vu que la notion de contexte qui est au cœur de la théorie des contreparties permet de fournir une objection à l'argument de désolidarisation. Lewis, qui défend un monisme matérialiste, ne rejette, rappelons-le, ni la théorie des désignateurs rigides, ni le schéma général de l'argumentation de désolidarisation. Il rejette en revanche la thèse de nécessité de l'identité en soutenant que l'identité est contingente.

Reformulons l'argument de désolidarisation dans la théorie des contreparties de Lewis :

(A) $a = b \rightarrow \forall w\ \forall x\ \forall y\ \Big((w$ est un monde $\wedge\ x$ est dans $w\ \wedge$ y est dans $w\ \wedge\ x$ entretient une relation de contrepartie avec $a\ \wedge$ y entretient une relation de contrepartie avec $b) \rightarrow x = y\Big)$	nécessité de l'identité traduite en termes de contreparties
(B) $\exists w\ \exists x\ \exists y\ \Big((w$ est un monde $\wedge\ x$ est dans $w\ \wedge\ y$ est dans $w\ \wedge$ x entretient une relation de contrepartie $\mathcal{R}$ avec $a\ \wedge$ y entretient une relation de contrepartie $\mathcal{R}$ avec $b) \wedge x \neq y\Big)$	hypothèse d'une distinction numérique possible des contreparties
(C) Donc $a \neq b$;	*modus tolens* appliqué à (1) et (2)

Notons que l'hypothèse que a et b sont des désignateurs rigides disparaît non parce qu'elle serait rejetée, mais parce qu'elle est absolument inutile en termes de contreparties. En effet, rappelons que la théorie des contreparties stipule que tout individu est exclusivement attaché à un monde, qu'il n'existe que dans un seul et unique monde. Rappelons aussi que les relations de contreparties sont sensibles au contexte. Or l'argument précédent n'est valable que sous la prémisse (2) que la relation de contreparties utilisée l'est dans un contexte où les désignateurs a et b débouchent sur des contreparties différentes, la relation visant une contreparties sous la description du type a et l'autre sous la description du type b. Pour observer ceci de plus près, revenons au raisonnement de désolidarisation de Swinburne :

(1)	J'aurais pu être immatériel.	hypothèse
(2)	Mon corps n'aurait pas pu être immatériel.	hypothèse
	———————————————	
(3)	Donc je ne suis pas identique à mon corps.	argument de désolidarisation appliqué à (1) et (2)

La relation de contrepartie utilisée dans ce contexte est relative à deux contextes, et implique conséquemment deux descriptions différentes. La prémisse affirme ainsi que *considéré sous la description d'être immatériel*, j'ai une contrepartie qui n'est pas un corps alors que *considéré sous la description d'être matériel*, j'ai une contrepartie qui est un corps. On peut bien sûr se demander si de tels contextes sont possibles. Mais même en l'admettant, il appert de toute façon qu'ils n'ont rien de nécessaire. Car il est également possible que dans un autre contexte, une relation de contreparties « agrippe » les désignateurs a et b sous le même type de description, et dans ce cas, il advient que (A) doive être remplacée par :

(A*) $a = b \rightarrow \forall w \; \forall x \; \forall y \; \Big((w$ est un monde $\wedge \; x$ est dans $w \; \wedge$ y est dans $w \; \wedge \; x$ entretient une relation de contrepartie $\mathcal{R}$ avec $a \; \wedge$ y entretient une relation de contrepartie $\mathcal{R}'$ avec $b) \wedge x \neq y \Big)$

Dans ce cas, il n'y a plus aucune raison logique d'inférer la conclusion (C) à partir des prémisses (A*) et (B). Car comme l'écrit Lewis,

> Les prédications modales *de re* ne sont pas, en général, transparentes. La dénotation du terme-sujet n'est pas seulement ce qui importe, mais aussi la relation de contreparties qu'il sélectionne. Si nous substituons un autre terme-sujet avec la même dénotation mais un sens différent, cela pourrait changer la valeur de vérité de la prédication modale en sélectionnant une relation de contrepartie différente. Ainsi, même si la chose dénotée là dans notre monde demeure la même, nous avons une manière différente de suivre la carrière de cette chose dans d'autres mondes. (LEWIS, 2007, p. 43)

Si les propriétés étaient associées par une relation d'identité entre des mondes différents, deux propriétés différentes ne pourraient pas être identiques à une propriété actuelle. Mais c'est tout à fait possible avec une relation de contrepartie. Un même contexte d'interprétation peut dépendre de deux aspects de similarité co-sélectionnés. Les relations de contreparties, à l'opposé de l'identité, n'ont pas besoin d'être une-à-une. C'est un avantage évident de la théorie des contreparties lorsqu'il s'agit de dire que « j'aurais pu avoir un jumeau ». Lewis étend cette capacité de dédoublement à toutes les propriétés.

> Je dis aussi que ces références différentes tendent à évoquer deux façons différentes de représenter, de sorte qu'un même monde peut donner lieu à deux représentations *de re* conflictuelles d'une même chose. Vous pouvez dédoubler, ici dans un même contexte, les façons dont un

> monde représente un unique objet en plastique et en forme de bassine. Ou bien vous pouvez dédoubler les entités elles-mêmes, en distinguant le morceau de plastique de la bassine. Je soutiens que le premier dédoublement est plus crédible que le second. (LEWIS, 2007, p. 253)

Dans *Counterparts of Persons and Their Bodies* (1971), Lewis a amendé sa théorie initiale des contreparties de telle manière que plusieurs relations de contreparties puissent simultanément lier le même individu à deux individus d'un même monde. La relation de contrepartie est ainsi libérée du joug un-à-un de l'identité entre individus. Selon Lewis, les relations de contreparties doivent être vues comme des relations de similarité globales variant sous les aspects qui sont visés. En variant les facteurs d'importance des aspects sous lesquels sont considérés ressemblances et dissemblances, nous obtenons différentes relations de contrepartie. C'est ainsi que, si l'on assigne une importance élevée à l'aspect « personne », nous viserons une certaine contrepartie, et que si l'on sélectionne au contraire l'aspect « corps », nous viserons une toute autre contrepartie. L'aspect « personne » établira une relation de contrepartie-sous-l'aspect-de-personne. « Only a person, or something very like a person, can ressemble a person in respect of personhood and personnal traits enough to be his personal counterpart » (LEWIS, 1971, p. 208) De même, l'aspect « corps » établira une relation de contrepartie-sous-l'aspect-de-corps. Or, « Only a body, or something very like a body, can resemble a body in respect of bodyhood and bodily traits enough to be its bodily counterpart » (LEWIS, 1971, p. 208).

Finalement, Lewis considère la théorie de l'identité transmondaine comme un cas particulier de contreparties, où les relations de contreparties considérées sont exclusivement limitées aux origines biologiques. Mais de fait, la théorie des contreparties est plus riche, puisqu'elle ne privilégie pas d'emblée un type particulier de relation plutôt qu'une autre, ce qui permet d'exprimer plus de possibilités :

> [...] je suggère que les philosophes qui prônent l'essentialité des origines ont absolument raison - dans le contexte de leur propre discours. Ils se donnent raison : leur discours constitue un contexte dans lequel la modalité *de re* est gouvernée par une manière de représenter (une relation de contrepartie, comme je le suppose) qui exige l'égalité des origines. Mais si je demande ce qui ce serait passé si Saul Kripke n'était pas né de la fécondation d'un œuf, mais avait été apporté par une cigogne, cela a également un sens. Je crée un contexte qui donne à ma question un sens, et de la sorte ce contexte rend les origines non essentielles. (LEWIS, 2007, p. 382)

Doit-on nécessairement accepter la théorie métaphysique de Lewis pour alléguer ce genre d'objection au raisonnement de désolidarisation ? Nous pourrions rejeter tout le bagage métaphysique de la théorie des contreparties, mais en retenir l'esprit épistémique. Autrement dit, on pourrions contester que le caractère essentiel des modalités et n'en garder que le caractère descriptiviste. Nous pourrions, d'une façon toute humienne, *pace* Kripke, rabattre ainsi les modalités aléthiques sur les modalités épistémiques, autrement dit, rabattre la possibilité métaphysique sur la concevabilité. Nous ferions ainsi allégeance à Hume, quand il bascule l'essentialisme par dessus bord avec sa célèbre formule : « La nécessité est quelque chose qui existe dans l'esprit, pas dans les objets, et il ne nous est jamais possible d'en former l'idée la plus lointaine si nous la considérons comme une qualité qui se trouve dans les corps. » (HUME, 1968b, p. 1.3.14).

La théorie de Lewis échoue cependant à caractériser l'identité personnelle parce qu'elle aussi présuppose d'emblée la notion de personne. Elle peut certes expliquer pourquoi nous n'attribuons pas les mêmes prédicats aux termes de personne et d'organisme, mais elle ne parvient pas à caractériser de manière non circulaire l'identité personnelle en tant que telle.

Individuation biologique et doubles

D'autres prétendants interviennent pour associer l'identité d'une personne à son corps, comme la biologie. Mais l'individu biologique est-il la personne ? Peut-on ramener l'individuation de la personne à l'individuation biologique ?

Une première réponse consiste à la faire mettre en rapport avec la façon biologique dont l'individuation est acquise. Notamment en la faisant provenir d'une source interne, autrement dit des moyens-mêmes de reproduction du particulier. C'est ce que cautionne en particulier le paradigme de la génétique défendu notamment par Jacques Monod et François Jacob. Ces deux auteurs tendent vers une explication exclusivement interne de l'individuation des êtres vivants, renouant avec la théorie biologique de la préformation [9] au détriment de l'épigenèse [10]. Car comme l'écrit Monod :

> La structure d'un être vivant (...) ne doit presque rien à l'action des forces extérieures, mais tout, de la forme générale, jusqu'au moindre détail, à des interactions « morphogénétique » internes à l'objet lui-même. Structure témoignant donc d'un déterminisme autonome, précis, rigoureux, impliquant une « liberté » quasi totale à l'égard d'agents ou de conditions extérieures, capables certes d'entraver ce développement, mais non de le diriger, non d'imposer à l'objet vivant son organisation. (MONOD, 1970, p. 26-27)

On trouve la même idée exposée chez François Jacob :

> Pour la biologie moderne, tout être vivant se forme par l'exécution d'un programme inscrit dans ses chromosomes (JACOB, 1981, p. 23)

C'est donc un déterminisme intérieur qui l'emporte sur les conditions externes. L'individu est sujet à une *in-formation*, celle du code génétique, qui anticipe tout le développement de l'individu dans l'intimité de son intériorité.

Dans l'article intitulé « What Are Dandelions and Aphids ? » (JANZEN, 1977), le biologiste évolutionniste Daniel H. Janzen a montré que la définition de l'individuation par la compétition spécifique impliquait de reconsidérer l'acception usuelle du terme. L'individu biologique est conçu par Janzen comme ce à qui peut être attribué une valeur adaptative reproductrice. Un pissenlit particulier ne doit pas être pris pour un individu biologique parce qu'il en n'est qu'une petite partie, l'individu étant la somme totale des pissenlits ayant le même bagage génétique. En effet la reproduction asexuée du pissenlit donne lieu à des copies génétiques parfaites, et qui ne peuvent donc entrer en compétition pour l'évolution de l'espèce :

[9] Théorie faisait découler le développement individuel d'un organisme à partir d'un organisme miniature, dit parfois animalcule ou homoncule, qui préexiste à l'individu biologique dans le zygote.

[10] Théorie biologique selon laquelle les organes se développent progressivement au cours de la croissance embryonnaire, en partie sous l'influence de contraintes externes.

> Henceforth I will refer to it as the "evolutionary individual", or EI. Instead of viewing the set of short-lived dandelion plants in a habitat as a many-membered population with a very high growth rate, I suggest a quite different view. I suggest that the dandelion population contains a small number of highly subdivided EIs with very long lives and very low population growth rates and which exist through the harvest of a highly predictable resource. (JANZEN, 1977, p. 586)

Ainsi, l'individuation du pissenlit ne doit pas se faire au niveau du végétal particulier, de tel pissenlit avec sa tige et ses pétales, mais au niveau de l'ensemble complet des végétaux ayant le même bagage génétique. En arrachant un pissenlit, on arrache bien un pissenlit particulier, mais il s'agit néanmoins d'une infime partie d'un individu biologique beaucoup plus vaste.

L'identité biologique d'un individu est ainsi, la plupart du temps, définie de manière générale par l'ADN et en particulier par la notion histocompatibilité. En dépit des altérations inévitables de l'organisme, son intégrité est préservée par certains mécanismes qui procèdent en éliminant tout corps étranger, et cela, sans dommage pour le « soi ». De fait, toute partie organique étrangère greffée sur un corps est susceptible de subir un rejet. Ce n'est la greffe elle-même qui est responsable du rejet, comme l'indique la parfaite tolérance des auto-greffes. Le rejet est la conséquence de la manière dont le corps réagit afin de réguler le « soi », par la destruction de ce qui ne lui est apparenté. L'immunologie est l'étude scientifique, chez les animaux, des mécanismes qui régissent ce phénomène de destruction sélective. La réussite d'une transfusion sanguine dépend par exemple d'une compatibilité déterminée par les trois types sanguins A, B, et O. Pour ce qui concerne la greffe d'un organe, les conditions sont plus sévères : elles dépendent des groupes tissulaires HLA (antigènes des leucocytes humains) qui sont définis, en termes de biologie, par 12 caractères, dont 6 sont hérités de la mère et 6 du père. La liste complète de ces caractères définit un code, l'histocompatibilité propre au sujet. Le code HLA est apparent à la surface de chaque cellule et assure la reconnaissance de celles qui sont propres à l'organisme. Les lymphocytes tueurs discriminent ainsi les corps étrangers, qui sont ensuite fragmentés en peptides [11]. En définissant ainsi le soi par un code biologique, on le rend immuable, imperméable au changement. Les structures HLA constituent donc une « carte d'identité biologique, le sceau de notre individualité et, [...], le garant de notre intégrité » (DAUSSET, 1990, p. 20). Cette caractérisation est donc négative, définissant le soi par le rejet de ce qu'il n'est pas.

Une telle conception de l'individuation du soi par le code génétique pose cependant un certain nombre de problèmes. Entre autres, elle implique l'existence de doubles biologiques et d'individus « divisibles ». Il existe par exemple certains organismes biologiques capables de fission. Certaines plantes peuvent se défaire, sans périr pour autant, de parties qui peuvent ensuite donner vie à une nouvelle plante. Cette capacité est notamment exploitée par les techniques pépiniéristes du bouturage et du marcottage qui consistent à multiplier des végétaux au moyen d'organes respectivement séparés puis enracinés ou enracinés puis séparés. Affirmer que c'est le code génétique qui individu l'organisme vivant, c'est affirmer que, puisque la bouture est génétiquement identique à la plante d'origine, des plantes obtenues par division d'un unique bourgeon appartiennent en fait au même individu. Mais ne devrions-nous pas plutôt reconnaître, par effet de bouturage, l'apparition

[11] On pourrait cependant s'étonner que certaines greffes réussissent en dépit de la faible probabilité qu'un donneur ait exactement le même code HLA qu'un receveur. Cela tient à ce que les caractères n'ont pas tous la même importance pour les lymphocytes tueurs, et que leur action peut être inhibée de manière chimique.

de nouveaux individus après la disparition de l'original ?

Après l'invention de moyens technique d'enregistrement et de reproduction des symboles, des images, du mouvement, du son et des programmes, nous aboutirons sans doute à un nouveau stade de dupliques, celui de la reproduction des organismes vivants. La technique du clonage est une sorte d'adaptation de la technique du bouturage aux animaux : une partie infime d'un individu est utilisée pour le reproduire intégralement. En 1996, deux chercheurs de l'Institut Roslin, Ian Wilmut et Keith Campbell, sont parvenus à reproduire un mammifère à partir d'une cellule prélevée sur une brebis adulte de six ans. La naissance de Dolly représente un nouveau mode de procréation, qui est en réalité une *reproduction* au sens le plus strict du terme. Cette technique sera-t-elle un jour applicable aux humains ? C'est très probable. Quoi qu'il en soit, certaines œuvres de fiction ont anticipé cette possibilité à des fins narratives [12]. En outre, la technique du clonage a d'ors et déjà été appliquée en médecine régénératrice à des fins de greffes pour reproduire une partie d'un organisme plutôt que l'intégralité d'un individu. En 2018, l'équipe de Susan Kimber, de l'université de Manchester est parvenue à produire un mini-rein. Des équipes de biologistes étudient actuellement la reproduction en laboratoire d'organes à partir de cellules souches, c'est-à-dire de cellules indifférenciées capables de produire tout type de cellules spécialisées.

Mais plutôt que d'anticiper sur l'existence future des clones, nous pouvons aussi nous intéresser aux cas des jumeaux monozygotes [13]. L'étude des jumeaux est particulièrement intéressante pour l'analyse de la genèse de l'identité personnelle, ou psychogenèse. De fait, des jumeaux monozygotes partagent, par définition, un même patrimoine génétique et partagent aussi, en toute vraissemblance, un même patrimoine mémétique. Or ils ne sont jamais psychologiquement identiques, et témoignent souvent de personnalités bien contrastées [14]. Leur indiscernabilité génétique et leur environnement culturel commun permet de questionner l'élaboration de leur identité personnelle en analysant les écarts psychologiques qui les séparent, en déterminant ce qui relève au juste de l'hérédité et du milieu. Et de toute évidence, leurs différences psychologiques montrent que l'identité personnelle ne peut être réduite à l'identité biologique.

Néanmoins, le paradigme de l'individuation biologique par l'ADN est aujourd'hui très conversé. La biologie moderne considère que l'individu biologique est un individu évolutionnaire. Autrement dit, c'est le paradigme de la théorie de l'évolution qui édicte maintenant ce qui fait partie

[12]Notons que des films comme *Alien IV* (JEUNET, 1997), *The Island* (BAY, 2005) ou encore *Us* (PEELE, 2019) présentent, par rapport au thème romantique des *doppelgängers*, une inversion du rôle de la duplique dans le récit. Car c'est maintenant la duplique qui occupe le rôle principal, c'est elle qui se découvre, au gré de diverses péripéties, être la reproduction d'un individu original dont elle est destinée à servir de pièce de rechange. Parvenu au stade technique de la reproduction à son étiage le plus élevé, celui des êtres vivants, la perspective du double est ainsi renversée : l'angoisse est maintenant éprouvée par la copie, et non par celui qu'elle copie.

[13]Les croyances erronées au sujet des jumeaux, qu'elles soient plus ou moins folkloriques ou scientifiques, sont nombreuses et variées. Il vaut la peine de remarquer que parmi celles-ci existe une théorie affirmant que les jumeaux ne sont pas des cas exceptionnels, mais la règle générale. Tous les individus naîtraient en deux exemplaires. Les naissances uniques s'expliquent par la mort prématurée de l'un des jumeaux dans le corps de la mère. Mais tout individu serait le jumeau d'un autre, quitte à l'ignorer. Cette théorie apparaît notamment dans les croyances des Dogons d'Affrique, mais aussi dans les travaux de psychologie génétique d'Arnold Gesell comme l'indique René Zazzo (ZAZZO, 1993). Notons en outre qu'une coutume germanique, liée aux pratiques obstétriques traditionnelles, consiste à enterrer le placenta d'un nouveau-né auprès d'un arbre qui représentera, au long de sa vie, le double de cette personne.

[14]En outre, même s'ils partagent le même patrimoine génétique, il arrive que des jumeaux aient, dès leur naissance, des différences physionomiques. On peut en particulier constater un certain nombre d'asymétries. Une tâche de naissance, un grain de beauté, peut être située à gauche sur l'un des jumeaux, tandis qu'elle apparaît à droite sur l'autre.

ou non de la catégorie d'individu biologique, et non celle d'une seule essence codée au creux des cellules. Selon ce paradigme, l'individu est vu comme *unité d'interaction* dans le processus causal de la sélection naturelle. En ce sens, l'individualité concerne aussi bien les gènes, les cellules, les organismes que les espèces. Et cela implique, à nouveau, une hiérarchie de l'individuation fondée sur l'inféodation de la partie au tout. Mais il semble que le milieu environnant soit également un facteur important de la différenciation. De fait, il y a aussi une individuation dont la source est externe : par exemple par les accidents que connaît la carrière d'un individu. Deux voitures semblables sortant de la même chaîne de production ne sont pas individuées par leur mode de production, mais par les modifications, usures, rayures, qu'elles subissent durant leur carrière temporelle. C'est aussi ainsi que l'épigénétique décrit les différenciations entre jumeaux monozygotes : ils sont bien reconnus comme étant des individus particuliers, mais leur individuation provient de sources externes.

Dans *La mosaïque du soi : les chimères en immunologie*, Thomas Pradeu désigne les individus biologiques comme des « mosaïques », cela, en deux sens différents. Le chimérisme [15] désigne en biologie la présence de cellules génétiquement différentes au sein d'un même individu. Ce concept biologique été introduit en 1945 par la découverte de Ray Owen (OWEN, 1945) de deux veaux non monozygotes partageant le même placenta, et conservant chacun certaines cellules de l'autre. Ce partage leur permettait de tolérer des échanges de tissus sans rejet immunitaire. Des pratiques de chimérisme artificiel ont ensuite été produites en 1960 par Nicole Le Douarin qui introduisit avec succès des cellules d'embryon de caille dans l'embryon de poulet. Frank Burnet (BURNET, 1969) a soutenu que cette cohabitation est rendue possible par période située tôt dans la vie de l'organisme, durant laquelle il admet une certaine tolérance à l'égard des éléments exogènes sans réaction immunitaire, ce qui lui permet une certaine extension du soi. Intervient ensuite une période, définitive, de « fermeture du soi ». C'est durant cette seconde période que l'organisme acquerrait son immunocompétence. Selon ce modèle, la tolérance à l'égard des entités admises lors de la première période perdurerait, puisqu'elle serait antérieure aux mécanismes biologiques de reconnaissance du moi, mais de nouveaux éléments exogènes déclencheraient des réponses immunitaires.

Cette explication a été cependant remise en cause à partir des années 90 par la découverte de formes de chimérisme tout au long de la vie. Pradeu soutient que le chimérisme est « ubiquitaire chez les vivipares » (PRADEU, 2009, p. 24). D'abord, Pradeu attire l'attention sur le fait que les mères conservent toute leur vie des cellules de leur enfant, et réciproquement, que les enfants conservent des cellules de leur mère. Il en déduit que :

> On peut en conclure que l'identité d'un organisme n'est pas définie une fois pour toutes, ni dans ses gènes, ni dans un quelconque programme, ni non plus après la période d'immaturité immunitaire. Nous ne sommes sans doute pas toute notre vie aussi ouverts à l'environnement que l'est un embryon, mais nous ne sommes pas pour autant complètement fermés à toute influence exogène. (PRADEU, 2009, p. 29)

Le chimérisme semble en complète contradiction avec la notion paradigmatique du soi et du non-soi biologique, définit par l'immunologie. Or pour Pradeu, cela montre que la notion d'indi-

[15] Rappelons qu'à l'origine, le mot désigne, dans la mythologie grecque, une entité constituée d'une tête de lion, d'un corps de chèvre et dont la queue est un serpent.

vidu comme unité biologique indépendante n'a tout simplement pas de sens et que « nous sommes tous des chimères » (PRADEU, 2009, p. 19). De fait, cette identité biologique ouverte implique qu'un organisme devrait être vu moins comme une unité autonome et indépendante que comme un être par nature double et composite :

> Le chimérisme montre qu'un organisme est une entité hétérogène. Etymologiquement, l'hétérogène ne désigne pas seulement une diversité, mais l'origine extérieure de la diversité, c'est-à-dire le fait de « venir de l'autre » (en l'occurrence, de venir de l'environnement). Le chimérisme montre que l'organisme est bien plus que la somme des constituants issus de la cellule œuf : il accueille en lui de l'étranger, qui devient une part de lui-même. (PRADEU, 2009, p. 24)

La théorie immunologique de l'identité du Moi est donc contestable. Les jumeaux et les clones ne sont certainement pas des répliques pures au sens de personnes, et la personne n'est donc pas réductible à l'individu biologique.



Définir l'identité personnelle par l'identité d'une substance matérielle ou non, mène à une voie sans issue. Quelle autre possibilité nous reste-t-il alors ? Nous pouvons ramener la question « à quelles conditions une personne reste-t-elle la même ? », question qui présuppose d'emblée la notion qu'elle prétend définir, à une autre : à quelles conditions une personne survit-elle ? Ce changement d'interrogation est motivé par la conception ontologiquement neutre de la notion de personne initiée par John Locke, que nous allons analyser dans la section suivante.

7.2 Pour un désengagement ontologique de la notion de personne

Depuis Locke, des philosophes ont développé l'idée que l'identité personnelle n'est rien, ou du moins, qu'elle n'importe pas. Ce qui compte, c'est la persistance d'un flux psychologique continu. Il s'agit donc d'une thèse opposée à aux théories substantialistes définissant la personne par l'âme ou par le corps. Derek Parfit souligne que ce qui compte, ce n'est finalement pas l'identité personnelle, mais la survie de la personne. Le réductionnisme de Parfit (PARFIT, 1984) a pour thèse que les faits les plus fondamentaux à propos d'une personne peuvent être décrits de manière impersonnelle, c'est-à-dire, sans faire appel à un point de vue subjectif. En ce sens, notre conception de l'identité personnelle comme individuation serait complètement erronée. La subjectivité n'aurait fondamentalement aucune importance en ce qui concerne la notion de personne. La conséquence qu'en tire Parfit est que, contrairement à l'identité personnelle qui bride l'individu à l'unité, la thèse de survie implique qu'un même individu puisse survivre dans *plusieurs personnes*, et donc, avoir des doubles.

Comment adopter un point de vue ontologiquement neutre sur l'identité personnelle ? Nous commencerons par re-situer la théorie de Locke et par commenter ses objections. Nous évaluerons ensuite la conception que Parfit a tiré de l'identité psychologique.

7.2.1 Mémoires doubles ou double mémoire

La conception réductionniste de l'identité personnelle

Dans la querelle au sujet de la nature des personnes qui oppose d'une part la conception de Descartes, fondée sur le dédoublement entre substance étendue et substance immatérielle et d'autre part le matérialisme de Hobbes, la position de Locke est neutre quant au type de substrat porteur des propriétés psychologiques. Selon Locke, le critère d'identité de conscience est indépendant de tout engagement ontologique, car nous ignorons si la pensée provient ou non de la matière : « Il est indéniable que nous avons en nous quelque chose qui pense : nos doutes même sur ce que c'est confirment la certitude de son être, bien que nous devions nous contenter de l'ignorance sur le genre d'être dont il s'agit » (LOCKE, 1998b, IV, chap. 3, § 6).

En conséquence, c'est par une voie exclusivement psychologique que Locke démarque le concept de personne, ou plutôt qu'il renonce à en définir la nature ontologique. Dans l'*Essai concernant l'entendement humain* (1689), Locke définit la personne comme un « être pensant, intelligent, qui a raison et réflexion et qui peut se regarder soi-même comme soi-même, comme la même chose qui pense en différents temps et lieux ; ce qu'il fait uniquement par la conscience qui est inséparable de la pensée, et qui lui est à mon avis essentielle (car il est impossible à quiconque de percevoir sans percevoir qu'il perçoit) » (LOCKE, 1998b, pp. 521–522). L'une des conséquences importantes de la caractérisation psychologique du Moi de Locke est l'égocentrisme, c'est-à-dire la capacité du sujet à recentrer ses pensées sur son propre moi. Il est en effet tout à fait essentiel à l'identité personnelle du sujet qu'il parvienne à s'auto-identifier. Contrairement à Descartes, chez Locke, la personne n'est donc pas définie par une substance pensante, mais par le fait qu'elle « peut se regarder soi-même comme soi-même », c'est-à-dire dans sa capacité à se traiter à la première personne, à parler de soi en tant-que-soi. D'une certaine manière, Locke retient la quintessence du *cogito*, la saisie d'un Soi au creux d'une pensée, mais il refuse de délibérer de sa nature ontologique.

Ayant délimité les contours de la notion de personne, Locke propose ensuite de lui conférer un critère d'identité :

> (…) puisque la conscience accompagne toujours la pensée, et que c'est là ce qui fait que chacun est ce qu'il nomme soi-même, et par où il se distingue de toute autre chose pensante : c'est aussi en cela seul que consiste l'*identité personnelle*, ou ce qui fait qu'un Être raisonnable est toujours le même. (LOCKE, 1998b, p. 522)

La persistance d'une même personne ne relève donc plus de la persistance d'une entité à la territorialité propre, celle d'une âme ou d'un corps. Elle relève d'une continuité psychique de souvenirs, du sentiment continu d'être le même. De ce point de vue, la démarcation du moi n'a pas de limites matérielles mais mémorielles.

Critiques de la définition psychologique de l'identité personnelle

Dès son apparition dans les écrits de Locke, le critère psychologique d'identité personnelle a fait l'enjeu de critiques *via* des expériences de pensée destinées à éprouver sa résistance. C'est

d'ailleurs Locke lui-même qui inaugura cette méthodologie heuristique en élaborant différentes hypothèses contrefactuelles afin d'infirmer la théorie substantialiste de l'identité personnelle. En particulier, la continuité mémorielle encourage d'envisager des expériences de pensées dans lesquelles sont permutés les corps de différentes personnes.

Que serait, pour l'identité personnelle, une âme sans souvenir ? Rien, le partage d'une même âme serait tout à fait inutile. En effet, dans une première expérience de pensée, Locke montre par l'absurde que l'âme ne peut être tenue responsable de l'identité personnelle. Locke admet le cas de deux corps séparés dans le temps, mais informés par la même âme, comme dans le cas d'une réincarnation. Admettons par exemple que l'âme d'Emmanuel Macron soit la même que celle de Nestor, le personnage de la mythologie grecque. Si Macron n'a pas le souvenir, en première personne, d'avoir participé à la guerre de Troie, se pourrait-il qu'il soit la même personne que Nestor ? Le critère d'identité assure que la réponse est négative puisque la continuité mémorielle est une condition nécessaire à l'identité. En outre, la continuité mémorielle est également une condition suffisante à l'identité. Ainsi, si une personne avait le souvenir vrai, en première personne, d'avoir vu l'arche de Noé durant le Déluge, cette personne serait manifestement la même que le personnage dont il est question dans l'ancien testament.

Locke indique encore une seconde expérience de pensée qui, elle, indique que le corps matériel n'est pas plus apte à assurer l'identité. Admettons que deux consciences « distinctes et incommunicables » (LOCKE, 1998b, II, chap. 27, §23) qui agiraient dans un même corps, de manière alternée, l'une de jour et l'autre de nuit, un peu à la manière du docteur Jekyll et de Mister Hyde. Bien qu'il n'y aurait qu'un seul corps, il faudrait néanmoins reconnaître deux personnes différentes, aussi distinctes entre elles que Socrate l'est de Platon.

Dans une troisième expérience de pensée, Locke donne envisage une conscience agissant de manière intermittente dans deux corps différents. Locke conclut qu'on aurait affaire à une unique personne dispersée dans deux corps différents, « tout comme un homme reste le même dans des costumes différents » (LOCKE, 1998b, II, chap. 27, §23).

Une quatrième expérience de pensée est proposée par Locke. C'est sans doute la plus connue car la plus discutée. Elle met en scène l'esprit d'un Prince transféré dans le corps d'un savetier (voir l'illustration 7.1) et l'esprit du savetier transféré dans le corps du Prince. Après substitution des esprits, à quelle personne aurions-nous affaire en présence du corps contenant l'esprit du Prince ? Devrait-on reconnaître le Savetier en donnant l'avantage au critère d'identité matériel, ou plutôt reconnaître le Prince en faisant au contraire primer le critère d'identité psychologique ? D'après Locke, cela dépend de l'aspect sous lequel nous visons l'individu. En tant qu'homme, il faudrait donner raison au critère d'identité matériel : il s'agirait du Savetier (avec l'esprit du Prince). Mais en tant que personne, c'est le critère d'identité personnelle qui primerait et il s'agirait du Prince dans le corps du savetier.

De telles expériences de pensée révèlent qu'avec un tel critère d'identité, la question de la nature de la substance devient superflue, parce qu'elle en est manifestement indépendante de l'identité personnelle. Ces expériences de pensées convergent en effet vers l'idée que le critère d'identité personnelle est indépendant de tout engagement ontologique à l'égard de la nature des personnes. Le critère de Locke dissocie ainsi l'identité matérielle (l'homme) de l'identité psychologique (l'esprit). En comparant le Prince et l'individu constitué du corps du savetier et de l'esprit du Prince,

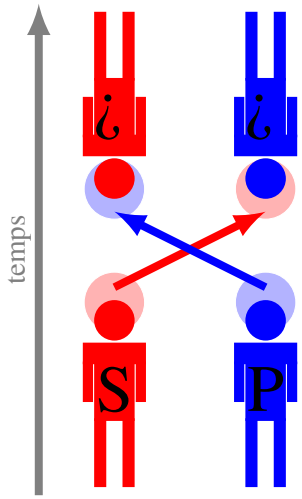


FIGURE 7.1 : Les « auras » représentent les esprits respectifs du savetier et du prince

l'on dira qu'il s'agit de la même personne, mais pas du même homme, ou bien l'on dira qu'il s'agit du même homme, mais pas de la même personne.

On pourrait objecter que cette caractérisation de l'identité enfreint le principe d'indiscernabilité des identiques. En effet, si le Prince dans le corps du Prince est la même personne que le Prince dans le corps du savetier, alors, par application du principe d'indiscernabilité des identiques, ces deux individus devraient partager *toutes* leurs caractéristiques, et parmi elles, celle d'être le *même homme*. Or, le Prince dans le corps du Prince n'est pas le même homme que le Prince dans le corps du Savetier. On retrouve donc ici une forme de relativisation de l'identité, puisqu'un individu a peut être le même F que b sans être le même G que b.

Examinons plus en détail la définition de l'identité personnelle par la continuité psychologique. On pourrait l'interpréter ainsi : pour deux moments $t_1 < t_2$, l'individu y à t_2 est la même personne que l'individu x à t_1 si et seulement si y s'attribue, à la première personne, toutes les actions de x « *comme des actions qui lui sont propres*, aussi loin que cette conscience s'étend et pas plus loin. Autrement dit, si y se souvient, à t_2 d'avoir été la personne x à t_1. Une conséquence juridique importante en découle : la responsabilité des actes d'un individu doit être attribuée à la personne, non à son corps. Ainsi,

> Que si le même Socrate veillant et dormant ne participe pas à une même conscience, Socrate veillant et dormant n'est pas la même personne. Et il n'y aurait pas plus de justice à punir Socrate veillant pour ce qu'aurait pensé Socrate dormant, et dont Socrate veillant n'aurait jamais eu aucun sentiment, qu'à punir un jumeau pour ce qu'aurait fait son frère et dont il n'aurait aucun sentiment, parce que leur extérieur serait semblable qu'on ne pourrait les distinguer l'un de l'autre ; car on a vu de tels jumeaux. (LOCKE, 1998b)

Objections à l'identité psychologique

Il découle de la définition de Locke qu'une personne demeure la même tant qu'elle éprouve, à la première personne, le sentiment d'être, tout au long du temps, la même. D'une certaine manière l'identité personnelle est définie par un *sentiment d'identité*. Ce critère nécessite donc que les souvenirs d'une personne s'étendent aussi loin que son existence. Or cela implique certaines difficultés. Il se pourrait en effet qu'un individu perde ce sentiment d'identité en raison d'une déficience de mémoire, comme c'est par exemple le cas des personnes atteintes de la maladie d'Alzheimer. Et il pourrait également arriver que ce sentiment d'identité, quoique véritablement éprouvé par un individu, le soit cependant de manière indue. Dès sa publication, la théorie psychologique de l'identité personnelle de Locke a été sujette à de nombreuses critiques particulièrement axées sur ces deux objections que nous allons maintenant détailler.

Le problème de la discontinuité des souvenirs a été soulevé par Thomas Reid (REID, 2011, III, chap. 6). L'argument consiste à imaginer une personne, à trois âges différents, ayant des souvenirs fluctuants. Un officier pourrait se rappeler d'avoir été puni, enfant, d'un larcin qu'il avait commis. Parvenu à la retraite, le soldat pourrait avoir définitivement perdu ce souvenir, tout en ayant en mémoire sa carrière d'officier. Le critère de continuité mémorielle étant une condition suffisante à l'identité personnelle, il assurerait que l'adulte serait bien identique à l'enfant et que le vieillard serait identique à l'officier. En revanche, la condition de continuité de mémoire étant nécessaire à l'identité, il viendrait alors que le vieillard ne serait pas la même personne que l'enfant. Or, pour servir de critère d'identité, la continuité psychologique devrait être transitive, ce qui n'est manifestement pas le cas dans la situation envisagée ici.

Une deuxième objection au critère d'identité lockéen a été adressée par Bishop Butler. Elle consiste à souligner la circularité qui réside dans le fait de définir l'identité personnelle par la continuité psychologique. On ne peut tirer l'identité personnelle de la conscience qu'en présupposant d'emblée l'identité de cette conscience. Selon Butler, ce n'est pas parce qu'une personne possède des souvenirs d'une de ses occurrences temporelles qu'elle lui est identique. C'est parce qu'une personne est identique à ses occurrences temporelles qu'elle a la possibilité de s'attribuer des souvenirs (BUTLER, 1975, pp. 99-105). L'identité est plus fondamentale, affirme Butler, que la continuité mémorielle. La première explique la seconde, et non l'inverse. Le concept d'identité, soutient Butler, est logiquement premier par rapport au critère psychologique.

Par ailleurs, Butler souligne que la caractérisation de Locke revient à dire qu'une personne n'a existé qu'aux moments où elle était consciente d'elle-même et dont elle se souvient. L'identité d'une personne serait alors discontinue, du fait des phases de sommeil, de pertes de mémoire, d'épisode de folie ou de périodes de coma. C'est sans doute tout à fait juste, mais est-ce que cela constitue vraiment une objection ?

La définition de l'identité personnelle de Locke est bien sûr nominale, ce qui implique qu'elle ne dit rien sur la nature métaphysique des personnes, sur leur essence. Comme l'indique Bernard Williams (WILLIAMS, 1973, p. 14), la question est alors de savoir s'il l'on peut effectivement admettre qu'un amnésique n'ayant aucun souvenir, mettons par exemple un patient très sévèrement atteint par la maladie d'Alzheimer, serait considéré par la communauté comme n'ayant plus d'identité personnelle. Mais cette objection semble mélanger deux notions, d'une part l'hypothèse d'une essence métaphysique, d'une nature des personnes, et d'autre part, un moyen épistémique

d'identification. Dans le cas de l'Alzheimer, nous perdrions sans doute un moyen d'identification de la personne (mais évidemment pas de son corps). Le critère ne nous permettrait plus d'identifier notre Alzheimer comme une certaine personne [16]. Il me semble qu'il s'agisse moins d'une objection que d'un constat : l'identité personnelle n'est rien en dehors d'un crible de reconnaissance, elle est seulement épistémique, mais pas métaphysique. Il pourrait cependant être tentant, pour le partisan de l'identité métaphysique, de considérer que la personne est constituée non pas sur la continuité de mémoire, mais par certaines qualités psychologiques, des goûts récurrents, un tempérament, ou plus généralement une personnalité qui résisteraient à la perte de mémoire, comme c'est souvent le cas pour les malades d'Alzheimer, dont de nombreux traits persistent dans la maladie. Mais en ce cas, le métaphysicien devrait admettre que deux individus ayant les mêmes qualités psychologiques ne formeraient qu'une seule personne, ce qui semble absurde.

La condition suffisante d'identité par la continuité mémorielle se heurte également à la question des identifications erronées que nous avons esquissée plus haut. Le souvenir n'est pas prémuni contre l'erreur d'identification. L'auto-identification d'un sujet est nécessaire à son identité personnelle. En revanche, il est douteux que cette condition soit suffisante, puisque le fait de s'auto-identifier subjectivement à Napoléon ne fait pas véritablement de soi Napoléon. De fait, le critère de Locke encourage l'idée que ce qui importe pour l'identité d'une personne, ce n'est pas l'identité métaphysique que l'impression de continuité d'un flux psychique éprouvé à la première personne. Si l'identité d'une personne repose sur le fait qu'elle, en première personne, se pense être cette personne-là, comment distinguer alors entre les cas où elle croirait à tort avoir vécu les faits d'un certain souvenir à la première personne qui appartiendrait en fait à quelqu'un d'autre ? Une telle hypothèse est illustrée dans le roman *Do Androïds Dream of Electric Sheep ?* de Philip K. Dick (DICK, 1968), dans lequel des androïdes sont programmés avec de faux souvenirs provenant de véritables humains dont ils constituent des répliques. Les répliquants possèdent leurs souvenirs à la première personne, sans les avoir réellement vécus dans leur chair. Rachael, un androïde adulte, possède ainsi les souvenirs de la nièce de son créateur Tyrell. Le critère de Locke incite alors à identifier Rachael à la nièce de Tyrell, ce qui illustrerait une radicalisation de la troisième expérience de pensée de Locke évoquée plus haut. Nous déboucherions en effet sur le cas d'une personne vivant dans deux corps différents non plus de manière intermittente, mais simultanée.

La médecine reconnaît différentes pathologies de l'identité personnelle dans lesquelles un patient devient incapable de s'attribuer des souvenirs. Rappelons ici les travaux séminaux de Pierre Janet autour des troubles dissociatifs, classés en amnésie dissociative, trouble de l'identité et dépersonnalisation. Ces troubles sont globalement caractérisés par l'effet de dissociation entre certaines fonctions psychiques et le reste de la personnalité. Précisément, ils sont diagnostiqués lorsqu'un individu devient incapable de rattacher à son « Moi » les souvenirs, les perceptions externes et internes, et les représentations et affections dont il fait l'expérience. On peut ainsi admettre l'hypothèse d'une personne qui serait régulièrement prise de crises délirantes durant lesquelles elle se prendrait pour quelqu'un d'autre, sans que cette autre personnalité ne sache qui elle est. Considérons par exemple le cas d'une personne x, soufrant de troubles de personnalité multiple, qui s'identifierait à un autre individu y. Admettons en outre que y ne s'identifierait jamais à x, voire

[16] C'est d'ailleurs parfaitement cohérent avec ce que vivent les personnes familières d'un individu atteint de maladie d'Alzheimer, qui, souvent, ne se plaignent pas tant de ne plus être reconnus, mais plutôt de ne plus parvenir, eux-mêmes, à le reconnaître.

que y ignorerait complètement être x. Combien de personnes devrions-nous dénombrer ? Considérons le raisonnement suivant :

(1)	x sait qu'il est y	hypothèse
(2)	y croit qu'il n'est pas x	hypothèse
	———————————————	
(3)	$x = y$	de (1) en appliquant le principe de connaissance vraie
(4)	$y \neq x$	de (2) parce que l'auto-identification est nécessaire à l'identité personnelle

Comme on peut le remarquer, puisqu'on a à la fois que $x = y$ et que $y \neq x$, le critère de continuité mémorielle nous conduit ici non plus à une transitivité défaillante, comme dans le cas exposé par Reid, mais à une anti-symétrie.

Réponses aux objections

Quelles réponses donner aux objections précédentes ? Sydney Shoemaker (S. SHOEMAKER, 1970) a montré que l'objection de discontinuité peut être levée en affirmant que le critère de Locke ne demande pas la restitution complète du souvenir des épisodes vécus, mais seulement l'existence d'une relation mémorielle assurant une forme de continuité entre les épisodes remémorés. La relation mémorielle ne repose pas sur une identité de conscience, mais sur une chaîne continue de connexions entre différents moments remémorés de l'existence d'un individu. L'identité psychologique ne demande pas une mémoire globale des événements vécus, mais suffisamment de chevauchements des souvenirs entre des périodes contiguës. Le problème de la non-transitivité de continuité des souvenirs est ainsi obvié.

Quant à l'objection de circularité énoncée par Butler, elle peut être évitée par l'introduction des quasi-souvenirs. À la différence du souvenir qui nécessite toujours une représentation à la première personne de l'événement produit, le quasi-souvenir d'un événement nécessite seulement un témoin. La notion de quasi-souvenir est donc constituée à partir de celle de souvenir, mais en rejetant la contrainte d'auto-attribution : le quasi-souvenir est un souvenir impersonnel, il s'agit d'un souvenir *de re*. Il n'est pas une connaissance *de se* à l'abri de toute erreur d'attribution. Plus précisément, Shoemaker définit le quasi-souvenir de la manière suivante : (D. SHOEMAKER, 1984, pp. 81–84) :

- on a l'impression d'avoir le souvenir d'un événement ;
- le contenu de ce souvenir correspond à un événement vécu par quelqu'un ;
- ce souvenir apparent est causalement dépendant, d'une manière appropriée, de l'événement en question.

La première condition du quasi-souvenir conserve une certaine indétermination par l'emploi du pronom indéterminé « on » afin d'éviter d'attribuer directement un souvenir à une personne particulière, ceci afin de contrer l'objection de circularité. Il n'y a plus d'auto-attribution du souvenir par le moyen d'un indexical. L'idée est de considérer qu'il suffit que n'importe qui soit, à la première personne, le témoin de l'événement, ce qui ne présuppose donc pas l'identité d'une personne en particulier. Il est souvent objecté que la notion de quasi-souvenir est plutôt étrange, sinon absurde : comment admettre en effet qu'un souvenir puisse être à la première personne sans être nécessairement le souvenir de cette même personne ? Mais nous avons vu à travers l'exemple de *Blade Runner* cité plus haut que les quasi-souvenirs n'ont cependant rien d'illogiques, qu'on peut tout à fait concevoir des scénarios présentant de tels types de souvenirs. Dans le film de Ridley Scott , la répliquante Rachael se souvient avoir vu elle-même une araignée dévorée par ses progénitures, sans que cela soit vraiment le cas, ce souvenir appartenant en réalité à la nièce de Tyrell dont Rachael a hérité la mémoire. Rachael a bien le souvenir *à la première personne* d'une araignée dévorée, mais ce n'est pas *son* souvenir. Ce que montre cette illustration, c'est qu'il est logiquement possible qu'un souvenir soit éprouvé à la première personne tout en appartenant à une autre personne. Il y a donc une forme d'indépendance entre le fait d'avoir le sentiment d'un souvenir à la première personne, et l'identité de la personne rattachée à ce souvenir. La notion de souvenir peut être désolidarisée du « soi » qui a ce souvenir. Autrement dit, certaines pensées peuvent être à la fois *impersonnelles et subjectives*.

La caractérisation psychologique de l'identité personnelle implique l'idée qu'il n'est pas logiquement impossible qu'une personne change de corps. Or, si l'on réfléchit, d'une manière toute cartésienne, en termes d'amputations successive, nous sommes conduits à concevoir une limite organique de la persistance d'une personne. En effet, contrairement aux autres organes, l'amputation des hémisphères cérébraux a pour conséquence la destruction de la personne. L'hypothèse la plus crédible pour l'identité personnelle est donc de localiser les personnes à l'emplacement de leur cerveau. Aussi, une expérience de pensée de Sydney Shoemaker (S. SHOEMAKER, 1963, p. 23) considère les circonstances favorables à un changement de cerveaux entre deux personnes, par transplantation [17]. Deux hommes, Brown et Robinson, font l'objet d'une permutation cérébrale. Le cerveau de Brown est placé dans la tête de Robinson et inversement, le cerveau de Robinson est placé dans la tête de Brown, comme illustré sur la figure 7.2. Admettons ensuite que le corps duquel a été extrait le cerveau de Brown décède en raison de complications de l'opération. Appelons « Brownson » le survivant, c'est-à-dire le corps de Robinson greffé du cerveau de Brown. De quelle personne s'agit-il maintenant ? Est-ce Brown, est-ce Robinson, est-ce les deux à la fois ou n'est-ce aucun d'entre eux ?

Nous pourrions attendre du survivant le comportement suivant : il désignerait sans doute le cadavre en disant qu'il s'agit là de son propre corps. Brownson aurait probablement le sentiment d'être Brown, avec l'impression que son corps a été remplacé par celui de Robinson. Si ce n'était pas le cas, autrement dit, si Brownson s'identifiait plutôt à Robinson qu'à Brown, nous serions quittes à réviser nos conceptions sur l'importance que joue le cerveau pour le sentiment de continuité psychologique. C'est-à-dire que nous devrions remettre en question les liens présumés entre

[17] Les premières tentatives de greffes datent du XIXe siècle et concernaient les reins. Des succès de greffes de poumons et d'yeux, furent obtenus dans les années 60. Le professeur Christiaan Barnard réussit en 1967 la première greffe cardiaque. Il fut bientôt suivi, en 1968, d'une seconde greffe de cœur, réalisée à Paris par le professeur Christian Cabrol.

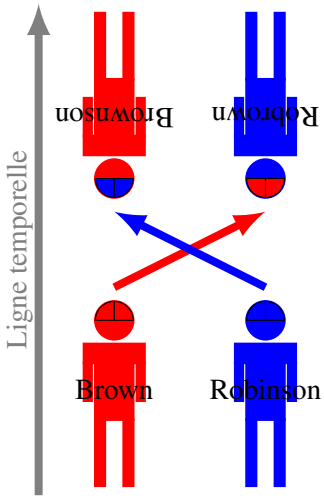


FIGURE 7.2 : Permutation des cerveaux

les états cérébraux d'un individu et ses états psychologiques.

D'après Shoemaker, ce que montre cette expérience de pensée, ce n'est pas tant qu'une âme, un esprit, une conscience ou quoi que ce soit d'autre migre, *via* la transplantation cérébrale, d'un corps à un autre. C'est plutôt que le corps est tout entier réductible au cerveau. La bonne interprétation de cette expérience de pensée n'est pas d'y voir un cerveau greffé sur une autre personne, mais plutôt de considérer que tous les organes manquants au cerveau d'une personne sont greffés autour d'elle. Finalement, cette expérience de pensée est similaire à celle consistant à imaginer qu'on remplace successivement tous les organes de Brown, à l'exception de son cerveau, par ceux de Robinson, à la manière des substitutions dont il est question dans les paradoxes du bateau de Thésée et de Chisholm. La théorie de Shoemaker conduit donc à un réductionnisme physique : l'identité de Brown est réductible à l'identité de son cerveau.

Identifier ainsi les personnes à leur cerveau a le mérite de donner une réponse aux situations ambiguës liées aux aberrations tératologiques. De ce point de vue, de présumés frères siamois dotés d'un unique cerveau ne forment véritablement qu'une seule personne aux membres dédoublés. En revanche, un corps d'apparence normale mais doté de deux cerveaux autonomes, voire d'hémisphères autonomes, devrait être considéré comme logeant deux personnes distinctes. En outre, l'identité personnelle est finalement ramenée ici à l'identité de corps, sinon que le corps est ici réduit aux dimensions du cerveau.

Les individus fissionnés

Le critère d'identité personnelle par identité des cerveaux est confronté aux expériences de pensée qui admettent le cas d'individus fissionnés. De tels individus éclatés semblent logiquement possibles. Dans des conditions favorables, les morceaux d'une amibe tranchée en deux sont capables de régénérer leurs parties manquantes. Ce cas de figure conduit ainsi à la production de

deux amibes complètes, dont on peut se demander, pour chacune d'elles, quelle est au juste la relation qui la lie à l'amibe originale. De telles questions peuvent être étendues à l'identité des personnes. Cette éventualité a d'ailleurs été exploitée dans la littérature du double. Dans la nouvelle de Nicholas Gogol intitulée *Le Nez* (GOGOL, 1966), le major Kovaliov constate, à son réveil, la stupéfiante disparition de son nez, dont la place habituelle est devenue une surface parfaitement lisse. Cependant, Kovaliov ne tarde pas à retrouver son nez au détour d'une rue. Celui-ci est vêtu comme un conseiller d'état. Kovaliov cherche maladroitement à convaincre son nez de reprendre sa place habituelle, invoquant sa charge de major, et les obligations mondaines qui sont les siennes et qui ne peuvent se passer d'un nez au beau milieu du visage : « je suis surpris, Monsieur, vous devriez, il me semble, un peu mieux connaître votre place... puisque je vous retrouve... avouez que... (...) Enfin Monsieur, n'êtes-vous pas mon propre nez ? » (GOGOL, 1966, p. 603-604) Mais le nez objecte aux sollicitations pressantes de Kovaliov qu'il « n'appartient qu'à lui-même », ne manquant pas, au passage, de faire remarquer à son interlocuteur que la différence de leurs uniformes manifeste une évidente appartenance à des classes sociales distinctes. Le récit de Gogol illustre donc un cas de duplication par fission d'une personne fragmentée par dé-solidarisation de ses parties matérielles. Or Locke évoque lui-même ce cas de figure, sinon que chez lui, il ne s'agit pas d'un nez, mais d'un doigt :

> Si la conscience suivait le petit doigt, lorsqu'il vient à être coupé, le même soi qui hier était intéressé pour tout le corps, comme faisant partie de lui-même, ne pourrait que regarder les actions qui furent faites hier, comme des actions qui lui appartiennent présentement. Et cependant, si le même corps continuait de vivre et d'avoir, immédiatement après la séparation du petit doigt, sa conscience particulière à laquelle le petit doigt n'eût aucune part, le soi attaché au petit doigt n'aurait garde d'y prendre aucun intérêt comme à une partie de lui-même, il ne pourrait avouer aucune de ses actions, et l'on ne pourrait non plus lui en imputer aucune. (LOCKE, 1998b, p. 170)

En 1967, David Wiggins (WIGGINS, 1967, p. 50) prolonge l'expérience de pensée des cerveaux transplantés en imaginant la scène suivante : chacun des deux hémisphères d'un même cerveau est transplanté à l'intérieur de deux corps comme l'illustre la figure 7.3. En faisant l'hypothèse que les souvenirs d'un individu sont « stockés » dans chacun des hémisphères de son cerveau, cette transplantation devrait donc produire deux individus ayant les mêmes souvenirs. En admettant l'hypothèse lockéenne de continuité psychologique, nous devrions reconnaître, tout de suite après la transplantation, une même personne « éclatée » dans deux corps différents. Bien sûr, cette identité ne serait que provisoire, et les différences commenceraient très vite à apparaître dans les carrière respectives. Nous voici en présence d'une situation de fission proche de celle du récit de Gogol, mais favorisant sans doute des conditions plus vraisemblables. Appelons William et Wilson les deux personnes transplantées. La caractérisation de Locke implique de considérer que, puisque la continuité des souvenirs est une condition suffisante de l'identité personnelle, les deux individus transplantés seraient identiques à l'individu original. Dans ce cas, par transitivité de l'identité, il faudrait également admettre l'identité entre William et Wilson, ce qui paraît étrange, mais peut éventuellement être admis si l'on accepte l'hypothèse d'individus-amibes non-connexes. Mais un autre problème, irrémédiable, subsisterait : William de son côté, et Wilson de l'autre développeraient au fur et à mesure de leur carrière des souvenirs différents, ce qui, cette fois par condition de nécessité du critère de Locke, impliquerait des personnes différentes. Aussi,

ni William ni Wilson ne pourrait être identique à l'individu original.

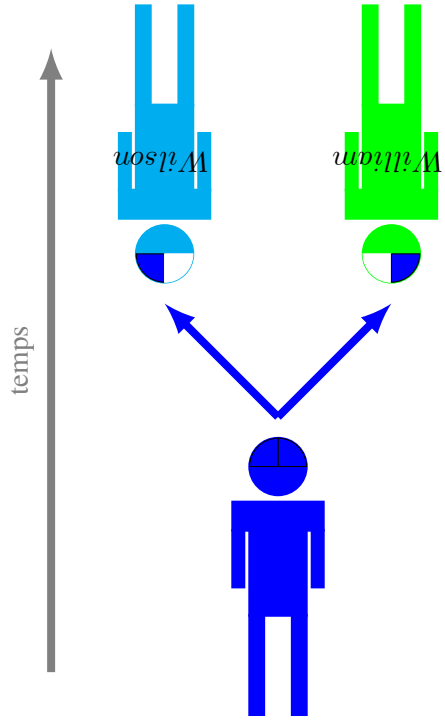


FIGURE 7.3 : Transplantation de deux hémisphères

La solution, écrit Parfit, consiste à dire que les deux personnes transplantées en formeraient encore une troisième qui, celle-ci, serait la seule identique à l'original, entité distribuée dans deux corps différents, « de même que les trois couronnes du pape n'en forment qu'une seule » (PARFIT, 2015, p. 4). Mais à dénombrer ainsi trois personnes pour deux corps, des difficultés apparaissent d'emblée. Admettons un duel entre William et Wilson, chacun se tenant d'un côté d'une pièce. De quel côté serait William-Wilson ? Et si William ou Wilson venait à mourir, combien resterait-il de personnes.

Le cas des fissions est sans doute l'écueil le plus grave de la caractérisation psychologique de l'identité. Si l'on admet la possibilité logique de cette expérience, et qu'on souhaite éviter de souscrire à l'idée d'une même personne présente simultanément dans deux corps, il semble qu'on doive alors choisir parmi les dénombrement suivants :

1. On affirme que l'existence de l'individu original persiste exclusivement à travers l'un des deux corps. Il reste alors à savoir lequel. Pour y répondre, on peut :

   (a) revenir à la théorie du meilleur candidat en discriminant de manière contextuelle, l'individu le plus apte à manifester une identification à l'original ;

   (b) ou alors renoncer à déterminer lequel est le même que l'original.

2. On admet, avec la théorie de la perdurance, qu'on a, à tout moment, deux individus différents qui partagent de manière provisoire certaines de leur parties temporelles, à savoir, toutes celles qui précèdent la transplantation.

3. On soutient qu'il y a trois personnes différentes. L'individu original ayant cessé d'exister, il n'est identique à aucun des deux individus transplantés, et ces deux individus sont des personnes différentes.

La première possibilité avec l'option du vague admet une identité sans admettre qu'il soit possible de la déterminer. Autrement dit, nous aurions une identité métaphysique épistémiquement inaccessible. Une telle identité, par nature invérifiable, serait sans doute une relation superfétatoire puisqu'elle n'aurait aucun intérêt pour nos moyens d'identifications. La seconde possibilité, défendue par David Lewis, admet des entités temporairement coïncidentes. La troisième possibilité consiste à admettre que l'individu original ne persisterait ni dans l'un ni dans l'autre des individus transplantés.

Parfit soutient que ce que nous empêche de choisir parmi des possibilités aussi invraisemblables les unes que les autres, c'est que nous traduisons indûment la situation dans le langage de l'identité. Or il faut distinguer l'identité de la *survie* :

> Peut-être la personne originelle perd-elle effectivement son identité. Mais il peut y avoir d'autres façons de le faire que de mourir. L'une de ces façons pourrait être de se multiplier. Regarder ces choses comme identiques, c'est confondre zéro et deux. (PARFIT, 2015, p. 8)

Selon Parfit, nous devrions tout simplement admettre qu'il n'y a pas de réponse à la question de l'identité personnelle. Elle consiste à abandonner le langage de l'identité. On peut imaginer que William-Wilson survive à la fois dans William et dans Wilson sans admettre pour autant leur identité. Ce qui compte au fond, écrit Parfit, ce n'est pas l'identité, c'est la survie de la personne, sa continuité, quitte à ce que celle-ci soit réalisée dans plusieurs corps à la fois. La personne originale n'est identique à aucune des deux personnes transplantés, mais elle survit dans ces deux personnes. Ainsi, écrit Parfit, « certaines questions importantes présupposent effectivement une question d'identité personnelle, mais on peut les libérer de cette présupposition. Et une fois qu'elles le sont, la question de l'identité n'a plus aucune importance » (PARFIT, 2015, p. 2). Nous devons donc admettre, conclut Parfit, un concept de survie qui autorise qu'une personne puisse survivre dans deux autres. Ce que montre l'expérience de fission, c'est que la survie n'est pas, comme l'identité, une relation un-un. Mais plus encore, Parfit soutient que la survie n'est pas une relation absolue mais relative. Le malade d'Alzeihmer n'est sans doute pas identique à la personne qu'il était avant, mais il survit à un certain degré à la personne qu'il était.

Les oublis de soi

Parfit envisage également la possibilité de quasi-intention, notion forgée de manière similaire à celle des quasi-souvenirs, mais temporellement inversée. Une personne pourrait avoir la quasi-intention d'accomplir les intentions appartenant en réalité à une autre personne. L'expérience de transplantation des cerveaux mène alors à des intentions contradictoirement réalisées : William

pourrait réaliser la quasi-intention de mener une vie ascétique alors que Wilson pourrait réaliser la quasi-intention de William-Wilson de mener une vie hédoniste. Joëlle Proust, qui reprend la notion de Parfit, affirme que les quasi-intentions sont des réalités cliniques :

> Le patient semble adhérer à une intention supposée, dont il ne peut dire si elle provient d'autrui ou de lui-même, ce qui revient à dire que les frontières entre les instances de formation des intentions sont, provisoirement, perdues. (J. PROUST, 1996)

De ce point de vue, il pourrait paraître surprenant que nous ne perdions pas, à tout bout de champs, notre identité qui semble un phénomène bien provisoire et fragile.

David Hume a repris à son compte, en l'amplifiant, la thèse de Locke :

> Nous avons une idée distincte d'un objet qui demeure invariable et ininterrompu durant une variation supposée de temps ; cette idée, nous l'appelons identité ou *sameness* (HUME, 1968b)

Le soi est une illusion parce que nulle impression sensible ne peut corroborer cette idée, parce que nos pensées sont en flux constant et que seules existent des perceptions particulières :

> Pour ma part, quand je pénètre le plus intimement dans ce que j'appelle *moi*, je bute toujours sur une perception particulière ou sur une autre, de chaud ou de froid, de lumière ou d'ombre, d'amour ou de haine, de douleur ou de plaisir. Je ne peux jamais me saisir, *moi*, en aucun moment sans une perception et je ne peux rien observer que la perception. (HUME, 1968b, p. 342-344)

Il est généralement objecté à ce point de vue que, lorsque Hume dit s'examiner au plus intime de lui-même, il y a bien là quelque chose qui s'examine. Pourtant, cette objection semble manquer sa cible, car ce que Hume affirme constater, c'est l'absence d'un *sentiment* d'identité personnelle. Il dit ne pas avoir de perception d'un *soi* comme il a des perceptions des objets ordinaires. Et cette insensibilité au *Moi* n'exclut pas la capacité à s'identifier.

Hume continue ainsi :

> Quand mes perceptions sont écartées pour un temps, comme par un sommeil tranquille, aussi longtemps, je n'ai plus conscience de moi et on peut dire vraiment que je n'existe pas. Si toutes mes perceptions étaient supprimées par la mort et que je ne puisse ni penser ni sentir, ni voir, ni aimer, ni haïr après la dissolution de mon corps, je serais entièrement annihilé et je ne conçois pas ce qu'il faudrait de plus pour faire de moi un parfait néant. (HUME, 1968b, p. 342-344)

S'identifier n'est rien d'autre que la capacité de se ré-identifier entre deux moments de conscience donnés. Mais comment, s'il y a eu interruption de conscience, nous retrouvons-nous ? Comment est-on certain de reprendre sa place ? Comment redevient-on soi, entre deux moments d'absence au monde ? Qu'est-ce que cela veut dire précisément ? Avons-nous un critère d'identité pour ne pas nous méprendre sur notre « moi véritable » ? Comme le constate aussi Vincent Descombes après Hume, la caractérisation lockéenne a des conséquences sur les moments d'inconscience :

« il n'y aurait rien d'illogique à ce que, tous les matins, au moment où je me réveille, je tienne à vérifier si je suis bien *dans le même corps* que la veille, si je suis bien *le même homme* que celui avec lequel je coïncidais hier » (DESCOMBES, 2013, p. 112).

Notons que l'inverse s'avère tout aussi problématique. On pourrait tout aussi bien s'interroger quant au fait de retrouver le « moi » de la veille. Dans un passage de *La recherche du temps perdu*, Marcel Proust s'étonne de cette possibilité, interrogeant le fait que nous nous retrouvions comme l'on retrouve le même vêtement gisant au pied du lit. Ce qui cause l'étonnement de l'écrivain, c'est l'expérience sans cesse renouvelée d'une continuité de conscience liée à un corps identique plutôt que diverses incarnations, dans le même corps, de différents soi :

> Comment alors, cherchant sa pensée, sa personnalité comme on cherche un objet perdu, finit-on par retrouver son propre « moi » plutôt que tout autre ? Pourquoi, quand on se remet à penser, n'est-ce pas alors une autre personnalité que l'antérieure qui s'incarne en nous ? On ne voit pas ce qui dicte le choix et pourquoi, entre les millions d'être humains qu'on pourrait être, c'est sur celui qu'on était la veille, qu'on remet juste la main. Qu'est ce qui nous guide quand il y a eu vraiment interruption [...] ? (M. PROUST, 1987, pp. 89-90)

La question pourrait paraître incongrue en ce qu'il semble difficile de nous tromper sur notre propre identité. Nous pouvons certes nous méprendre sur l'identité à la troisième personne, celle d'un individu que nous prendrions pour quelqu'un d'autre. Mais il semble, en revanche, qu'en dehors de cas pathologiques graves, nous soyons complètement immunisés contre toute erreur d'identité à notre propre égard. Nous savons que nous sommes le même sans avoir à y réfléchir, et le fait même de soupçonner une erreur pourrait s'avérer symptomatique d'une quelconque pathologie psychotique. Il n'en demeure pas moins que cette question possède un intérêt philosophique. Pourquoi l'auto-identification ne permet-elle pas l'erreur ? Il semble pourtant que nous ne pouvons pas nous fier complètement à notre mémoire. Nous pouvons par exemple avoir des quasi-souvenirs, forgés par des histoires qui nous ont été racontés a propos de fait remontant loin dans notre enfance. Nous pouvons croire avoir vécu ces faits parce qu'on nous l'a dit, et que nous avons construit un scénario que nous prenons pour un authentique souvenir de ces événements. Mais ils pourraient tout autant s'avérer purement imaginaires. Et inversement, nous pourrions avoir oublié de très nombreux événements auxquels nous avons participé.

L'hypothèse la plus probable est que nous nous inventons en partie pour nous identifier. Cette invention procède le plus souvent de l'imitation d'un tiers. Clément Rosset constate ainsi que :

> Ce manque à être de l'identité personnelle trouve son palliatif le plus ordinaire dans l'acquisition d'une identité d'empreint. L'imitation de l'autre permet seul à ma personnalité de se constituer. [...] l'autre dont je m'inspire doit continuer à m'influencer à tout instant : si son influence cesse, je cesse d'être moi. À moins naturellement - et c'est encore une fois le cas le plus fréquent - que son influence ne cesse au profit de celle d'un autre : auquel cas mon moi ne cesse pas d'être, mais se trouve plus ou moins considérablement modifié. Mais qu'il change ou non, mon moi ou ce que je prends pour tel ne cessera pas d'être un moi d'empreint. Incapable d'exister par moi-même, j'emprunte son identité à un autre dont j'adopte le moi, et en quelque sorte me "paye la tête", encore que dans un sens très différent, et même diamétralement opposé, de celui de l'expression habituelle. (ROSSET, 1999, pp. 41-43)

Contre une théorie substantialiste de l'identité personnelle, Alfred Ayer préconisait de considérer que le moi n'est rien d'autre qu'une construction logique élaborée à partir du matériau des perceptions : « We know that a self, if it is not to be treated as a metaphysical entity, must be held to be a logical construction out of sense-experiences. » (AYER, 1952, p. 125) Mais la notion de construction logique ne doit pas être comprise comme l'idée que nous serions des objets abstraits, par exemple des ensembles. Il ne s'agit pas non plus pour Ayer de dire que nous serions des objets fictifs, dans le sens ou nous dirions que Sherlock Holmes est un personnage fictif. Ce que veut dire Ayer par « construction logique », c'est que tout énoncé qui se prononce sur un « moi » pourrait être reformulé sans perte de sens en un énoncé ne mentionnant pas d'être pensants, mais uniquement des flux de perception.

Nous avons vu que l'identité personnelle n'était pas une donnée immuable, mais quelque chose qui peut se perdre (cas de l'amnésie) ou qui peut se créer à partir d'un autre (cas des q-souvenirs). En cela, l'identité personnelle peut être rapprochée d'une création, et en particulier, d'une création imaginaire. Le récit de soi contribue ainsi à l'émergence de l'identité personnelle.

7.2.2 Le moi-mêmeté et le moi narratif

Définir l'identité personnelle par la continuité psychologique, c'est conférer au récit de soi une importance primordiale dans le processus d'auto-identification. Se souvenir, c'est se raconter, à soi-même, ce soi passé. Le soi est ainsi dédoublé par le récit qu'il se fait de lui. Or, ramener ainsi le souvenir de soi à un récit de soi implique qu'il est fondamental, pour que soit préservée l'identité personnelle de l'individu qui s'y trouve ainsi définie, d'en conserver l'authenticité. C'est toute la différence qui sépare le récit auto-biographique du récit d'auto-fiction. Dans l'un, les événements décrits sont présumés avoir eu lieu, de manière objective, dans le monde commun. Dans l'autre, il ne s'agit au fond que de réarranger la réalité en constituant un monde similaire mais fictionnel, où l'individu actuel s'y trouve dédoublé en une contrepartie.

Si l'identité personnelle est essentiellement narrative, elle est alors sans doute aussi extrêmement fragile. La nécessité de préserver le fil du récit pour persévérer son rapport à soi est illustré par les histoires que Shéhérazade, dans *Les Milles-et-une Nuits*, se propose de conter, de nuit en nuit, afin d'assurer sa propre survie. Rappelons que dans le conte, le Sultan, furieux de l'infidélité de l'une de ses courtisanes, a pris la décision de ne plus risquer aucun adultère, en faisant occire, au lever du soleil, chacune de ses infortunées épouses de la veille. Shéhérazade, qu'il vient juste d'épouser, parvient néanmoins à proroger de jour en jour la sentence qui la menace, en aiguillonnant la curiosité du Sultan, impatient de connaître la suite d'une histoire qu'elle a entrepris de raconter. Telle serait la fragilité de l'existence des personnes : briser le fil narratif, c'est perdre l'identité.

Si la vie de Shéhérazade tient au fil, fébrile, de son récit, c'est que l'identité narrative est tout entière dépendante de l'histoire que l'on raconte. Les personnes sont alors moins dépendantes d'animaux humains, de corps, de cerveaux, que d'informations. C'est ce qui convie sans doute Daniel Dennett à observer que la seule théorie qui pourrait prétendre à une perdurance de l'individu par delà la destruction matérielle serait une théorie de l'individu comme information abstraite :

> If you think of yourself as a center of narative gravity, on the other hand, your existence

> depends on the persistence of that narative (rather like the Thousand and One Arabian Nights, but all a single tale), which could theoreticaly survive indefinitely many switches of medium, be teleported as readily (in principle) as the evening news, and stored indefinitely as sheer information. If what you are is that organization of information that has structured your body's control system (or, to put it in its more usual provocative form, if what you are is the program that runs on your brain's computer), then you could in principle survive the death of your body as intact as a program can survive the destruction of the computer on which it was created and first run. Some thinkers (e.g., Penrose, 1989) find this an appalling and deeply counterintuitive implication of the view I've defended here. But if it is potential immortality you hanker for, the alternatives are simply indefensible. (Daniel DENNETT, 1992, p. 439)

Les deux démarches ont pourtant des démarcations troubles. Dans le film de Ridley Scott *Blade Runner* adapté du roman de Philip K. Dick déjà cité dans ce mémoire, quelques scènes décrivent l'importance des artefacts scripturaux pour l'identité d'une personne. Les répliquants, androïdes dont le programme est doté d'une fausse mémoire affective, prennent chacun grand soin à préserver quelques photographies jaunies qui leur appartiennent. Ces photographies sont censées témoigner d'un passé qu'ils n'ont pourtant jamais vécu. Mais cela ne les empêchent pas de fonder leur rapport à eux-mêmes sur de telles reproductions contrefactuelles, faute d'autres moyens. Lorsque le monde actuel ne permet pas de définir l'identité, c'est la pure possibilité qui en reprend le rôle.

L'exemple est certes tiré d'une fiction, et on ne manquera pas de constater qu'il s'agit même de science-fiction. Mais il semble que c'est la réalité qui l'inspire ici. Nous disons parfois que « nous avons ramenés des souvenirs d'un séjour et que nous les avons disposés sur des étagères », comme si nous pouvions poser des états mentaux sur des planches de bois. En outre, parmi les marchandises des marchés aux puces ou des étals des brocanteurs, on trouve de vieux portraits photographiques, que certaines personnes avouent acheter, ni par intérêt esthétique ni par intérêt historique, mais motivées par l'idée qu'ils promeuvent une identification généalogique, qu'ils constituent une sorte d'identité familiale fictionnelle. En achetant un portrait, on s'achète des aïeux, autrement dit, une profondeur généalogique, une origine narrative du soi. Bien sûr, de tels souvenirs sont fantaisistes, leur acquéreurs ne sont pas dupes d'eux-mêmes. Mais ces photographies ne sont pas des souvenirs « usurpés », elles ne sont pas des souvenirs du tout. Elles sont des supports qui servirent aux souvenirs d'autres personnes. Mais, que ces identifications soient hallucinatoires plutôt que réelles, fictionnelles plutôt que factuelles, force est de constater que certains y trouvent une forme de reconnaissance de soi.

Cette invention narratologique basée sur les souvenirs d'autrui peut sans doute être rapproché de la notion d'ipséité que Paul Ricœur développe dans *Soi-même comme un autre*. La notion d'identité personnelle y est médiatisée par celle d'identité narrative, c'est-à-dire d'identité de personnage. À en croire Ricœur, une question d'identité est avant tout une question concernant une action. Demander l'identité d'un individu, d'une communauté, c'est demander « qui a fait cette action ? ». « Qui a ouvert la porte ? » ou « qui a mangé la pomme ? » demande-t-on. Le sujet, chez Ricœur, est d'emblée actanciel, il est l'auteur de faits, de gestes. La porte a été ouverte, la pomme a été mangée, c'est bien par quelqu'un, un sujet dont on peut raconter le récit. En définissant le personnage de manière actancielle, Ricœur suit les préconisations structuralismes de Vladimir Propp et d'Algirdas Greimas : « est personnage celui qui fait l'action dans le récit » (RICŒUR, 1990, p.

170) C'est tout aussi vrai de personnes factuelles que de personnages de fiction :

> Quel est le support de permanence du nom propre ? Qu'est-ce qui justifie qu'on tienne le sujet de l'action, ainsi désigné par son nom, pour le même tout au long d'une vie qui s'étire de la naissance à la mort ? La réponse ne peut être que narrative. Répondre à la question « qui ? », comme l'avait fortement dit Hannah Arendt, c'est raconter l'histoire d'une vie. L'histoire racontée dit le *qui* de l'action. *L'identité du* qui *n'est donc elle-même qu'une identité narrative.* (RICŒUR, 1985, p. 442-443)

Parallèlement à l'identité-mêmeté, que Ricœur définit par le quadruplet identité numérique, identité qualitative, continuité, et principe d'individuation, l'auteur de *Soi-même comme un autre* introduit la notion d'identité-ipséité. La question posée est alors « l'ipséité du moi implique-t-elle une forme de permanence dans le temps qui ne soit pas réductible à la détermination d'un substrat, même au sens relationnel que Kant a assigné à la catégorie de substance, bref, une forme de permanence dans le temps qui ne soit pas simplement le schème de la catégorie de substance ? [...] Une forme de permanence dans le temps qui soit une réponse à la question : "qui suis-je ?" » (RICŒUR, 1990, p. 143). Ricœur introduit ici deux modalités de la permanence de l'ipséité dans le temps : le caractère et la parole tenue.

> Mon hypothèse est que la polarité de ces deux modèles de permanence de la personne résulte de ce que la permanence du caractère exprime le recouvrement quasi-complet de l'une par l'autre de la problématique de l'*idem* et de l'*ipse*, tandis que la fidélité à soi dans le maintien de la parole donnée marque l'écart extrême entre la permanence du soi et celle du même, et donc atteste pleinement l'irréductibilité des deux problématiques l'une à l'autre. (RICŒUR, 1990, p. 143)

Selon l'auteur, cette polarité nécessite d'introduction de l'identité narrative comme élément fondamental de la construction de l'identité personnelle, « à la façon d'une médiété spécifique entre le pôle du caractère, où *idem* et *ipse* tendant à coïncider et le pôle du maintien de soi, où l'ipséité s'affranchit de la mêmeté. » (RICŒUR, 1990, p. 143).

Dans *Soi-même comme un autre*, Ricœur constate qu'il a deux manière de se désigner, qu'il appelle respectivement la référence et l'auto-désignation. L'action de se désigner soi-même comme l'on désignerait encore un autre, cette action de se prendre objectivement comme objet du discours doit être différenciée de la référence à soi faite en parlant à la première personne. Le « soi » dont on parle à la troisième personne n'est pas le « soi » qui parle à la première personne. Il semble donc ici y avoir une sorte de dédoublement, comme le commente Descombes :

> Un tel résultat est évidemment des plus déconcertants, car on s'attendrait, en principe, à ce qu'un locuteur qui parle de lui-même à la première personne fasse référence à l'individu même dont les autres parlent à la troisième personne. (DESCOMBES, 2004, p. 135)

Le dilemme entre la permanence d'un moi cartésien et la pluralité des volitions et des sensations prend fin, pour Ricœur, si l'on dissocie le moi-mêmeté (*idem*) du moi narratif (ipséité) : « La différence entre *idem* et *ipse* n'est autre que la différence entre une identité substantielle ou formelle et l'identité narrative. » (RICŒUR, 1985, p. 443) Cette dissociation permet à Ricœur de rendre à la personne la possibilité du changement :

> À la différence de l'identité abstraite du Même, l'identité narrative, constitutive de l'ipséité, peut inclure le changement, la mutabilité dans la cohésion d'une vie. Le sujet apparaît alors constitué à la fois comme lecteur et comme scripteur de sa propre vie, selon le vœu de Proust. Comme l'analyse littéraire de l'autobiographie le vérifie, l'histoire d'une vie ne cesse d'être refigurée par toutes les histoires véridiques ou fictives qu'un sujet raconte sur lui-même. Cette refiguration fait de la vie elle-même un tissu d'histoires racontées. (RICŒUR, 1985, p. 443)

Il est certainement possible que le livre, que l'œuvre soit une forme de conjuration de la destruction matérielle de l'esprit. L'observation d'une toile, la lecture d'un essai ou d'un roman permet, au moins partiellement, de ressusciter les états mentaux de son auteur, d'en invoquer le caractère et la sensibilité. On a ainsi moult fois indiqué le lien qui unit l'auteur au narrateur, le double que constitue le récit pour son auteur. Outre les reflets spéculaires, les ombres portées et les échos, le texte est encore un mode d'appréhension du soi particulièrement important. Le texte permet à son auteur de se concevoir, de s'imaginer, de se comprendre, de se saisir. Remarquons d'ailleurs que dans les récits de doubles de Gogol et de Dostoïevski, les personnages dédoublés ont tous un rapport direct à l'écriture : ils sont tous deux copistes.

Constatons aussi que dans *Le Double* de Dostoïevski, le personnage principal prend progressivement la place du narrateur, faisant ainsi progresser le récit du statut d'hétérodiégèse au statut l'homodiégèse. Le texte autorise une forme de duplication métaleptique qui constitue, de manière logique mais avec ses marges inévitables d'incohérence, l'identité personnelle de celui qui s'y prête.

Dans les pratiques littéraires et les discussions critiques qu'elles ont occasionnées, il y eut une sorte de double renversement de l'identification : alors que la théorie littéraire des années 70 récusait le caractère mimétique des fictions, en contestant fermement l'idée que les personnages sont des copies de personnes réelles, inversement, les fictions post-modernes et en particulier les autofictions élaborent leur effets de fictionnalité par l'identification des auteurs à leurs personnages. Le principe d'une imperméabilité de la fiction à l'égard du réel a été progressivement détrôné par le principe de l'alimentation du réel par la fiction.

En tant que créature de fiction, et plus généralement d'œuvre d'art, au début du XXe siècle, le personnage est encore analysé en rapport à la biographie de son créateur : comprendre une fiction, comme toute œuvre d'art, c'est comprendre les méandres de la vie de son créateur. C'est donc la biographie qui dévoile la compréhension de l'œuvre, qui en explique non seulement la genèse, mais également la signification. Le personnage est expliqué par la personne. C'est pourquoi il devient légitime de reprocher aux auteurs le comportement de leurs protagonistes. Gustave Flaubert, à travers Madame Bovary, est accusé de manquer de « morale publique » et de « religion ». Dans son réquisitoire contre l'écrivain, l'avocat impérial Ernest Pinard souligne les tendances libertines, non de l'auteur, mais du personnage, indiquant son attitude délibérément déviante, irrésolue aux convenances de la morale bourgeoise :

> Messieurs, madame Bovary a-t-elle aimé son mari ou cherché à l'aimer ? Non, et dès le commencement il y a ce qu'on peut appeler la scène de l'initiation. À partir de ce moment, un autre horizon s'étale devant elle, une vie nouvelle lui apparaît. (TRIBUNAL CORRECTIONNEL DE PARIS, p.d.)

Il s'agit donc de démontrer la perversion qui habite le personnage au delà même de ce qui est simplement écrit dans le texte de Flaubert. Car c'est bien l'immoralisme de Bovary qu'on accuse au procès de Flaubert.

Le mythe de l'introspection

Pour éclairer les conséquence de l'invention du Moi cartésien, il peut être éclairant de faire appel à une étude archéologique de l'identité personnelle. Dans *L'individu dans la cité*, Jean-Pierre Vernant reprend la tripartition élaborée par Michel Foucault dans *Le Soucis de Soi* pour caractériser la notion d'individu. Il développe trois acceptions différentes, mises en correspondance avec trois genres littéraires. La première notion caractérise l'individu *stricto sensu*. C'est l'individu tel qu'il est vu depuis la troisième personne, l'entité dont on peut suivre la trace à travers l'histoire collective. Autrement dit, c'est la définition de l'individu comme membre d'une société lui conférant une place à travers ses faits et gestes et à travers le regard que lui portent ses concitoyens. Cette première notion correspond au genre littéraire de la biographie. La seconde notion d'individu le définit comme *sujet* s'exprimant à la première personne, en son propre nom, et manifestant en cela sa singularité propre. Cette notion correspond au genre de l'autobiographie ou des Mémoires autodiégétiques. Enfin, la troisième notion d'individu le définit comme un moi, comme une personne. Cette notion confère à l'individu un sens introspectif. La personne est caractérisée par son unicité et son originalité, « un individu singulier dont la nature authentique réside tout entière dans le secret de sa vie intérieure, au cœur d'une intimité à laquelle nul, en dehors de lui ne peut avoir accès, car elle se définit comme conscience d'elle-même. » (VERNANT, 1989, pp. 215-216). L'individu est placé d'emblée au creux d'une intimité qui ne se manifeste qu'à lui seul et dont lui seul peut rendre compte. Cette dernière instance correspond au genre littéraire des confessions.

Si l'on trouve bien, en Grèce antique, des biographies et des Mémoires autodiégétiques, Vernant souligne en revanche que la société grecque antique ignore les confessions et les journaux intimes. La raison, dit Vernant, est que la notion de vie « intérieure » n'a tout simplement aucun sens pour un grec antique. Certes, il serait absurde de soupçonner que les anciens grecs n'aient pas eu conscience d'eux-mêmes. En revanche, cette expérience était organisée d'une manière très différente de la nôtre, parce que les limites du moi y étaient nettement moins démarquées. L'expérience que le grec antique fait de lui-même est entièrement « orientée vers le dehors, non vers le dedans. L'individu se cherche et se trouve dans autrui, dans ces miroirs reflétant son image que sont pour lui chaque *alter ego*, parents, enfants, amis. » (VERNANT, 1989, p. 224) L'expérience que l'individu grec fait de son moi est certes réelle, mais elle n'est pas unifiée comme l'est celle de l'individu moderne. L'individu n'a pas de territoire intérieur clos, il ne se pense pas en pénétrant sa propre conscience introspective. « De même que l'œil ne se voit pas lui-même, l'individu pour s'appréhender regarde vers l'ailleurs, au dehors. Sa conscience de soi n'est pas réflexive, repli sur soi, enfermement intérieur, face à face avec sa propre personne : elle est existentielle. L'existence est première par rapport à la conscience d'exister » (VERNANT, 1989, p. 225). Chez le grec antique, l'appréhension du moi n'est pas encore l'appréhension d'un « je » mais l'appréhension d'un « il ».

Ces réflexions d'un historien tendent à indiquer l'inanité de vouloir définir de manière substantielle l'individu. La personne, au sens où elle se définit de manière introspective, est une invention

récente, moderne, dont la philosophie cartésienne marque certainement les élans inchoatifs. Mais comme le fait remarquer Vernant lui-même, il pourrait être objecté que, dans certains textes de Platon, l'être propre de Socrate est assimilé à une âme, à une *psukhé*. Dans l'*Alcibiade*, Platon écrit par exemple que : « [...] du moment que ce n'est ni le corps, ni le composé des deux qui est l'homme, il reste, je crois, ou bien que l'homme ne soit rien, ou bien, si c'est quelque chose, que l'homme ne soit rien d'autre qu'une âme » (PLATON, 1950a, p. 242 130 c)) Vernant souligne que la notion d'âme dont Platon use ici est empruntée à celle développée par les mages grecs. Les mages figurent, à côté des héros, comme des individus à part, démarqués des citoyens ordinaires par le choix d'une vie différente et singulière, ordonnée par des régimes particuliers. Par les exercices auxquels ils se soumettent, ces hommes sont vus par leurs concitoyens comme des hommes-dieux, des *théoiándres*, qui mènent leur existence à l'immortalité non par des faits héroïques, mais par des pratiques ascétiques visant la purification pneumatique et la capacité à s'extraire de son propre corps. Soucieux de remplacer la notion traditionnelle d'ombre du mort, de double venant suppléer l'absence du mort par une forme inconsistante venant hanter l'Hadès, ces mages avaient développé l'idée que l'âme n'est pas un double de la personne qui s'y substituerait avec sa destruction matérielle, mais une émanation propre de la personne, dissociable de sa matérialité physique. Néanmoins, Vernant souligne que l'âme ainsi définie n'est encore qu'impersonnelle, « elle est *l*'âme en moi plutôt que *mon* âme ». D'abord, parce que cette âme se définit par son opposition radicale au corps et à tout ce qui s'y attache, qu'elle exclut par conséquent ce qui relève en nous des particularités individuelles, de la limitation propre à l'existence physique. Ensuite, parce que cette *psukhé* est en nous un *daimon*, un être divin, une puissance surnaturelle dont la place et la fonction, dans l'univers, dépassent notre personne singulière (VERNANT, 1989, p. 228). Dans la tradition que Platon emprunte aux mages, le nombre d'âmes dans l'univers est fixé une fois pour toute : c'est exactement le nombre d'astres. Ainsi, cette métaphysique rompt avec l'idée que chaque homme soit dédoublé par une ombre venant le représenter dans l'Hadès, en la remplaçant par une âme impersonnelle venant résider dans différents hommes. La métempsychose implique de renoncer à l'idée que la personne *est* une âme, au profit de l'idée que la personne *ait* une âme. La duplication ne ressortit plus à celle d'une substance immatérielle de substitution de la personne physique, mais d'une duplication des personnes physiques logeant la même substance immatérielle.

> Chaque homme trouve donc à sa naissance une âme qui était déjà là depuis le commencement du monde, qui ne lui est aucunement particulière et qui ira, après sa mort, s'incarner en un autre homme, ou un animal, ou une plante, si elle n'est pas parvenue, dans sa dernière vie, à se rendre assez pure pour rejoindre dans le ciel l'astre auquel elle est attachée. (VERNANT, 1989, p. 228)

La psychologie, l'histoire personnelle, est donc complètement désolidarisée de la notion d'âme. Quand les mages pratiquent des exercices de réappropriation du souvenir de vies passées, ce n'est pas pour reconstruire la trace d'un individu plus grand, disons « transcorporel », comme somme des vies particulières menées par la même âme. Il s'agit plutôt d'un exorcisme : sortir à tout prix du cycle des vies pour s'intégrer à l'anonymat du cosmos [18]. L'activité de remémoration n'est pas

[18]On comprend au passage le caractère résolument contradictoire des pseudo-philosophies dites *New Age* ou des théories du « développement personnel » qui tentent vainement d'ajuster la théorie de la métempsychose à l'idée d'une préservation, voire d'un retour à une authentique identité personnelle.

liée au soucis de reconstituer la trace d'un méta-individu, mais d'en défaire l'individuation.

On pourra concéder que la notion de psyché chez Platon prend toutefois une coloration subjective qui sera complètement réalisée ensuite dans les religions monothéistes. Mais comme l'indique Vernant, chez Platon, la personnalisation de l'âme n'est pas encore de mise car « [...] cette ouverture en direction du psychologique s'effectue à travers des pratiques mentales engagées dans la cité et orientées vers ce monde-ci » (VERNANT, 1989, p. 228). L'individu doit parfois se pencher sur ses propres pensées, sur son *Daimon*. En particulier, les stoïciens développeront des pratiques introspectives destinées à affranchir l'individu du joug des autres autant que du sien propre. Mais dans le monde antique, ces exercices de discussions intérieures, ceux dont témoignent par exemple les écrits de Marc Aurèle, ne sont pas une activité d'anachorète, mais des pratiques ayant pour fin de rejaillir sur la vie en société.

Autobiographie et portrait

L'affranchissement du Moi hors des structures sociales émerge avec le surgissement de l'anachorète, type d'individu désengagé du monde dans une quête désespérée d'un moi authentique et non réductible aux seules actions et relations sociales. Comme le dit encore Vernant « recherche de Dieu et recherche du moi sont les deux dimensions d'une même épreuve solitaire » (VERNANT, 1989, p. 231).

L'extraversion du sujet, chez les Grecs antiques, est parfois repoussée jusqu'aux dernières limites. Il n'est pas rare qu'on affranchisse l'individu des limites de sa corporalité pour l'identifier aux actes et aux œuvres dont il est responsable. À ce titre, Jacques Brunschwig remarque que :

> On a du mal à croire qu'Aristote ait pu penser, en psychologique et moraliste, que l'être actuel du producteur soit l'œuvre d'elle-même (fusse en un sens seulement) et que l'œuvre de Socrate, selon l'expression de Michel d'Éphèse ne soit "rien d'autre que Socrate lui-même en acte". Mon œuvre (mais aussi bien mon ami, mon obligé, mon enfant, mon reflet, mon ombre) peut être bien *quelque chose de moi*, ma projection, mon expression, mon objectivation, ou mon "extranéation" ; il paraît absurde et sauvage de dire qu'*elle est moi*, que je suis là où elle est, qu'elle est mon être [...] Mon rapport à moi-même n'est assimilable à aucun rapport que je puisse avoir avec quelque objet que ce soit ; tout ce qui est objet pour moi est par principe autre que moi. Je suggérerai, pour finir, qu'il y a là une sorte d'obstacle épistémologique (disons "cartésien", pour faire court) dont il faut se défaire, si l'on veut comprendre un certain nombre de pensée grecques. Il serait intéressant à plus d'un égard de relever les traces, dans la pensée grecque, d'une sorte de *cogito* paradoxal qui pourrait se formuler ainsi : je me vois (dans mon œuvre ou dans quelque autre des projections de moi-même qui ont été énumérées ci-dessus) donc je suis ; et je suis là où je me vois ; je suis cette projection de moi que je vois. (BRUNSCHWIG, 1983, p.375)

L'autobiographie, qu'elle prenne la forme morale de confessions (Augustin, Rousseau) ou la forme psychologique de journaux intimes (Henri-Frédéric Amiel, Witold Gombrowicz) est animée par la quête d'un objet impossible : trouver un moi. Si l'injonction est ancienne, comme en témoigne la fameuse inscription au fronton de l'oracle de Delphe, « connais-toi toi-même », il est

sans doute révélateur qu'il n'existe pas d'autobiographie ou de journaux intimes avant le début du Moyen Âge.

On pourra toujours soupçonner de mauvaise foi l'auteur qui se livre à l'exercice de l'autobiographie. On pourra aussi y voir une sorte d'entreprise mégalomane comme en attestent les premières lignes bien connues des *Confessions* où Rousseau affirme être parvenu à réaliser une entreprise sans précédent. Citons-là dans son intégralité pour en faire ressortir l'ampleur de son caractère névrotique :

> Je forme une entreprise qui n'eut jamais d'exemple et dont l'exécution n'aura point d'imitateur. Je veux montrer à mes semblables un homme dans toute la vérité de la nature ; et cet homme, ce sera moi. Moi seul. Je sens mon coeur, et je connais les hommes. Je ne suis fait comme aucun de ceux que j'ai vus ; j'ose croire n'être fait comme aucun de ceux qui existent. Si je ne vaux pas mieux, au moins je suis autre. Si la nature a bien ou mal fait de briser le moule dans lequel elle m'a jeté, c'est ce dont on ne peut juger qu'après m'avoir lu. Que la trompette du jugement dernier sonne quand elle voudra ; je viendrai, ce livre à la main, me présenter devant le souverain juge. Je dirai hautement : Voilà ce que j'ai fait, ce que j'ai pensé, ce que je fus. J'ai dit le bien et le mal avec la même franchise. Je n'ai rien tu de mauvais, je n'ai rien ajouté de bon ; et même s'il m'est arrivé d'employer quelque ornement indifférent, ce n'a jamais été que pour remplir un vide occasionné par mon défaut de mémoire. J'ai pu supposer vrai ce que je savais avoir pu l'être, jamais ce que je savais être faux. Je me suis montré tel que je fus ; méprisable et vil quand je l'ai été, bon, généreux, sublime, quand je l'ai été : j'ai dévoilé mon intérieur tel que tu l'as vu toi-même, être éternel. Rassemble autour de moi l'innombrable foule de mes semblables ; qu'ils écoutent mes confessions, qu'ils gémissent de mes indignités, qu'ils rougissent de mes misères. Que chacun d'eux découvre à son tour son coeur au pied de ton trône avec la même sincérité ; et puis qu'un seul te dise, s'il l'ose : Je fus meilleur que cet homme-là. (ROUSSEAU, 1973)

Mais une telle entreprise ne relève sans doute pas de seuls intérêts psychologiques et moraux. Elle ressortit aussi à un intérêt ontologique. Car le récit de soi n'est pas le récit d'un événement extérieur qui n'attendrait qu'à être couché sur le papier. La mise en portrait est concomitante à la recherche d'un soi, il s'agit d'une quête dont l'étape première est justement le constat de l'absence du sujet même de cette quête. En ce sens l'autobiographie ne saurait être rien d'autre qu'une invention d'un moi constitué par et au travers de cet objet littéraire. Georges Gusdorf souligne ainsi que :

> Par delà les domaines de la psychologie ou de l'éthique, la connaissance de soi met en cause des préoccupations d'un ordre différent. Quête de l'être dont l'intention est d'établir l'identité ontologique de l'intéressé, en vertu d'une inquiétude proprement métaphysique, déjà présente chez l'adolescent qui se lance dans l'aventure des écritures intimes. [...] Le commencement des écritures du moi correspond toujours à une crise de la personnalité ; l'identité personnelle est mise en question, elle fait question ; le sujet découvre qu'il vivait dans le malentendu. (GUSDORF, 1991, p. 23)

Il se pourrait que cette crise ne corresponde pas à une défaillance du moi, à une désolidarisation éprouvée entre un moi social et un moi intime. Cette crise serait plutôt la découverte plus ou moins

consciente de la vacuité du moi, de son inexistence. Par cette expérience de recherche dont Hume parle lorsqu'il affirme ne trouver en lui aucun sujet mais que des flux de sensations et de pensées. Le vide demanderait alors à être comblé par un moi imaginé plutôt que décrit.

Pour Gusdorf, l'autobiographie est sollicitée par une présence à soi toujours donnée d'emblée : « Les écritures du moi ont pour postulat de départ l'affirmation de l'existence du moi, enjeu de l'entreprise. » (GUSDORF, 1991, p. 26). Mais Gusdorf poursuit en constatant que le point de départ centré sur la préexistence d'un moi ne va pas entièrement de soi, puisque « ceux qui s'engagent dans cette recherche paraissent mus par le désir de démontrer à leurs propres yeux, et de justifier, leur existence, comme une assurance contre les épreuves qui peuvent entraîner tout un chacun à douter de soi » (GUSDORF, 1991, p. 26). Si l'entreprise du récit de soi émerge, c'est sans doute par la constatation d'une absence qui nécessite d'être comblée. La recherche du moi serait alors, fondamentalement, une invention du moi. Comme le souligne lui-même Gusdorf, « L'affirmation de l'identité et de la permanence du moi n'est pas le résultat d'une constatation banale ; elle résulte plutôt d'un *projet* et d'un désir, elle énonce une résolution métaphysique, démentie par des témoignages contraires. [19] (GUSDORF, 1991, p. 26).

Beaucoup répondent à l'appel du moi. Mais peu d'élus parviennent à en fonder autre chose qu'un « on » ou qu'un « nous ». C'est souvent un « on » impersonnel, communautaire, qui parle dans les récits de soi. En revanche, les traces d'un personnage sont rares. C'est ce que constate Gusdorf en critiquant les théoriciens des autobiographies de masse qui nivellent tous les récits de soi au même niveau d'intérêt :

> Les auteurs majeurs dans ce domaine ne sont pas ceux qui parlent comme tout le monde, noyés dans la masse, et porte-parole, non pas d'eux-mêmes, mais de tout un chacun. Ceux-là proposent non pas des écritures du moi, mais des écritures du « nous », du « on » ; leur première personne n'est pas la première personne du singulier, mais celle de l'universel ou du communautaire. (GUSDORF, 1991)

Par quels moyens se produit l'invention d'un moi ? Plusieurs auteurs soutiennent que le rapport à la création est fondamental à son émergence. Goethe place au centre des bienfaits la création et l'encouragement de la croissance d'un être intérieur, d'une *personnalité*. Le processus passe chez lui par la création, et plus précisément, par la création littéraire. L'autobiographie de Goethe, c'est l'ensemble de sa production littéraire, son œuvre complète. L'auteur considère en effet que l'émergence et le développement du soi dans l'individuation artistique est la « jouissance la plus haute des enfants de la terre ». C'est aussi l'invitation lancée par le mot de Nietzsche « deviens qui tu es ».

7.3 La construction psycho-sociologique du moi narratif

C'est un fait que les théories qui identifient les personnes à des substances expliquent mal le fait qu'une personne ne débute pas, et parfois ne finisse pas, avec la substance censée la gouverner. Il semble plus adéquat de dire que le Moi n'est pas un produit « tout fait », qu'il n'apparaît qu'au

[19] C'est moi qui souligne.

terme d'un certain nombre de processus, et qu'il se développe sans doute tout au long de la vie d'un organisme vivant.

Nous pouvons en effet considérer que le moi n'est pas une substance, mais relève de l'acquisition de moyens de localisation d'un corps, le sien, dans le monde, et le moyen d'y référer par des expressions appropriées. Il reste à savoir alors par quel processus nous apprenons à nous identifier, à nous centrer dans le monde. Je soutiens que c'est par une forme de duplication épistémique, par une projection dans autrui et par des processus d'imitation que nous faisons l'acquisition progressive de la réflexivité. C'est la sociologie qui nous guidera vers cette fin.

7.3.1 L'émergence du moi par le jeu de rôle chez Mead

Le rôle des doubles épistémiques dans la constitution du moi a été introduit en sociologie par George Herbert Mead. On trouve en effet dans son œuvre une théorie de l'émergence du moi (*self*) ancrée sur la capacité de dédoublement conférée par l'imagination et la projection. Mead considère que le moi est complètement inféodé à la réflexivité, qu'il est pour cette raison expérimenté comme un tout, contrairement au corps qui n'est que partiellement accessible à la connaissance. Se reconnaître, constate Mead, c'est sortir de soi pour se découvrir comme objet. Mais comment est-ce possible ?

> How can an individual get outside himself (experientially) in such a way as to become an object to himself? This is the essential psychological problem of selfhood or of self-consciousness; and its solution is to be found by referring to the process of social conduct or activity in which the given person or individual is implicated. (MEAD, 1934, p. 137)

A l'instar de Descartes, Mead part du principe rationaliste que nous avons un sentiment global de notre moi, et que nous pouvons accéder directement à sa connaissance par la pensée. En outre, pas plus pour Mead que pour Descartes, le Moi n'est réductible à l'organisme humain, pour la raison que, contrairement au moi, le corps ne peut pas faire une expérience intégrale de lui-même. Certes la main droite peut sentir la main gauche et les yeux peuvent voir les genoux. Mais dans des conditions ordinaires, le dos et les yeux restent invisibles.

> We can distinguish very definitely between the self and the body. The body can be there and can operate in a very intelligent fashion without there being a self involved in the experience. The self has the characteristic that it is an object to itself, and that characteristic distinguishes it from other objects and from the body. It is perfectly true that the eye can see the foot, but it does not see the body as a whole. We cannot see our backs; we can feel certain portions of them, if we are agile, but we cannot get an experience of our whole body. There are, of course, experiences which are somewhat vague and difficult of location, but the bodily experiences are for us organized about a self. The foot and hand belong to the self. We can see our feet, especially if we look at them from the wrong end of an opera glass, as strange things which we have difficulty in recognizing as our own. The parts of the body are quite distinguishable from the self. We can lose parts of the body without any serious invasion of the self. The mere ability to experience different parts of the body is not different from the experience of a table. The table presents a different feel from what the hand does when one hand feels another, but

> it is an experience of something with which we come definitely into contact. The body does not experience itself as a whole, in the sense in which the self in some way enters into the experience of the self. (MEAD, 1934, p. 136)

Comme pour Descartes, le moi de Mead est donc quelque chose dont on fait une expérience globale, quelque chose qui est mieux connu que le corps qui reste partiellement indéterminé.

L'originalité des travaux de Mead ne tient donc ni à sa question de départ, qui n'est rien d'autre que la vieille antienne du rapport dialectique du sujet à l'objet, ni sa conception de la nature non corporelle du moi. Son intérêt est ailleurs, dans sa méthodologie behaviouriste qui détermine les expériences par lesquelles un sujet se prend pour son propre objet par l'observation d'activités de jeunes enfants. La thèse que tire Mead de ses observations, est que l'enfant parvient à construire lui-même des pensées réflexives par l'imitation et la projection dans un moi déjà constitué, celui d'un adulte. C'est donc une forme initiale de dédoublement qui conduit progressivement à accéder à la connaissance de soi à la première personne.

Plus précisément, Mead fait émerger les connaissances à la première personne de deux types d'interactions sociales, d'une part le jeu de rôle (*play*) et d'autre part l'activité ludique (*game*) où s'engage l'imagination. Dans la première catégorie de jeux, il s'agit principalement de reproduire les comportements d'une personne en occupant son rôle (par exemple le rôle de la mère ou du père, du docteur, etc.) ou en parlant à sa poupée de la même manière que ses parents. Comme le souligne Mead, il s'agit là de la façon la plus simple « d'être un autre ».

La voix est donc le média que privilégie Mead dans l'accession à la conscience de soi, à la fois parce qu'elle véhicule le langage, mais aussi parce qu'elle permet le discours intérieur, silencieux, et enfin par les effets à distance qu'elle permet de produire par l'injonction [20] :

> it has been the vocal gesture that has pre-eminently provided the medium of social organization in human society. It belongs historically to the beginning of the act, for it arises out of the change in breathing rhythm that accompanies the preparation for sudden action, those actions to which other forms must be nicely adjusted. (MEAD, 1925, p. 271)

Mead souligne qu'il ne s'agit pas tout à fait de jeux d'imitation car la simulation engage l'appel d'un effet retour sur l'enfant qui se livre ici à un processus d'interaction :

> the child is continually exciting in himself the responses to his own social acts. In his infant dependence upon the responses of others to his own social stimuli, he is peculiarly sensitive to this relation. Having in his own nature the beginning of the parental response, he calls it out by hisown appeals. The doll is the universal type of this, but before he plays with a doll, he responds in tone of voice and in attitude as his parents respond to his own cries and chortles. This has been denominated imitation, but the psychologist now recognizes that one imitates only in so far as the so-called imitated act can be called out in the individualby his appropriate stimulation. That is, one calls or tends to call out in himself the same response that he calls out in the other. (MEAD, 1925, p. 269)

[20] Notons que c'est d'ailleurs une autre originalité de la théorie de Mead que de conférer ainsi à la voix le rôle majeur dans le processus du dédoublement menant à l'auto-identification, contrairement aux nombreuses théories spéculaires, comme celle de Lacan, qui confèrent ce rôle à la vision.

Ce que l'enfant cherche à déclencher, c'est une certaine réaction qui lui permet de comprendre la signification de sa propre activité. En outre, parmi les jeux d'enfant qui relèvent du processus de construction de soi, le jeu du compagnon imaginaire occupe une place intéressante pour notre propos. Car Mead souligne que ce compagnon, qui peut être incarné par une poupée ou par tout autre artifice, n'est autre que l'enfant lui-même, dont la personne se dédouble comme pour mieux s'élaborer en s'observant « de l'extérieur ».

Ainsi, par de telles activités ludiques comme les jeux de rôle ou le jeu du compagnon imaginaire, l'enfant développe sa capacité à se considérer lui-même comme objet, à faire l'expérience de sa propre personne, c'est-à-dire, à développer son aptitude à la réflexivité qui est au cœur du processus introspectif de conscience de soi. Mais si de tels processus consistant à prendre le rôle d'autrui et à en reproduire les réactions s'avèrent indispensables à la constitution du moi, ils ne sont pas encore suffisants. Car selon Mead, il faut encore que s'y ajoutent des activité ludiques (*games*), qui succèdent aux jeux de rôles (*plays*) avec d'autres participants. Dans ces jeux apparaissent en plus des procédures et des règles. L'enfant ne doit plus seulement s'approprier un rôle, mais participer aussi à l'organisation des rôles des autres joueurs. En particulier, dans ce type de jeu, l'enfant ne doit plus seulement imiter la manière de réagir des autres, mais aussi l'utiliser pour contrôler les différents rôles du jeu collectif et structuré. Il ne passe plus de manière arbitraire d'un rôle à un autre, mais doit maintenir son propre rôle sans usurper celui des autres. C'est par de telles pratiques que Mead conçoit que se constitue progressivement un moi, mais cette constitution passe par l'élaboration préliminaire d'un autrui :

> If he plays first base, it is as the one to whom the ball will be thrown from the field or from the catcher. Their organized reactions to him he has imbedded in his own playing of the different positions, and this organized reaction becomes what I have called the "generalized other" that accompanies and controls his conduct. And it is this generalized other in his experience which provides him with a self. (MEAD, 1925, p. 269)

C'est donc la création d'un « autre généralisé », c'est-à-dire un autre qui n'est pas encore un individu concret, mais un autre-moi (dont le statut hésite encore entre un moi autre et un autre que moi), qui produit l'instance du Moi, autrement dit, le dédoublement imaginaire est constitutif de la conscience de soi. Le caractère de cet « autre généralisé » varie selon le type de jeu ou selon la position qu'occupe l'enfant dans le jeu.

La connaissance du moi relève ainsi d'une aptitude logique ancrée sur l'emploi des pronoms réflexifs, qui permettent de considérer un objet du monde sous deux modes de présentation, comme objet et comme sujet :

> It is the characteristic of the self as an object to itself that I want to bring out. This characteristic is represented in the word "self", which is a reflexive, and indicates that which can be both subject and object. (MEAD, 1934, p. 137)

Le moi est une structure essentiellement sociale qui émerge dans une expérience sociale du language « After a self has arisen, it in a certain sense provides for itself its social experiences, and so we can conceive of an absolutely solitary self. But it is impossible to conceive of a self arising outside of social experience » (MEAD, 1934, p. 142). Ce sont les actes et les paroles d'un

homme qui fondent les interprétations qu'il se donne de lui-même. Ne voyant pas ses propres actes et n'entendant pas ses propres paroles comme le font les autres, il se doit de se les représenter pour en prendre la juste mesure, pour s'en donner une perception. Or c'est cette mise à distance de ses propres actes qui fonde, selon Mead, la réflexivité du Moi.

La conséquence de cette caractérisation socio-linguistique de l'identité personnelle est, chez un même individu, différentes personnalités peuvent émerger en rapport aux contextes d'interactions sociales :

> A multiple personality is in a certain sense normal [...]. There is usually an organization of the whole self with reference to the community to which we belong, and the situation in which we find ourselves. [...] Normally, within the sort of community as a whole to which we belong, there is a unified self, but that may be broken up. To a person who is somewhat unstable nervously and in whom there is a line of cleavage, certain activities become impossible, and that set of activities may separate and evolve another self. Two separate "me's" and "I's", two different selves, result, and that is the condition under which there is a tendency to break up the personality. (MEAD, 1934, p. 142)

Les doubles épistémiques qui favorisent l'émergence du moi sont nombreux, et peuvent devenir concurrentiels. Mead en tire la conséquence qu'il peut arriver que alors l'unification normale de la personnalité diverge en courants de personnalités multiples et antagonistes, en deux moi séparés, donnant ainsi les bases pour une analyse des conditions sociales qui favorisent l'apparition de maladies psychiatriques liées aux troubles dissociatifs de l'identité.

7.3.2 Le *doppelgänger* social

En insistant sur la notion de jeu de rôle, l'identité d'une personne est finalement définie par Mead comme celle d'un personnage. Il est d'ailleurs bien connu que le terme de « personne » provient du latin *persona* qui désignait dans la Rome antique le masque porté par les acteurs de théâtre. Aujourd'hui, le terme de personnage est employé dans différents contextes : bien sûr dans la fiction, mais également en politique, en économie, en marketing et surtout en sociologie.

Ce n'est d'ailleurs pas seulement la notion de personnage qui émerge dans la sociologie du début du XIXe siècle, mais directement celle de *dopperlgänger*, tant les deux notions tendent à être solidaires l'une de l'autre. En effet, dès 1924, l'anthropologue allemand Helmuth Plessner présente explicitement l'identité sociale sous la notion de *doppelgänger* :

> Durch dir Rolle, die er vor allem als Träger eines Ranges in Amt und Beruf spielt, hat der modenre Mensch seinen Status, stellt er etwas dar. Er ist für sich und andere etwas, er ist, "wer". Dieses fundamentale Doppelgängertum braucht der einzelne, um sich in der Leistunggsesellschaft zurechtzufinden : neben einer mehr oder weniger nebelhaften Privatexistenz macht jeder, so gut er kann, eine öffentliche Figur - oder mehrere - unter dem Diktat von Freizeitgestatung und anderen öffentlichen Betreuungen. [...] So stützt sich der funktionelle doch auf den anthropologischen Rollenbegriff, der das Verhältnis des Rollenträgers zu seiner Rolle im Auge hat und damit das Doppelgängertum des privaten und des öffentlichen Menschen als seine Voraussetzung festhält. Dieses Doppelgängertum erläutert sich, da Rolle hier als Maske

> verstanden wird, am Bilde des Schauspielers. Rolleünd SSpielßind nicht mehr als Bezeischnungen für den funktionellen Beitrag zu einem Wirkungszusammenhang verwandt, sondern als Formen des Verhaltens, welche die Gesellschaft dem Menschen zumutet. Er selbst schlüpft in diese Rolle und hat dafür zu sorgen, dass er sie gut spielt. Wer aus der Rolle fällt und zum Spielverderber wird, stört die Gesellschaft und macht sich in ihr unmöglich, in den niedersten wie in den höchsten Rängen. (PLESSNER, 2003)

Plessner soutient que le concept de personne est indéterminé (*Unbestimmtheit*), par une sorte d'imprécision sociale. Ce qui pousse les hommes à participer à un rôle social, c'est le désespoir de leur intériorité (*Verzweiflung seiner Innerlichkeit*). Cela les motive à rechercher une sorte d'invulnérabilité sociale (*Wege zur Ungreifbarkeit*) (PLESSNER, 2003, p. 92) qui remplisse le vide laissé par une identité personnelle vacante. En produisant un double social afin d'accomplir un rôle, les hommes empruntent des masques qui couvrent et ce vide intérieur. Ils n'accèdent ainsi à un soi qu'à la condition de se constituer un doppelgänger (PLESSNER, 1974, p. 33)

La vie intérieure n'est donc rendue possible que par la vie sociale. De fait, la conscience de soi ne peut émerger du seul rapport de soi à soi. Comme le constate Luigi Pirandello dans *Un, personne et cent mille*,

> [...] devant votre miroir, dans l'acte de vous mirer, vous cessez de vivre [...]. Parce qu'il vous faut suspendre un instant votre vie, pour vous voir vivre. Comme devant un appareil photographique. Vous vous figez dans une attitude ; et se figer, c'est devenir statue pour un instant. La vie est en perpétuel mouvement, et ne peut jamais s'arrêter pour se contempler [...]. Quand on vit, on est occupé à vivre et l'on ne se voit pas. Se connaître, c'est mourir. (PIRANDELLO, 1982)

Si ce connaître c'est mourir, c'est qu'il ne faut pas s'attendre à trouver un moi immuable au profond de son intériorité. Toute attitude solipsiste d'observation de soi ne mène, comme le dit aussi en substance Hume, à un vide, à une absence de soi. Le soi n'est pas connaissable « par soi », il nécessite un double. Or Pirandello rappelle que le reflet spéculaire et la photographie ne sont pas encore des doubles épistémiques suffisants pour livrer la connaissance de soi, car observer un reflet ou une photographie de soi, c'est arrêter la vie, c'est constituer le moi dans une rigidité cadavérique. Rappelons encore une fois que ce que montre Zazzo par ses expériences du miroir, c'est que c'est moins la reconnaissance d'une forme figée qui témoigne chez le jeune enfant de la reconnaissance de soi, que l'appréhension d'une synchronicité motrice : se trouver dans le monde, ce n'est possible que dans et par le mouvement, par la relation aux positions et aux mouvements relatifs.

7.3.3 Pathologies du moi dédoublé

Si la sociologie permet de reconnaître l'importance du doppelgänger dans le processus de construction de l'identité narrative, elle indique aussi que cette construction n'est jamais à l'abri de dissensions internes pouvant affecter de manière plus ou moins importante la capacité d'auto-identification. Dans les cas les plus graves, on court alors le risque de ne plus parvenir à se centrer dans le monde.

Le thème littéraire de l'individu socialement fragmenté est récurrent au XIXe siècle, le *Peter Schlemihl* d'Adelbert von Chamisso (CHAMISSO, 1973), *Le double* de Fiodor Dostoïevski (DOSTOÏEVSKI, 1969) et *Le nez* de Nicholas Gogol (GOGOL, 1966) en sont des manifestations notoires. La scission intervient en particulier quand le sujet ne se retrouve pas dans l'objet que les autres prétendent voir en lui. C'est encore une fois Pirandello qui parle ici :

> Donc, les autres voyaient en moi un être qui m'était inconnu, qu'eux seuls pouvaient connaître en me regardant du dehors, avec des yeux qui n'étaient pas les miens ; ils me prêtaient un aspect destiné à me demeurer toujours étranger, bien qu'étant celui qui je revêtais pour eux (par conséquent, un « moi » qui m'échappait complètement) ; ils m'attribuaient une vie qui me demeurait impénétrable. Cette idée ne me laissa plus de répit. Comment tolérer en moi la présence de cet étranger ? Cet étranger que j'étais moi-même pour moi ? Comment renoncer à le voir ? À le connaître, Comment rester condamné à le porter avec moi, visible pour les autres, et cependant invisible pour moi ? Souvent, mes yeux avaient croisé, par hasard, dans une glace, ceux d'un autre qui m'y regardait aussi. J'étais vu sans me voir, tout comme l'autre qui ne voyait pas son visage mais apercevait le mien et surprenait mon regard posé sur lui. En cherchant à me mirer dans cette glace, j'aurais peut-être continué d'être vu par l'autre, mais toutefois sans la possibilité de le voir. On ne peut simultanément fixer des yeux sa propre image et celle d'un autre qui nous regarde dans ce même miroir. (PIRANDELLO, 1982, pp. 27-28)

Certaines études sociologiques ramènent ainsi le thème fictionnel du double à une problématique identitaire, à une scission de l'identité personnelle due à la pression exercée par la société sur l'individu, soit que cette pression soit trop contraignante (MIYOSHI, 1969), soit qu'elle occasionne le refoulement des pulsions sexuelles (ROGERS, 1970), soit qu'elle cause un clivage du sentiment d'appartenance nationale (MARTINIÈRE, 2008) .

Du concept d'identité personnelle au concept d'identitaire, on passe d'une relation absolue (c'est ou le même individu, ou un autre) à une relation plus vague (l'individu est plus ou moins reconnu comme [...]). Plus précisément, on passe d'un problème d'identification numérique à un problème de reconnaissance sociale. Le double y tient la place d'une scission psychologique, d'une fission dans le sentiment d'avoir non une, mais plusieurs identités, voire de subir différentes identités contradictoires [21].

En mêlant identité personnelle et sentiment identitaire, William James introduit la notion de caractère pour définir le sentiment d'identité, et il avance l'idée d'une possibilité d'erreur dans l'identification du soi. Car il y aurait, dit James, un « vrai » soi, un soi qui serait plus authentique que d'autres. Dans un lettre à sa femme, James écrit par exemple que :

> I have often thought that the best way to define a man's character would be to seek out the particular mental or moral attitude in which, when it came upon him, he felt himself most

[21] Masao Miyoshi (MIYOSHI, 1969) montre ainsi que la littérature victorienne du double repose sur le conflit dramatique entre les aspirations et les passions des sujets et la rigidité puritaine d'une morale y contredisant. On pourra aussi conférer aux approches complémentaires de (MILLER, 1985) et (HERDMAN, 1990). De fait, la perception du monde n'est pas dissociable d'un contenu politico-idéologique. En particuliers, les mondes fictionnels reproduisent, dans une large mesure, les conditions sociales du monde réel qu'ils imitent.

> deeply and intensively active and alive. At such moments, there is a voice inside which speaks and says, 'This is the real me.' (W. James, 1920, p. 199)

La mise en soupçon du caractère absolu de l'identité personnelle a plus particulièrement été développée par le psychologue américain Erik Erikson. Des soldats revenus du front le consultaient en affirmant que leur vie avait perdu toute cohésion, quoiqu'ils parussent, sinon, médicalement indemnes. Cette perturbation psychologique, Erikson la désigna sous le nom de « crise d'identité », et appela identité subjective ce qui était encore appelé « caractère » ou « personnalité » avant lui [22]. Comme le souligne Descombes, dans le cas de l'identité subjective, la question n'est plus simplement « qui es-tu ? », dont la réponse pourra être donnée par la présentation d'une carte d'identité, mais la question « qui es-tu quand tu es toi ? », question qui manifeste l'idée qu'il y a des moments où nous ne sommes pas nous-mêmes. Il y a ici une exigence de saisir une identité précise alors même que l'individu n'est pas définissable. En outre, l'identité relèverait d'un élément décisionnel, accidentel, et non plus essentiel : je peux décider d'être tel ou tel. Mais cela ne suffit par pour définir l'identité subjective, car une décision d'identité n'est pas en soi suffisante, comme le montre l'exemple de Don Quichotte qui échoue à être considéré, selon son choix, comme un justicier errant. Cette identité doit aussi être reconnue.

Celui qui est victime d'une crise d'identité, constate Erikson, sait très bien qui il est. Mais il souffre de n'être pas reconnu en tant que tel. Ce genre de crise résulte souvent de migrations contraintes de populations. L'individu déplacé est en effet placé entre deux systèmes normatifs. Celui qu'il quitte, et celui dans lequel il n'est pas encore entré. C'est un cas similaire qui se présente avec la crise d'adolescence, qui démarque la sortie de l'enfance et l'entrée dans l'âge adulte :

> Pour Erikson, le problème de l'adolescent est de développer *une* et *une seule* identité, condition de sa bonne insertion dans un groupe, lequel lui procure le soutien de son éthos, de ses idéaux collectifs. La fragmentation de son « identité » serait pathologique, elle serait à bien des égards l'équivalent de ces dédoublements de la personnalité que décrivait la psychologie du XIXe siècle. (DESCOMBES, 2013, p. 38)

Nous avons là une problématique qui est très proche de celle que nous avons développé à propos des personnages de fiction : nous pouvons admettre des identifications diverses, mais cela n'implique pas n'importe quelle identification, celles-ci étant toujours réglées par des consensus communautaires.

Comme le constate encore Descombes, il est généralement admis que la solution à la crise d'identité est la duplicité, l'idée que l'identité est multiple. Certes, l'individu ne saurait être réduit à un rôle, à une fonction. Mais c'est encore autre chose de parler d'identité plurielle :

> Me demander d'avoir plusieurs identités, c'est me demander, en réalité, de me multiplier, c'est-à-dire d'être plusieurs fois identifiable. Et me demander de me multiplier, c'est me demander d'exister à plusieurs exemplaires. Pour me multiplier, pour donner satisfaction à toutes

[22] Je m'inspire ici des analyses de Vincent Descombes qui a retracé, dans *Les embarras de l'identité* et dans *Le parler de soi*, l'émergence et le développement de l'identité psychologique sous l'aspect de l'identitaire. Voir aussi la conférence du Forum philosophique 2019 « Le Monde » de l'université du Mans à l'adresse suivante : https://www.youtube.com/watch?v=JY4am8oIKEc.

> les demandes que je reçois quand elles sont hétérogènes les unes aux autres, et bien, il faut que j'existe plusieurs fois. Et bien je vous demande, si vous deviez vous multiplier, comment feriez-vous ? Comment pourriez-vous exister à plusieurs exemplaires ? Essayez d'exister à plusieurs exemplaires, et si vous pouvez, et bien, vous m'aurez réfuté. En réalité, l'idée d'une identité plurielle présente comme la solution de la crise d'identité ce qui est la source de la crise d'identité, c'est-à-dire, précisément, cette exigence impossible adressée à l'individu. Le problème que se pose l'individu qui passe par une crise d'identité c'est de savoir comment il va pouvoir exister comme individu, donc à un seul exemplaire. (DESCOMBES, 2019, 21 :56)

Le double est constitué par la fission du Moi, qui prend la forme de l'*alter-ego*. Comme dans le mythe de l'androgyne raconté par Platon dans le *Banquet*, l'alter-ego présente une réification d'une part contradictoire de la personnalité. L'individu fissionné et son double sont alors vus comme deux figures antithétiques, des couples d'oppositions binaires.

Erikson n'admet finalement qu'une seule identité authentique, qui doit se fixer parmi différents choix possible, mais qui coexiste difficilement avec d'autres identités sociales. La pathologie du double naît quand cette fixation ne se fait pas, et que d'autres identités concurrentielles émergent. La conception d'Ervin Goffman [23] privilégie au contraire différentes identités coexistantes, ou plus exactement, l'idée d'une identité qui serait toujours multiple, en raison des changements successifs de rôles que doit occuper un même individu. La pathologie du double naît alors de la fixation d'une unique identité sociale au détriment des autres. Les membres d'une société sont considérés comme des acteurs sociaux jouant un rôle à l'intérieur d'une scène sociale. Les différents éléments constitutifs du théâtre, rôle, scène, acteur, public, coulisse, mais aussi divers éléments ludiques ou cinématographiques (par exemple la notion de cadre) sont repris par Goffman comme des arguments sociologiques pour asseoir une position interactionniste [24].

La pleine reconnaissance des conséquences de la conception sociale de l'identité a toutefois été progressive dans sa désolidarisation avec l'idée de substance du moi. Et les sociologues sont en effet parfois passés d'une identité réduite à une multiplicité de rôles à une identité multiple [25] S'il convient de reconnaître les avancées de Mead, elles sont parfois réinterprétées par une forme de retour aux identités substantielles.

L'introduction du concept d'identité en sociologie est tardif, il est dû à Erik Erikson dans son essai *Adolescence et crise. La quête de l'identité* dans les années 70. Mais notons que l'apparition de ce concept est précédé de l'importation de la figure du double, par les avatars dans les jeux de simulation dans la théorie de Mead, dans les années 20, puis par la figure du Doppelgänger chez Helmuth Plessner dans les 60. La construction sociologique de la notion de personne est donc elle-même tirée de la figure du double, elle est d'abord bâtie à partir d'elle. La sociologie de l'identitaire a certes raison d'insister sur l'importance que jouent les rôles sociaux pour l'individu.

[23] Notons que c'est à partir de la publication de l'essai *Stigma* d'Erving Goffman que le mot « self » est progressivement substitué au mot « identity » dans les études sociologiques.

[24] D'une certaine façon, c'est déjà anticipé dans l'existentialisme sartrien quand l'auteur de *L'Être et le Néant* (SARTRE, 1976) évoque le serveur jouant au garçon de café. Mais à la différence de Goffman, Sartre critique ce jeu comme relevant de la mauvaise foi : jouer à être garçon de café, c'est éluder la liberté d'être autre chose, c'est fuir l'angoisse existentielle devant cette liberté des possibles.

[25] On notera le parallélisme avec ce qui s'est aussi produit en psychanalyse : alors la théorie de Freud fissionne l'identité en trois composantes que sont le ça, le moi et le surmoi, Freud ne se défait pas complètement de l'idée d'âme, d'une substance singulière.

Mais certains auteurs, dont Erikson pourrait être le modèle, ont eu tendance à réifier ces rôles, leur conférant finalement un poids ontologique malvenu. L'identité personnelle n'est alors plus basée sur des doubles épistémiques, mais sur des natures, des doubles ontologiques, sur l'idée d'un *real me* en contradiction avec les rôles multiples, d'une concurrence douloureuse entre des identités octroyées, relationnelle. Il y aurait une identité qui serait plus vraie, plus authentique, plus fondamentale. [26]. Somme toute, parmi les identités multiples, il y aurait une identité qui serait plus authentique que les autres. De telles sociologies tendent à revenir à une conception substantialiste de ce que l'on est de manière « profonde », ou pour parler en métaphysicien, ce que l'on est de manière essentielle. En concurrence avec la notion d'identité narrative, la sociologie identitaire produit ainsi un concept d'identité sociale dont la métaphysique n'a parfois rien à envier à son homologue cartésien, sinon que la dualité des substances est remplacée par une multiplicité d'identités réifiées.

7.3.4 Imiter le désir d'un autre

Cette réification de l'identité personnelle en objet, Girard la retrouve également en psychiatrie, qu'il accuse d'être « incapable de poser correctement le problème du double, car elle ne peut pas mettre en question les structures sociales. Elle cherche à guérir le malade en le ramenant au "sens de l'objectivité". Mais "l'objectivité" de ce malade est, à certains égards, supérieure à celle des êtres "normaux" qui l'entourent. » (GIRARD, 1983a, p. 67)

C'est que psychiatrie et psychanalyse succombent la plupart du temps au mensonge romantique qu'on peut énoncer ainsi : « il y a un moi qui désire donc il y a quelque chose, un sujet du désir, une substance désirante ». Dans *Mensonge romantique et réalité romanesque* (GIRARD, 1983b), Girard a montré que l'illusion d'un sujet désirant autonome n'était qu'un contre-effet de l'idée, toute aussi fausse, d'une autonomie du désir :

> Le vaniteux romantique veut toujours se persuader que son désir est inscrit dans la nature des choses ou, ce qui revient au même, qu'il est l'émanation d'une subjectivité sereine, la création *ex nihilo* d'un Moi quasi divin. Désirer à partir de l'objet équivaut à désirer à partir de soi-même : ce n'est jamais, en effet, désirer à partir de l'*Autre*. (GIRARD, 1983b, p. 30)

La critique de Girard ne porte donc pas sur la menace de dissolution du Moi absolu dans le monde, mais sur le rejet romantique de l'imitation dans la constitution du moi.

Le double épistémique, à travers l'imitation d'autrui, permet l'occupation d'un rôle. En ce sens, il est liée à l'empathie, au partage d'intentions similaires entre des individus. L'identification

[26] Quelques remarques biographiques peuvent expliquer ce fait. "Erikson" est le nom de famille que le sociologie s'est choisi lors de sa naturalisation américaine lorsqu'il a quitté l'Allemagne nazie. Notons au passage que ce nom vient redoubler la sonorité de son prénom de naissance, Erik, comme une forme d'écho, et qu'elle semble témoigner de l'idée que l'individu est construit par lui-même puisque le nom signifie « Erik fils d'Erik ». Cette remarque pourrait bien sûr confirmer qu'Erikson défend bien une conception entièrement sociale de l'identité. Mais il faut sans doute nuancer cette affirmation. Erikson s'est certes construit une nouvelle identité sociale, mais d'un autre coté, il l'a fait de manière contrainte, en raison de la situation politique de son pays d'origine. Cela transparaît encore dans sa théorie de l'identité qui manifeste une forme de nostalgie d'une identité d'une nature différente, qui serait plus authentique que les autres. C'est certainement le sens que confère Erikson à ce qu'il a étudié sous le nom de crise d'adolescence.

à autrui passe par la reconnaissance d'intentions partagées, en particulier, de désirs communs. Girard constate la caractérisation positive qui est généralement attribuée à la similarité des penchants dans les phénomènes d'empathie : « nous avons les mêmes goûts, nous aimons les mêmes choses, nous sommes faits pour nous entendre » (GIRARD, 1993, p. 218).

Mais si la ressemblance des intentions semble fournir une explication à l'harmonie sociale, Girard soutient qu'elle débouche toujours sur une violence déchaînée qui ne trouve son apaisement que dans le rituel religieux du sacrifice via la victime suppliciée, le *bouc émissaire*. La violence sacrificielle n'apparaît que chez les hommes, les animaux des autres espèces ayant tendance à rompre la tension conflictuelle avant la destruction de l'adversaire. Il s'agit donc d'une composante importante de notre espèce. Contre l'explication de Freud, qui introduit la notion d'instinct de mort pour expliquer cette caractéristique, Gérard soutient que le désir partagé débouche de manière systématique sur la violence, et donc qu'il n'est pas nécessaire de faire appel, en plus, à une pulsion de destruction. La violence ne serait donc pas la résultante d'une pulsion propre, mais les conséquences même du partage d'attitudes intentionnelles.

L'analyse de Girard part de la constatation que la violence, à l'intérieur d'une société, est intimement liée au phénomène du désir. Analysant la notion de sacrifice, Girard fait appel aux travaux ethnologiques de Godfrey Lienhardt (LIENHARDT, 1961) et Victor Turner (TURNER & INSTITUTE, 1968). Pour ces deux anthropologues, le sacrifice rituel n'est rien d'autre qu'une opération de transfert collectif faite aux dépends de la victime sacrificielle, et dont l'objet est la résorption des tensions internes menaçant la stabilité politique d'une communauté. Dans toute société naissent des conflits d'intérêts qui, s'aggravant, débouchent sur une violence déchaînée qui n'est apaisée que par le sacrifice rituel d'un tiers, le bouc émissaire. Partant des ces données ethnologiques, Girard procède à une herméneutique de la violence sacrificielle consacrée aussi bien aux textes bibliques et mythologiques qu'à la littérature réaliste du XIXe siècle.

Contrairement à la perspective classique qui fait découler la violence sociale de la convoitise d'un objet pour ce qu'il représente par lui-même, Girard considère qu'un objet ne devient désirable que par la médiation d'un autre sujet désirant. L'objet désiré ne l'est jamais pour ce qu'il représente en tant que tel, ni même pour ce qu'il peut représenter pour soi, mais pour ce qu'il représente pour « l'*Autre* ». De sorte qu'aucun objet n'est, de lui-même, désirable. Pour qu'il le devienne, il faut l'intervention d'un autre sujet désirant. C'est donc parce qu'un objet est désiré par un autre qu'il devient désirable pour soi, et non le contraire. Comme l'écrit Girard dans *La violence et le Sacré*,

> la rivalité n'est pas le fruit d'une convergence accidentelle des deux désirs sur un même objet. *Le sujet désire l'objet parce que le rival lui-même le désire.* [27] En désirant tel ou tel objet, le rival le désigne au sujet comme désirable. Le rival est le modèle du sujet, non pas tant sur le plan superficiel des façons d'être, des idées, etc., que sur le plan plus essentiel du désir. (GIRARD, 1993, p. 216-217)

Le désir est un processus d'imitation entre rivaux qui implique une part grandissante de tensions qui s'accroissent jusqu'au déchaînement de la violence.

La rivalité mimétique introduit un paradoxe, lié à la double injonction que lancent les rivaux : « imite-moi » mais « ne m'imite pas », impératif contradictoire que Girard relie à la théorie de l'in-

[27] C'est Girard lui-même qui souligne.

jonction paradoxale qualifiée de *double bind* par les anti-psychiatres de l'école de Palo Alto. Alors que chez Gregory Bateson (BATESON, 1972), l'injonction paradoxale de la double contrainte est responsable de maladies plutôt marginales comme la schizophrénie [28], Girard en fait *a contrario* la cause de tous les processus de violence intra-communautaires, en élargissant son rôle sous la dénomination de « *cyclothymie* ».

Notons que les rôles d'imitateur et d'imité ne restent pas fixes. En effet, parvenus à ce stade du processus mimétique, écrit Girard, les rivaux sont amenés à permuter successivement leurs positions sans constater la ressemblance qui les relie finalement. La cyclothymie exprime l'ambivalence qui dynamise le conflit à travers un cycle de rôles alternés, dans lequel les substitutions se font à un rythme de plus en plus rapide.

L'accélération des alternances implique un aveuglement des différences. C'est alors que se produit le phénomène hallucinatoire du *double*, dont Girard affirme que Dostoïevski est l'un des premiers à avoir perçu la réalité. Dans *La Violence et le Sacré*, René Girard introduit la notion de *doubles monstrueux* pour compléter sa théorie du désir mimétique. Le double est toujours monstrueux, et le monstre, toujours double : « Il n'y a pas de monstre qui ne tende à se dédoubler, il n'y a pas de double qui ne recèle une monstruosité secrète » (GIRARD, 1993, p. 237).

> Dans l'expérience collective du double monstrueux, les différences ne sont pas abolies mais brouillées et mélangées. Les doubles sont tous interchangeables sans que leur identité soit formellement reconnue. Ils fournissent donc, entre la différence et l'identité, le moyen terme équivoque indispensable à la substitution sacrificielle, à la polarisation de la violence sur une victime unique qui représente toutes les autres. Le double monstrueux fournit aux antagonistes incapables de constater que rien ne les sépare, c'est-à-dire de se réconcilier, exactement ce dont ils ont besoin pour arriver à ce pis-aller de réconciliation qu'est l'unanimité *moins un* de l'expulsion fondatrice. C'est le double monstrueux, ce sont tous les doubles monstrueux en la personne d'un seul (...) qui font l'objet de la violence unanime. (GIRARD, 1993, p. 238)

Homme-cheval, serpent-lion, les doubles monstrueux résultent d'un processus où les différences s'effacent quand la violence s'exacerbe, et où son hystérisation implique le mélange des contraires par la confusion des genres, où n'importe quoi peut tenir le rôle de n'importe qui. Ainsi naît la créature du bouc émissaire qui institue l'identification de l'animal à l'humain, celle aussi des hybrides que sont les centaures et l'homme-cancrelat de Kafka, toutes figures dont le sacrifice sert à apaiser la communauté.

[28] Manuel Rebuschi, Maxime Amblard et Michel Musiol soutiennent que le déni de réalité du schizophrène n'est pas dû à un manque de rationalité : « la folie n'exclut pas la rationalité du sujet, même si celle-ci est déviante, et que par voie de conséquence, un point de vue à la première personne sur la pathologie est défendable » (REBUSCHI, AMBLARD & MUSIOL, 2013, p. 128). Les conversations menées avec des schizophrènes conduisent les auteurs de l'article à poser l'hypothèse que le dysfonctionnement du schizophrène ne relève pas d'une défaillance logique : « il n'y a pas plus de contradictions dans la représentation sémantique d'un schizophrène que dans celle d'un sujet normal – même si du point de vue de ce dernier, il y a des contradictions (ce qui rejoint certains traits symptomatiques habituels de la pathologie) » (REBUSCHI et al., 2013, p. 140). Selon ces auteurs, les représentations d'un schizophrène sont constituées de quasi-croyances qu'ils baptisent « schizocroyances ». Parce que le sujet non-schizophrène est incapable d'en produire, le monde du schizophrène n'est accessible pour le sujet normal ni de façon intellectuelle, ni de façon empathique. L'inconsistance du schizophrène n'est pas logique, mais pragmatique, elle relève d'usages défaillants dans la pratique de la communication que les auteurs de l'article désignent de « déficit interactionnel » (REBUSCHI et al., 2013, p. 140).

Pour soutenir sa thèse du désir mimétique, Girard traite ainsi le désir en le vidant de toute substance. Et de fait, cela encourage également à vider le moi de tout contenu. S'il n'y a pas de désir autonome, c'est qu'il n'y a pas de sujet autonome.



La thèse de l'identité narrative peut être soutenue, nous l'avons vu, par les résultats de la psychologie sociale qui reconnaît l'importance des doubles épistémiques dans la constitution de l'identité sociale et de l'identité psychologique.

En particulier, nous avons vu que le moi est construit de manière interpersonnelle. Parvenir à s'identifier, ce n'est pas se découvrir une substance du Moi préexistante, mais parvenir à centrer sa position par l'occupation de rôles dans des relations interpersonnelles.

Le processus de création du moi narratif est donc tributaire, de processus d'imitation. Mead a décrit l'apprentissage de l'enfant par l'imitation des paroles de ses parents et par l'occupation de rôles dans des jeux de simulation. Chez Girard, le rival mimétique permet à l'individu de se constituer un désir, et par là-même, de s'identifier comme sujet de son désir. De telles analyses soulignent l'importance de l'imitation dans la représentation de soi.

Il reste maintenant à analyser plus finement les mécanismes intentionnels qui génèrent la production de doubles épistémiques. Autrement dit, nous avons maintenant à éclairer la généalogie des doubles épistémiques. Je m'attacherai principalement à l'analyse de deux phénomènes : se voir comme un autre et s'imaginer comme un autre. Nous consacrerons respectivement les deux sections suivantes à leur étude.

7.4 Se voir comme un autre

Dans *Le parler de soi*, Vincent Descombes remarque qu'une célèbre scène décrite par Blaise Pascal dans les *Pensées*, où un homme le regarde passer dans la rue est l'exacte pendant d'une scène décrite par Descartes dans les *Méditations*. Quand Descartes est à la fenêtre et médite sur les manteaux et les chapeaux qui s'agitent en contre-bas, Pascal demande, quant à lui, ce qu'aurait vu un homme qui l'aurait vu passer de sa fenêtre. Entre les deux descriptions, les points de vue sont inversés :

> Qu'est-ce que le *moi* ? Un homme qui se met à la fenêtre pour voir les passants, si je passe par là, puis-je dire qu'il s'est mis là pour me voir ? Non ; car il ne pense pas à moi en particulier. (PASCAL, 1950, p. 903, frag. 323)

Ainsi, la philosophie égologique de Descartes se différencie de celle, égotiste, de Pascal, par un changement de point de vue. Chez l'un, un moi qui est le centre perceptif d'autres soi, chez l'autre, un moi qui est l'objet de la perception d'autrui. Or, cette modification de perspective a pour conséquence deux philosophies du sujet diamétralement opposées, celle de Descartes pour laquelle le

Moi se prenant comme tel est la connaissance ultime, là où Pascal le considère comme un vil objet digne de mépris (« le Moi est haïssable »).

Le changement de perspective n'est sans doute pas seulement une manière purement rhétorique et finalement anecdotique de confronter deux conceptions différentes de l'identité, mais un phénomène qui est au cœur même du processus d'identification. Dans les deux illustrations précédentes, le moi est nettement séparé des autres soi par la position qu'il occupe dans l'espace. Mais les pensées contrefactuelles autorisent des déplacements imaginaires du soi. Et comme le rappelle Bret Easton Ellis dans *White*, « [n]e pas être capable ou ne pas vouloir se mettre dans la peau de quelqu'un d'autre - afin de voir le monde d'une façon complètement différentes de la vôtre - est le premier pas en direction de l'absence d'empathie, et c'est la raison pour laquelle tant de mouvements progressistes deviennent aussi rigides et autoritaires que les institutions qu'ils combattent » (ELLIS, 2019, p. 146). Quelle légitimé donner à une intuition qui semble avoir tant d'importance pour régler notre comportement social ? De telles pensées semblent ancrées sur une hypothèse de dédoublement. Le changement de point de vue dont procède l'empathie consiste en l'occupation, en pensée, d'une position spatio-temporelle contrefactuelle. Se mettre dans la peau de quelqu'un, ce n'est pas renoncer à soi, mais occuper une autre position. « J'aurais pu être une autre personne » se traduit ainsi par l'idée d'occuper une autre position tout en demeurant identique à soi. C'est principalement de cette question que nous souhaitons discuter dans la présente section.

En outre, je souhaite montrer que l'identité personnelle n'est pas une donnée, mais une construction, et en particulier, je souhaite montrer que l'identification à la première personne n'est pas fondamentalement déterminée à partir de ressemblances qualitatives intrinsèques, mais à partir de relations extrinsèques, en particulier, sur la coïncidence spatio-temporelle. L'image spéculaire, qui est une sorte de réplique pure épistémique, est bien sûr un phénomène particulièrement important dans l'identification de soi, et offrira un point d'appui à notre analyse.

Comment répondre à la fameuse prescription « Connais-toi toi-même (*Gnothi seauton*) » qui servit de devise à l'oracle de Delphe ? Dans *L'Ombre et la différence : le double en Europe*, Wladimir Troubetzkoy souligne l'importance que revêt la vision de soi pour la connaissance de soi. Or la vision de soi nécessité un déplacement du regard, afin de passer de sujet connaissant à objet de connaissance :

> Comment se connaître soi-même ? [...] il ne sera possible de se connaître qu'en se regardant, en appliquant la procédure cognitive normale, en se posant soi-même à côté et en face de soi. (TROUBETZKOY, 1996)

L'injonction pourrait paraître paradoxale parce qu'elle demande de procéder à un changement de perspective, elle pousse celui qui souhaite s'y employer à faire comme un pas sur le coté *sans changer de place*, d'être à la fois, comme l'écrit Troubetzkoy, « à côté et en face de soi ». Poursuivant, l'auteur indique néanmoins une solution pour remédier à cette double injonction contradictoire : « L'esprit est un miroir où le moi, cette chose si haïssable et sans cesse à surveiller, se mire dès les grands moralistes du XVIIIe siècle » (TROUBETZKOY, 1996, p.) La solution, c'est bien sûr le miroir, dont le dispositif permet de mettre le sujet au devant de lui-même sans quitter sa place.

> Un moi objet est ce qui n'est pas concevable, ce qui n'est qu'imaginable sous la forme d'un simulacre vide, c'est prendre une ombre imaginaire, imaginée, pour le moi. C'est, en fait, se voir comme les autres me voient. Non pas se voir soi-même, mais voir ce que les autres voient de moi, se voir comme un autre, introduire dans son moi, à l'intérieur de son moi, une distance, une scission, un clivage, une altérité irrémédiables. C'est l'erreur par excellence que l'erreur narcissique, l'erreur du miroir, l'errance au miroir. (TROUBETZKOY, 1996, p.)

Mais peut-on ainsi prendre la place d'un autre, sans, par là-même, *devenir autre* ?

7.4.1 De l'auto-identification spéculaire à l'auto-identification imaginaire

Considérons le cas des visions de soi, ou *autoscopies*. De telles perceptions peuvent advenir à diverses occasions, celle d'un reflet spéculaire, d'une photographie ou d'un enregistrement vidéo, voire encore à l'occasion d'hallucinations.

Quand l'on porte le doigt levé à hauteur des yeux, et qu'on plonge son regard à l'infini, en direction de ce doigt, il apparaît dédoublé. Comme en témoigne également la perception en état d'ébriété, les images doubles ont mauvaise réputation du point de vue ontologique. Elles sont souvent considérées comme l'indice d'un manque de réalité. Deux doigts à la place d'un, c'est un de trop. Mais alors, l'un serait halluciné et l'autre pas ? Comment décider lequel ? Non, ils ne peuvent qu'être, *tous les deux*, hallucinés. Ce n'est donc plus le doigt qui est extérieur à mon esprit qui fait l'objet de ma vision, dirions-nous, mais plutôt deux doigts hallucinés n'existant que dans mon esprit. Il faut alors constater que, dans ma vision, deux doigts inexistants ont remplacé le doigt réel qui y occupait un emplacement médian. Mais cela semble absurde. Il nous suffirait de regarder les choses en louchant pour les rendre inexistantes. Cela pourrait s'avérer parfois extrêmement utile, mais c'est malheureusement, aussi complètement absurde. Il faut alors déduire que le doigt vu normalement n'est pas plus réel que le doigt dédoublé. Et donc que les images sont uniformément des entités mentales, et non la saisie directe de réalités extérieures. En outre, nous devrions accepter que même dans les cas de vision ordinaire, c'est bien deux images de doigt qui se présentent encore, à la différence que celles-ci demeurent superposées.

Comment contredire l'argument ? Certains affirment que nous n'avons pas affaire, dans la vision troublée, à deux doigts hallucinés, mais à la perception d'un doigt dédoublé. C'est, en substance, l'objection de R.J. Hirst (HIRST, 1959, p.49) contre l'argument des images doubles. On ne voit pas deux choses, mais on semble voir une chose dédoublée. Cet argument paraît étrange dans la mesure où la vision resterait une vision d'un doigt, mais où celui-ci aurait acquis la propriété de dédoublement, comme s'il s'agissait d'une propriété non cardinale comme être rouge ou comme, sans jeu de mot, être rond. Mais de fait, il y a bien deux visions de doigt, et non une seule. Ce n'est donc sans doute pas le contenu de l'expérience qui est double, mais l'état perceptif. Le double n'est pas ontologique (il n'y a qu'un doigt), mais bien épistémique (il y a deux visions de doigt).

Le reflet spéculaire offre une autre sorte de double épistémique, qui permet d'introduire une différence entre identification *de re* (c'est-à-dire à la troisième personne) et identification *de se* (à la première personne en tant que première personne). Considérons les deux expressions suivantes :

- **(A1)** Jean voit son image (dans un miroir, sur une photographie).

- **(A2)** Jean se voit (dans un miroir, sur une photographie).

La question qui va nous intéresser est de déterminer plus précisément ce qu'on prétend signifier avec de telles expressions réflexives.

L'expression (A1) insinue que Jean ne se voit pas lui-même, mais une image de lui. Or Jean pourrait voir son image sans s'y reconnaître, la voir à la troisième personne. Il pourrait ne pas s'éprouver comme sujet de sa vision. De fait, il s'agit là d'une histoire ancienne, celle de Narcisse, fils de Liriope et du fleuve Céphise, que conte ainsi Ovide :

> Là, le jeune homme, qu'une chasse ardente et la chaleur du jour avaient fatigué, vint se coucher sur la terre, séduit par la beauté du site et la fraîcheur de la source. Il veut apaiser sa soif, mais il sent naître en lui une soif nouvelle ; tandis qu'il boit, épris de son image, qu'il aperçoit dans l'onde, il se passionne pour une illusion sans corps ; il prend pour un corps ce qui n'est que de l'eau ; il s'extasie devant lui-même ; il demeure immobile, le visage impassible, semblable à une statue taillée dans le marbre de Paros. [...] Sans s'en douter, il se désire lui-même, il est l'amant et l'objet aimé, le but auquel s'adresse ses vœux [...]. (OVIDE, 1992, p. 120)

Avant de se reconnaître dans l'onde, Narcisse se voit, mais il n'identifie pas, à la première personne, comme l'individu qu'il voit. Sigmund Freud a rapporté deux témoignages de cette sorte d'autoscopie sans reconnaissance directe. Le premier témoignage provient d'Ernst Mach, qui mentionne, en 1900, dans son essai *Analyse der Empfindungen* l'histoire suivante :

> Lorsque j'étais tout jeune homme, je vis un jour dans la rue le profil d'un visage qui me parut au plus haut point déplaisant, et repoussant. Quelle ne fut pas ma stupeur de constater que c'était le mien ; je l'avais aperçu en passant devant un entrepôt de miroiterie, par un jeu de glaces inclinées l'une contre l'autre. Une autre fois, je montais dans un omnibus, harassé par une nuit très fatigante passée dans le train, et voilà que de l'autre côté entre également un homme : « Quel est donc ce professeur minable, pensai-je ». Ce n'était que moi, en face de qui se tenait un grand miroir. L'habitus de classe m'était donc plus familier que mon habitus spécifique. (MACH, 1991a)

Le deuxième témoignage provient d'une expérience que fit Freud lui-même :

> Je puis raconter une aventure analogue arrivée à moi-même. J'étais assis seul dans un compartiment de wagons-lits lorsque, à la suite d'un violent cahot de la marche, la porte qui menait au cabinet de toilette voisin s'ouvrit et un homme d'un certain âge, en robe de chambre et casquette de voyage, entra chez moi. Je supposai qu'il s'était trompé de direction en sortant des cabinets qui se trouvaient entre les deux compartiments et qu'il était entré dans le mien par erreur. Je me précipitai pour le renseigner, mais je m'aperçus, tout interdit, que l'intrus n'était autre que ma propre image reflétée dans la glace de la porte de communication. Et je me rappelle encore que cette apparition m'avait profondément déplu. (FREUD, 1985c, p. 257)

Hormis la différence de sentiment à l'égard du reflet, les trois récits ont ceci de commun que leur protagoniste ne le reconnaissent pas comme leur propre image. Ces situations montrent que

le contenu intentionnel *de se* de la vision de soi ne peut être réduit à un contenu intentionnel *de re*. Dans les cas les plus fréquents, ce que nous pourrions vouloir dire avec l'expression (A1), c'est que Jean voit son image comme étant la sienne propre, c'est-à-dire, qu'il la voit à la première personne. Nous pourrions reformuler l'expression sous une forme plus apte à traduire l'intentionnalité de la vision de Jean, par exemple en utilisant l'expression « Jean voit que son image est en face de lui ». Or ce qui se tient en face de Jean, c'est un morceau de verre opacifié (miroir) ou un morceau de papier (photographie).

Nous pourrions alors préférer la deuxième forme (A2), qui semble plus naturelle en ce qu'elle fait usage du pronom réflexif « se ». Remarquons d'abord que toutes les actions désignées par des expressions réflexives ne sont pas toujours des actions vécues de manière subjective. On pourrait par exemple dire, pour reprendre l'exemple d'Aristote, de tel médecin : « il s'est soigné lui-même ». Ce qu'illustre la proposition exprimée ici, c'est qu'une action transitive est rapportée de manière réflexive entre le sujet de la phrase et son complément. Mais la diathèse réfléchie employée ici ne relève en rien de la subjectivité. Certes, l'homme qui soigne effectue bien un geste sur lui-même, mais c'est de manière accidentelle qu'il l'exécute non tant que médecin, et de manière essentielle qu'il l'exécute en tant que malade. Car c'est en tant que malade, et non parce qu'il est médecin, qu'il est soigné. Cette distribution des descriptions empêche de considérer qu'il s'agisse là d'une pure attitude subjective, bien que la phrase contienne un pronom réflexif et que sa signification implique l'identité de l'agent et du patient. Mais encore une fois, il est tout à fait accidentel que le médecin et le malade coïncident. Le malade aurait pu être soigné par n'importe quel autre médecin que lui-même, de même que le médecin aurait pu soigner n'importe quel autre malade que lui-même. Ce premier type de réflexivité objective où l'agent coïncide de manière accidentelle avec le patient doit être complétée par une deuxième. Dans une conception subjective de la réflexivité, il est tout à fait essentiel que le geste réflexif porte sur soi-en-tant-que-soi, c'est-à-dire où il y a identité nécessaire entre le désignateur et le désigné. Il semble que cela soit le cas de l'expression (A2), qui s'interpréterait aussi par « Jean voit Jean », autrement dit, « Jean voit que Jean est devant lui ». Autrement dit, cette fois, c'est bien Jean qui se voit en tant que Jean. L'expression retient avec pertinence l'identification en première personne de Jean, qui se voit lui-même et non en tant qu'un autre. Mais cette fois, l'expression paraît étrange puisqu'elle tend à faire admettre, en quelque sorte, que Jean sorte de lui-même pour se voir. Mais nous devrions peut-être nous défier des termes, et, à l'instar d'Évrard de Conty, garder à l'esprit que le reflet n'est pas le modèle :

> [...] la chose oncques ou miroer ne se moustre en son lieu, ne le lieu de l'ymage n'est ou lieu de la chose. [...] et ainsy voit on aucunement la chose, maiz ce n'est pas en son droit propre lieu, et a la verité aussi, ne voit on pas a parler proprement la chose principal, maiz on en voit l'ymage et la semblance. (de CONTY, 1990, pp. 702-703)

Comme le souligne Alain Berthoz, les hallucination et les illusions relèvent toutes les deux de productions mentales de mondes possibles. Ce qui les différencie néanmoins, c'est que « les illusions, qui sont des solutions dans un monde ambigu, diffèrent des hallucinations, véritables créations de scénarios du futur à partir du passé par notre cerveau » (BERTHOZ, 2014).

Mais que verrait celui qui prétendrait se voir du dehors, mettons par exemple en l'absence de tout dispositif réfléchissant ? En quoi consisterait l'expérience inverse de la rencontre toujours

manquée avec soi-même ? Citons le témoignage de Goethe, qui, dans *Poésie et Vérité*, raconte que, chevauchant en direction du village de Drusenheim, il fut victime d'une telle hallucination autoscopique : « Je me vis non pas avec les yeux du corps mais avec les yeux de l'esprit, à cheval sur le même chemin, venant à l'encontre de moi-même, vêtu comme je ne l'avais jamais été, d'un habit gris brodé d'or. » (GOETHE, 1941). Et nous pourrions indiquer encore bien d'autres témoignages de tels dédoublements, comme les prétendues expériences de hors-corps que relatent certains patients à la suite d'une opération chirurgicale délicate, expériences dans lesquelles le sujet affirme avoir eu la sensation de flotter en dessus de son propre corps en faisant face à lui. Si l'on admet la sincérité de ceux qui affirment avoir eu, éveillés, de telles visions, nous devrions les considérer comme des sortes d'hallucinations. Nous pourrions ainsi concéder à Goethe qu'il ait effectivement eu *l'impression de voir* son propre corps à distance, et peut-être même d'une manière absolument indiscernable d'une vision authentique. Mais en revanche, nous devrions rejeter l'idée que Goethe ait *vu* son propre corps à distance, car le contenu de sa perception aurait tout simplement été vide. Goethe n'aurait tout simplement rien vu.

Au passage, il vaut certainement la peine de comparer ces hallucinations aux récits fantastiques qui relatent la rencontre avec des Doppelgänger. Les hallucinations de corps dédoublés et les récits à personnages doubles ont en commun ceci que la rencontre est bien souvent frontale, le double faisant face au sujet. Il peut s'avérer intéressant de noter précisément comment Dostoïevski décrit la rencontre de Goladkine avec son doppelgänger dans *Le double*. Dostoïevski nous présente Goliadkine marchant vers son logis. Le personnage est complètement absorbé dans ses pensées à la suite d'aventures défavorables dont il développe les circonstances comme pour les exorciser. Bien que le récit soit à la troisième personne, le narrateur décrit les pensées du personnage en suivant ses arguties de manière subjective. Notons que le procédé confère au récit un aspect étrange, paradoxal tendant à effacer les démarcations du moi de Goliadkine. Mais tout à coup, le personnage s'arrête, stupéfié, au milieu de la rue :

> Soudain il s'arrêta, comme cloué sur place, comme frappé par la foudre, puis se retourna très vite sur le passant qui venait de le croiser, - il se retourna absolument comme si quelque chose l'avait brusquement tiré en arrière ou comme une girouette bousculée par un coup de vent. (DOSTOÏEVSKI, 1969, p. 46)

Dans le récit, le fantastique s'immisce de manière subjective et comme après coup, la rencontre ayant déjà eu lieu lorsque le lecteur en est informé. Le temps de narration est ainsi légèrement décalé du temps de l'histoire. Il me semble que la stratégie de cette description tient aussi, et en particulier, du retournement à 180° de Goliadkine. En effet, il permet de susciter, chez le lecteur, à la fois l'idée d'un double en miroir par la rencontre frontale (ordre chiral, réflexion axiale) suivie d'un retournement avec des personnages répétés (ordre antichiral, translation).

Sous cette perspective, la suite du récit s'avère également intéressante. Dans la traduction française de Gustave Aucouturier, après que Goliakine a rejoint son double, différentes figures de style jouant sur l'ambiguité du caractère chiral ou non des personnages peuvent être relevées :

> M. Goliadkine avait tout à fait reconnu son son nocturne compagnon. Son nocturne compagnon n'était autre que lui-même…M. Goliadkine lui-même, un autre Goliadkine, mais tout à fait identique à lui-même…en un mot ce qui s'appelle son double sous tous les rapports. (DOSTOÏEVSKI, 1969, p. 50)

La reprise de la locution « son nocturne compagnon » en début et fin de phrase, qui figure un motif symétrique (appelé anadiplose), est suivie de diverses répétitions (plus précisément des épanalepses), comme la répétition du nom « Goliadkine » et le recourt anaphorique à la locution « lui-même ») qui viennent appuyer l'effet de redondance. Dans le même temps, le procédé rhétorique d'antilogie qui noue les termes « lui-même » et « un autre » a pour effet de rendre insistante l'absurdité d'une similitude parfaite entre les deux personnages : « il y avait là quelque chose d'anormal, d'insensé, d'incongru » (DOSTOÏEVSKI, 1969, p. 98). Ces figures de style indiquent trois caractères du double : la symétrie en miroir renvoie au caractère spéculaire, la répétition à la ressemblance parfaite, et l'oxymore à l'aspect contradictoire du double.

Notons aussi l'analogie avec ce que décrit Edgar Poe dans *William Wilson*. Lors du combat final qui oppose, face à face, le protagoniste à son doppelgänger, après avoir plongé, à plusieurs reprises, la lame de son épée dans le ventre de son adversaire, William Wilson, un moment distrait par un événement impromptu, s'éloigne de sa victime. Alors qu'il se retourne à nouveau vers son double, Wilson constate que

> Une vaste glace – dans mon trouble, cela m'apparut d'abord ainsi – se dressait là où je n'avais pas vu trace auparavant ; et, comme je marchais frappé de stupeur vers ce miroir, ma propre image, mais avec une face pâle et barbouillée de sang, s'avança à ma rencontre d'un pas faible et vacillant. (E. POE, 1932, p. 310)

Nous devrions ainsi considérer que le Doppelgänger de William Wilson en est discernable par le caractère énantiomère des deux personnages.

Dans les témoignages de visions de doubles, qu'il s'agisse d'hallucination visuelle ou de récit de fiction, il n'est généralement pas précisé si le double est ou n'est pas chiral par rapport à son modèle. L'hallucination prend-t-elle une forme spéculaire ou anti-spéculaire [29] ? Or c'est un caractère propre au double, et rarement remarqué dans la littérature critique, que le Doppelgänger est à la fois un double en reflet, un double spéculaire comme l'est une image en miroir, et à la fois un double anti-spéculaire parce qu'il est un modèle *répété*, un indiscernable.

7.4.2 Se voir voir et voir à travers

La représentation picturale n'a cessé de questionner le regard à travers ses manifestations métaleptiques. Le tableau figuratif est sans doute fait pour être vu comme une fenêtre qui donne sur des mondes fictionnels, mais il réalise cette médiation à partir d'un regardant qui est nécessaire à son activation. Aussi, les peintres ont intégré le regard du spectateur, non seulement en rapport avec l'intérêt que doit susciter le spectacle de l'œuvre, par le point de vue qu'il expose, etc., mais également en rapport avec la projection fictionnelle du spectateur à l'intérieur de l'œuvre.

Dans son livre *L'effet Sherlock Holmes, variations du regard de Manet à Hitchcock* (STOICHITA, 2015), Victor Stoichita a montré que la représentation du regard du spectateur à l'intérieur du tableau est essentiellement développée, dans la peinture occidentale, par deux procédés antagonistes : le premier est la figure-écho. Celle-ci désigne la représentation d'un personnage faisant

[29] J'appellerai dorénavant forme « anti-spéculaire » les reproductions qui, à la manière des photographies, respectent la chiralité entre le modèle et son double, par opposition aux reflets.



FIGURE 7.4 : Édouard Manet, *Un bar aux Folies Bergère*, 1880



FIGURE 7.5 : Édouard Manet, *Le Chemin de Fer*, 1874

directement face au spectateur du tableau, comme le serait son propre reflet dans un miroir. La figure-écho est donc un double spéculaire du spectateur, qui lui renvoie son regard. Le second procédé est celui de la figure-filtre, consistant à représenter un personnage vu de dos. Dans sa dernière œuvre majeure de 1880, *Un bar aux Folies Bergère* (voir figure 7.4), Édouard Manet a usé simultanément des deux procédés sur la même toile. Une serveuse regarde en direction du spectateur, et l'on peut voir également, légèrement à sa droite, et de manière contradictoire aux lois de l'optique, le reflet de son dos à travers un miroir. Le personnage est donc vu à la fois de face et à la fois de dos, présentant, d'une manière anticipant le cubisme, ses différentes « faces » de manière simultanée.

La représentation d'un personnage vu de dos interroge aussi, et surtout, la possibilité même du regard. Un autre tableau de Manet peint en 1874, *Le Chemin de Fer* (voir figure 7.4), donne à voir deux personnages. L'un, qui fait face, est une jeune femme assise qui interrompt sa lecture pour regarder le spectateur du tableau (figure-écho). À ses côtés apparaît une fillette, debout et vue de dos (figure-filtre), qui s'accroche de ses mains à une grille. Contrairement à la promesse du titre, aucun chemin de fer n'est représenté sur le tableau. En revanche, la scène dépeint la spectatrice d'un chemin de fer. Pour comprendre la scène, le spectateur est ainsi invité à prendre littéralement la place de la fillette. Celle-ci constitue ainsi une sorte d'avatar, sollicitant le spectateur à prendre sa place, à l'incarner, pour imaginer la scène qu'il regarde. La fillette devient le double du spectateur, et toute la lecture du tableau prend son sens par la projection de celui-ci dans celle-là.

7.4.3 L'illusion spéculaire

L'aspect contradictoire des figues-échos et les figures-filtres peut être rapproché de ce que René Zazzo a appelé « illusion spéculaire ». Une peinture de Magritte, *La reproduction interdite* (1937) en illustre le paradoxe. Elle représente un personnage qui fait face à un miroir. Mais la glace en reflète étrangement le dos à place du visage. Pour le spectateur comme pour le personnage, le visage de celui-ci est parfaitement invisible. Si le tableau paraît exposer une situation étonnante, constatons toutefois qu'elle ne fait en réalité que rappeler un fait très ordinaire, mais que nous oublions, à savoir que le visage d'une personne lui est toujours invisible. Si la remarque paraît triviale, après réflexion, elle semble pourtant souvent échapper à notre intuition, sans doute parce

que nous identifions de manière quasi-automatique l'image reflétée par un miroir avec celle de notre visage, et que nous connaissons celui-ci par celui-là. Comme le constate Zazzo,

> Vous n'étonnerez personne, sinon par la candeur de votre remarque, si vous faites remarquer que le dos est la partie invisible de notre corps. Mais si vous dites la même chose du visage, si vous déclarez "il nous est directement, à tout jamais invisible" vous déclencherez presque à coup sûr chez votre interlocuteur une réaction d'étonnement, une réflexion, de quelques secondes pour aboutir à un aveu de cette sorte "... évidemment, mais je n'y avais jamais pensé". (ZAZZO, 1993, p. 161)

S'il est plus probable qu'un miroir réfléchisse le visage de celui qui s'y mire plutôt que rien, il n'est pas encore évident que ce visage soit celui d'un familier. Denis Diderot mena dans la petite ville de Puiseaux une enquête de terrain en allant interroger un aveugle-né qu'il publia sous le nom de *Lettre sur les aveugles à l'usage de ceux qui voient*. Lors de sa visite, Diderot constata d'emblée la fascination de son hôte pour les miroirs. L'interrogeant sur ce qu'il entendait précisément par ce terme, celui-ci répondit qu'il s'agit d'une « machine [...] qui met les choses en relief loin d'elles-mêmes, si elles se trouvent placées convenablement par rapport à elle. » (DIDEROT, 1956, p. 84). Diderot commente ainsi les propos de son hôte :

> Il sait, sur le rapport des autres hommes, que par le moyen de la vue on connaît les objets, comme ils lui sont connus par le toucher [...]. Il sait de plus qu'on ne peut voir son propre visage, quoiqu'on puisse le toucher. La vue, doit-il conclure, est donc une espèce de toucher qui ne s'étend que sur les objets différents de notre visage et éloignés de nous. (DIDEROT, 1956, p. 85)

L'aveugle de Puiseaux souligne là encore cette réalité souvent négligée : notre propre visage nous est invisible.

L'aptitude à projeter notre moi dans l'image spéculaire est devenue, pour la plupart d'entre nous, tellement enracinée dans notre rapport à nous-même qu'elle est prise pour une connaissance *a priori*.

> L'illusion spéculaire consiste donc à croire, sans y avoir vraiment réfléchi, que la connaissance visuelle de soi est une donné immédiate de la perception, une expérience intime comme l'est la connaissance tactile et protopathique de notre corps. Alors le problème de la *construction* de l'image visuelle de soi, dans la prime enfance ne se pose pas. "Dès l'âge où la capacité visuelle existe, *dès le moment où l'enfant voit, il se voit* dans le miroir." Cette conviction, sans être formulée explicitement, se révèle dans l'expression courante "il se reconnaît" [...] alors que l'expression correcte est de dire "l'enfant s'est identifié au reflet du miroir". (ZAZZO, 1993, pp. 161-162)

Les expériences de Zazzo ont montré que, loin d'être évidente par elle-même, l'identification spéculaire est progressivement élaborée. Par l'importance qu'elles ont pour notre sujet, ces expériences méritent qu'on en rappelle le détail.

Le reflet en miroir et le soi

Dans la littérature fantastique qui lui est consacrée, l'une des caractéristiques du vampire est qu'il ne possède pas de reflet. La figure du vampire n'est d'ailleurs pas la seule, dans la littérature, à se trouver dépossédée d'image spéculaire. Dans la première version du *Horla*, Guy de Maupassant décrit ainsi le stratagème par lequel le personnage principal, hanté par une présence invisible et guettant les traces de son visiteur, tente de le surprendre :

> Je me dressai, en me tournant si vite que je faillis tomber. Eh bien !... On y voyait comme en plein jour... et je ne me vis pas dans la glace ! Elle était vide, claire, pleine de lumière. Mon image n'était pas dedans... Et j'étais en face... Je voyais le grand verre, limpide du haut en bas ! Et je regardais cela avec les yeux affolés, et je n'osais plus avancer, sentant bien qu'il se trouvait entre nous, lui, et qu'il m'échapperait encore, mais que son corps imperceptible avait absorbé mon reflet. (MAUPASSANT, 1979a, p. 828)

Au passage, notons à nouveau le mouvement de rotation du personnage qui laisse présager le caractère à la fois spéculaire et à la fois anti-spéculaire du double. Notons aussi que le personnage parle de son reflet et de lui comme d'un « nous », ce qui produit une identification particulièrement étrange, qui le dépossède d'emblée d'un moi unifié, et dont le Horla n'est finalement qu'un symptôme personnifié.

Le vampire n'est sans doute pas le seul individu qui ne se reflète pas dans le miroir. Commentant la *Dioptrique* de Descartes, Merleau-Ponty constate en effet que le sujet cartésien qui fait face au miroir ne se voit pas lui-même :

> Un cartésien ne *se* voit pas dans le miroir : il voit un mannequin, un « dehors » dont il a toutes les raisons de penser que les autres le voient pareillement, mais qui, pas plus pour lui-même que pour eux, n'est une chair. Son « image » dans le miroir est un effet de la mécanique des choses ; s'il s'y reconnaît, s'il la trouve « ressemblante », c'est sa pensée qui tisse ce lien, l'image spéculaire n'est rien *de* lui. (MERLEAU-PONTY, 1964, p. 38-39)

Le reflet n'est ressemblant que dans la mesure où la pensée de celui qui s'y mire en convient. Le reflet n'est pas de l'ordre de la ressemblance directe et naturelle, mais de la projection. La chose circulaire est reflétée en une ellipse, celle qui est rectangulaire apparaît en parallélogramme. Une image, qu'elle soit spéculaire ou picturale, n'est finalement le reflet d'une chose qu'à la condition de « ne lui pas ressembler » (DESCARTES, 1953a, IV, p. 204). Mais si la ressemblance reste stérile quant à l'effet de représentation suscitée par l'image, de quel artifice celle-ci tire-t-elle son effet de reproduction ? La réponse est que l'imagination est entièrement tributaire de la conception. L'image résulte d'une anamorphose et c'est la pensée qui constitue un rapport de ressemblances, et non la ressemblance qui commanderait la perception : Descartes inverse donc la position des simulacres qui étaient encore, chez Lucrèce, des copies fidèles des choses, des répliques pures de leurs apparences. Mais l'auteur de *L'Œil et l'Esprit* critique le désaveu cartésien du reflet spéculaire. C'est que le corps, affirme Merleau-Ponty, s'introduit toujours dans l'image, fut-elle virtuelle : « fumant la pipe devant le miroir, je sens la surface lisse et brûlante du bois non seulement là où sont mes doigts, mais aussi dans ces doigts glorieux, ces doigts seulement visibles qui sont au fond du miroir. Le fantôme du miroir traîne dehors ma chair, et du même coup tout

l'invisible de mon corps peut investir les autres corps que je vois. » (MERLEAU-PONTY, 1945, p. 33) La capacité à se projeter dans le reflet témoigne ainsi de la capacité à se projeter à la place d'un autre.

L'étude psychologique de la reconnaissance de soi par le reflet spéculaire a été développée par René Zazzo dans de nombreuses expérimentations faisant appel à l'usage de miroirs, d'ombres et de photographies. Ces expériences, dont l'objet premier fut de dater la reconnaissance de soi dans un miroir, ont fait suite aux études qu'Henri Wallon publia en 1934 dans *Les origines du caractère chez l'enfant*. Mais la première estimation de l'âge de reconnaissance de soi est due à Charles Darwin, qui avançait l'âge de 9 mois, dans un article intitulé *A biographical sketch of an infant* (1877). Les expérimentations de Zazzo ont montré que les enfants ne s'identifient que très rarement avant 18 mois. Ses diverses expérimentations sont décrites et commentées dans *Reflet de miroir et autres doubles* (ZAZZO, 1993). Elles tendent à montrer que la reconnaissance de soi dans un miroir est le fruit d'un long apprentissage dont la réussite n'apparaît guère avant l'âge de deux ans. De fait, le travail qu'engage Zazzo ressortit à deux questions : la première porte sur la connaissance visuelle de soi, et la deuxième est de déterminer ce que peut être un miroir pour le jeune enfant qui ne s'y reconnaît pas encore.

Pour commencer, notons qu'à la différence d'un proche, dont l'identification dans le miroir est favorisée par la possibilité de comparer ses traits à ceux de son image réfléchie, il n'est jamais possible de comparer notre propre visage à son reflet. L'image de soi, dans le miroir, est tout simplement orpheline d'un modèle de comparaison. La question qui se pose alors est la suivante : puisque nous ne voyons jamais notre visage, puisque celui-ci nous demeure parfaitement invisible, comment nous est-il possible de reconnaître notre reflet ?

Les expérimentations de Zazzo ont été conduites sur différents couples de jumeaux monozygotes et dizygotes. Ceux-ci furent placés dans deux types de situations, la situation « vitre » et la situation « miroir ». Dans la première, les jumeaux sont placés simultanément de part et d'autre d'une vitre transparente. Dans le type miroir, les jumeaux sont successivement placés devant un miroir réfléchissant [30]. Deux critères expérimentaux servent respectivement pour déterminer, chez l'enfant, le moment de l'élaboration de l'identification et le moment de la virtualisation de l'espace créé par le miroir. Le premier critère consiste à marquer l'enfant d'une tâche sur le nez. Si l'enfant le gratte, c'est signe qu'il s'identifie déjà à son reflet. Le deuxième critère consiste à utiliser un clignotant placé à l'arrière de l'enfant, et permettant d'attirer son attention. Si l'enfant se retourne, c'est qu'il a compris que l'espace réfléchi ne se continue pas au delà du miroir. Les expériences sont systématiquement conduites, pour tous les couples de jumeaux, à différents moments, selon l'ordre suivant :

1. Deux jumeaux sont mis en situation vitre.
2. Un jumeau est mis en situation miroir.
3. Expérience de la tache sur le nez.
4. Expérience du clignotant en arrière fond.

[30] Il faut remarquer que ces situations sont idéalisées, parce que des vitres peuvent aussi partiellement réfléchir l'environnement et partiellement laisser passer la lumière.

Les observations indiquent trois périodes successives dans l'élaboration de la reconnaissance de soi en miroir. Durant la première phase, avant 11 mois (âge de développement), dans la situation vitre, les enfants cherchent à entrer en contact avec leur frère/sœur en tapotant la vitre qui les sépare, et en envoyant des signes de contact social, comme des sourires. Dans la situation du type miroir, les enfants continuent à taper sur la vitre comme pour attirer l'attention d'un individu réel qui se tiendrait derrière un milieu transparent leur faisant obstacle, à la manière d'une vitre séparant deux parties de la réalité. Mais ils réagissent aussi par des réactions sociales plus intensives comme des baisers qui sont moins fréquents dans la situation vitre. Ainsi, si la nature des réactions semblent être la même dans les deux situations, Zazzo constate toutefois une différence d'intensité. Quand l'enfant est mis en situation miroir, il sollicite bien plus son image qu'il ne sollicite son frère dans la situation vitre. C'est certainement dû au fait que l'image de l'enfant répond systématiquement et parfaitement aux sollicitations sociales de celui-ci, entretenant chez lui une excitation continue. En revanche, dans la situation vitre, il apparaît des moments de rupture de communication entre les deux jumeaux, qui décrochent, puis se retrouvent. Soulignons aussi que l'intérêt d'utiliser des jumeaux est ici de produire différentes situation de reconnaissance. En effet, dans le cas de jumeaux monozygotes, l'enfant croit voir dans le miroir son co-jumeau, donc un individu connu, alors que pour des jumeaux dizygotes qui ne se ressemblent pas complètement, l'enfant croit voir un parfait inconnu.

La conclusion qu'il semble raisonnable de tirer est que durant cette première période, l'enfant perçoit son image comme s'il s'agissait d'un *autre* enfant lui faisant face. L'espace et l'image spéculaire sont perçus par l'enfant comme s'ils étaient réels [31].

Durant la deuxième période, autour de 15-16 mois, la réaction au miroir se différencie de la réaction à la vitre. Confronté au miroir, les enfants ne tapent plus comme pour entrer en communication avec leur reflet. En outre, ils manifestent des réactions de perplexité, d'étonnement et d'exploration. Leurs reflets semblent alors les incommoder. Zazzo observe que les enfants adoptent des comportements de fuite, de détournement du regard, de retournement. Soumis à l'expérience de la tache, les enfants parviennent maintenant à s'identifier en frottant leur propre nez. Mais durant cette même période, lorsqu'il sont soumis à l'expérience du cligotant, les enfants ne parviennent pas encore à comprendre que l'objet de distraction est situé derrière eux et non pas en avant du miroir. Le reflet paraît donc élaboré comme un double de soi, comme si l'enfant était à deux emplacements à la fois, la perception de l'image spéculaire demeure ambiguë. Ce n'est que lors de la troisième période qu'intervient la reconnaissance complète de l'espace virtuel que constitue l'image spéculaire.

Ce que tendent à montrer les expériences de Zazzo, c'est que l'appariement à un reflet n'est pas donné d'emblée à un sujet connaissant, mais élaboré progressivement par les différentes expérimentations qu'il en fait. C'est principalement le phénomène de *synchronie* qui répond, selon Zazzo, à cette question. Par opposition à l'image d'autrui, l'image spéculaire d'un individu lui est solidaire par le mouvement. Le synchronisme, par exemple celui de la main et de son reflet, aide l'enfant à comprendre la relation qui le lie à son image par la perception kinesthésique qu'il en a.

[31] Par esprit de contradiction envers les théories psychanalytiques qui définissent le stade narcissique comme moment de la constitution psychique de l'enfant, nous pourrions inversement qualifier la seconde phase observée par Zazzo de « narcissique » en ce que, comme le personnage de la mythologie, l'enfant croit voir un autre à travers son image spéculaire, et que c'est à l'endroit de cet autre que sont adressés ses baisers.

Et la découverte de soi passe plus généralement par la compréhension de l'unité dynamique qui relie le mouvement perçu au mouvement vécu subjectivement. Mais il demeure pas moins vrai, comme le reconnaît d'ailleurs Zazzo, que certaines parties du corps humain restent accessibles à la comparaison à leur reflet dans le miroir, notamment la main. Dans les expériences filmées par Zazzo [32], on constate ainsi que les jeunes enfants jouent beaucoup en observant leur main et leur reflet, en les approchant et en les éloignant progressivement du miroir. Car de fait, la main pressée sur le miroir, puis détachée de celui-ci, semble se dédoubler. Et il fait également noter que quelques parties du visage peuvent aussi être comparées à leur reflet, comme les lèvres que les enfants collent et décollent du miroir pour observer les effets de duplication qui se produisent.

Et de fait, comme le raconte Ovide, c'est bien pour cette raison que Narcisse finit par comprendre que l'onde lui renvoie sa propre image :

> Quand je te tends les bras, tu me tends les tiens de toi-même ; quand je te souris, tu me souris. Souvent même j'ai vu couler tes pleurs, quand je pleurais ; tu réponds à mes signes en inclinant la tête et, autant que j'en puis juger par le mouvement de ta jolie bouche, tu me renvoies des paroles qui n'arrivent pas à mes oreilles. (OVIDE, 1992, p. 121)

Avec la synchronie, Narcisse parvient bientôt à la certitude que ce qu'il voit dans l'onde n'est rien d'autre que lui-même :

> Mais cet enfant, c'est moi ; je l'ai compris et mon image ne me trompe plus ; je brûle d'amour pour moi-même, j'allume la flamme que je porte dans mon sein. Que faire ? Attendre d'être imploré ou implorer moi-même ? Et puis, quelle faveur implorer maintenant ? Ce que je désire est en moi ; ma richesse a causé mes privations. Oh ! que ne puis-je me séparer de mon corps ! (OVIDE, 1992, p. 121)

Le second résultat important des observations menées par Zazzo est que, contrairement à ce qu'on pourrait supposer d'emblée, le phénomène du *retournement* de l'enfant qui indique la réussite de la construction de l'espace virtuel du miroir n'est pas concomitant à la réussite de la reconnaissance de soi. Même quand l'enfant semble comprendre que l'image est la sienne, il existe une période durant laquelle l'espace du miroir semble n'être pas complètement virtualisé. L'enfant agit encore comme si l'espace projeté était en fait étendu par delà le miroir. Il contourne le miroir pour tenter de « se rejoindre ». C'est ainsi que Julien (5 ans et 1 mois), l'un des sujets des expérimentations de Zazzo, alors qu'il est encouragé à rejoindre sa mère, située derrière lui et dont il voit l'image dans le miroir, répond qu'il ne peut pas, parce que pour cela, il faudrait pouvoir traverser la glace (ZAZZO, 1993, p. 102). Mais ceci peut apparaître étrange, parce que la conséquence semble être que l'enfant s'éprouve comme étant situé au devant lui, à une place qu'il n'occupe pas. Le phénomène semble répondre ainsi à l'apostrophe lancée à un individu devenu par trop gênant : « vas voir là bas si j'y suis ! » :

> [...] la distinction entre vitre et miroir reste pendant longtemps incertaine, fragile. Tous deux, perçus par l'enfant comme des obstacles translucides, ont en commun apparemment de scin-

[32] Les expérimentations ont fait l'objet de quatre films dont le premier, datant de 1973 et intitulé *À travers le miroir – étude sur la découverte de l'image de soi chez l'enfant* peut être visionné à l'adresse suivante : https://www.canal-u.tv/video/cerimes/a_travers_le_miroir_etude_sur_la_decouverte_de_l_image_de_soi_chez_l_enfant.9037.

> der l'espace, mais l'un est réfléchissant, et l'autre pas. Réfléchissant, le miroir, donc pas translucide comme cela paraissait pourtant évident. Il faudra du temps pour s'en convaincre, le temps justement de substituer l'espace virtuel à l'illusion d'un espace réel. (ZAZZO, 1993, p. 181)

Et de fait, l'enfant, alors même qu'il s'identifie déjà à son reflet, tente de le contourner pour rejoindre un objet placé au devant du miroir.

Enfin, il convient d'insister sur un troisième résultat, celui de l'*évitement* que Zazzo renonce dans un premier temps à interpréter, même s'il le suppose présent chez tous les enfants, puis qu'il attribue aux hypothèses suivantes :

1. Le synchronisme révèle à l'enfant la connaissance visuelle de soi, or cette découverte de son visage provoque un émoi.
2. L'enfant est troublé par le face à face avec un inconnu.
3. La réaction de l'enfant répond au comportement lui-même étrange du personnage qui imite tous ses mouvements.

Or ce que montrent les expériences, c'est que la troisième hypothèse est la bonne. La deuxième est réfutée d'emblée par le fait que l'expérimentation fait usage de jumeaux monozygotes, dont l'image en miroir ne peut être associée à un inconnu. En outre, Zazzo a montré que des enfants sont indifférents au reflet d'un inconnu, mais bien gênés par la synchronicité de leur propre reflet. « Ce qui déconcerte, écrit Zazzo, ce n'est pas la présence incongrue d'un *étranger*, d'ailleurs tout un chacun en a vu d'autres, c'est le face à face avec un individu *étrange*, du fait de son comportement » (ZAZZO, 1993, p. 125). Zazzo en conclut :

> C'est le double spéculaire qui est troublant, et ce double est tout à fait autre chose qu'une imitation. Entre un geste modèle et son imitation, il y a toujours un décalage de temps, fut-il infime. De sorte que l'imité et l'imitant gardent chacun son quant à soi, une marge d'indépendance, de liberté. [...] Avec le double, la dépendance est totale. Impossible de casser la simultanéité qui fait de vous son prisonnier. Impossible, à moins de ne plus regarder, de se retourner, de fuir. (ZAZZO, 1993, p. 125)

Mentionnons encore que la gêne occasionné par le reflet se retrouve également dans le rapport que les jeunes enfants entretiennent à leur ombre, lorsqu'ils ne se l'approprient pas encore complètement.

L'illusion spéculaire est sans doute un cas particulier de l'illusion générale de transparence à soi, dont l'expression elle-même souligne déjà le caractère spéculaire. Mais comme l'indique ce qui précède, ce jeu de transparence ne peut se véritablement constituer qu'avec une relation projetant ce putatif Moi sur, justement, ce qu'il n'est pas.

Vallée de l'étrange et inquiétante étrangeté

Quel est l'origine du comportement d'évitement de l'enfant ? Quel trouble éprouve-t-il devant le miroir ? Zazzo soupçonne que le ce comportement provient d'une réaction face au comportement

lui-même étrange du reflet. De fait, il n'est pas coutume qu'un individu réponde aux solicitations d'un autre par une imitation parfaitement synchronisée. C'est donc la synchronie qui, selon Zazzo, produit le comportement d'évitement.

Je voudrais illustrer ceci par un épisode de la série télévisée *Twilight Zone*, « Mirror Image » (BRAHM, p.d.), où Millicent Barnes, une jeune femme en escapade, fait une escale dans une gare routière. Par le miroir des toilettes, elle aperçoit son propre reflet au dehors des toilettes, découvrant qu'elle a affaire à un double indiscernable. Non seulement sa duplique lui est, du point de vue physionomique, parfaitement ressemblante, mais en outre, elle agit en tout point à sa manière, se déplaçant aux mêmes endroits, effectuant les mêmes actions, à cette différence près que la copie devance systématiquement son modèle de quelques secondes. Ainsi, lorsque Barnes s'adresse à une personne, son double l'a déjà fait, prononçant les mêmes paroles, posant les mêmes questions. Barnes cherche alors à se confronter directement à son double afin de l'obliger à rendre raison. Mais le spectateur comprend qu'elle n'a malheureusement aucune chance d'y parvenir puisqu'un intervalle temporel sépare inexorablement ses actions de celles de son *doppelgänger*, ce qui rend impossible leur rencontre [33].

Nous pourrions donner un prolongement à cette situation imaginaire. Admettons par exemple que l'intervalle temporel qui sépare les agissements du double et de son modèle soit réduit à zéro. Nous aurions alors une parfaite synchronie entre le double et son modèle. Deux individus indiscernables se tiendraient simultanément au même emplacement au même moment. Nous aurions deux individus absolument coïncidents quoique numériquement distincts.

Il existe un conte turque qui suit la même idée, bien que le protagoniste renonce à rejoindre, par manque d'intérêt, son double :

> Nasr Eddin se rend au marché pour y vendre deux belles pastèques de son jardin. Il va ainsi, un fruit sous chaque bras, lorsqu'au détour du chemin, il voit, marchant devant lui, un homme vêtu exactement de la même manière que lui et portant lui aussi deux pastèques. - Par Allah ! s'écrie-t-il, si ce n'est pas moi, je me demande qui cela peut bien être !
>
> Il accélère un moment la cadence, mais quelques pas plus loin renonce définitivement à le rejoindre :
>
> - Au fait, se demande-t-il, à quoi cela servirait de le rattraper ? (ANONYME, 1998, p. 23)

Et on ne peut sans doute que donner raison à Nasr Eddin : quel serait, en effet, l'intérêt de gaspiller son temps et son énergie pour coïncider avec soi-même ?

Dans *Alentour et au-delà du miroir*, Umberto Eco s'interroge : « Celui que je vois dans le miroir c'est vraiment moi. Mais suis-je convaincu que c'est moi parce que je sais que cet objet est un miroir à quoi je fais confiance ; ou bien est-ce que je comprends que c'est un miroir parce je rencontre à sa surface quelque chose qui, sans aucun doute, est moi ? [34] » Selon Eco, notre fascination pour les miroirs viendrait de cette incertitude, de cette « question insignifiante ou paranoïaque ». Mais Zazzo constate que le trouble ne provient sans doute pas de l'absence de réponse

[33] De fait, il est pourtant théoriquement possible pour Barnes de rejoindre son double, puisque la condition de séparation stipulée par le scénario est une exclusion temporelle et non spatiale. En revanche, une telle rencontre ne pourrait pas être due entièrement à l'initiative de Barnes. En effet, si elle tentait de rencontrer son double, celui-ci l'aurait inévitablement déjà précédée, et la rencontre aurait déjà eu lieu avant l'initiative de Barnes.

[34] Passage traduit par René Zazzo dans (ZAZZO, 1993).

à la question de savoir qui est premier entre le fait de reconnaître un miroir ou le fait de reconnaître un reflet. Le trouble paraît plus profond.

Zazzo avoue que les réactions de l'enfant face aux doubles perceptifs, qu'il s'agisse de photographies, d'ombres ou de reflets spéculaires l'intéressent aussi « parce qu'elles permettent de dater les strates archaïques de notre propre pensée. Il est évident que les doubles mentaux que l'enfant se fabrique perdurent en nous » (ZAZZO, 1993, p. 143). De fait, il semble que la réaction de détournement soit plus qu'un étonnement face au reflet, mais d'une forme d'appréhension du double qu'il me semble intéressant de rapprocher à la fois du motif freudien de l'*unheimlich* et de la conjecture que le roboticien Masahiro Mori a qualifié en 1970 de « vallée de l'étrange » (*uncanny valley*). Mori affirme que le sentiment d'empathie à l'égard de robots androïdes dépend de leur ressemblance aux humains. Pour Mori, cela se traduit mathématiquement par une fonction continue exprimant le degré d'empathie en fonction de la ressemblance d'un artefact à un être vivant. La conclusion de Mori est que la courbe est croissante sur un premier intervalle, avant de devenir brutalement décroissante, atteignant un degré minimal de familiarité qu'il désigne par le nom de « vallée de l'étrange », avant que la courbe redevienne à nouveau croissante. Mori décrit ainsi sa métaphore de "vallée" :

> L'ascension d'une montagne offre un exemple de fonction qui n'augmente pas de façon continue. Compte tenu des collines et des vallées, l'altitude y d'une personne n'augmente pas nécessairement à mesure que la distance qui la sépare du sommet décroît. De la même manière, j'ai observé que plus les robots paraissent humains, plus notre sentiment de familiarité envers eux augmente, jusqu'à atteindre ce que j'appelle une vallée. J'ai nommé cette relation : "la vallée de l'étrange". (MORI, 2012, p. 27)

Selon Mori, plus l'apparence physique d'un robot androïde ressemble à celle d'un être humain véritable, plus ses imperfections mécaniques, ses mouvements imparfaits, nous paraissent monstrueux et dérangeants, jusqu'à un certain point au delà duquel l'identification empathique reprend à nouveau lorsque l'objet devient quasiment indiscernable de son modèle humain. En particulier, Mori constate que le mouvement de l'automate, qui figure, selon l'expression d'Henri Bergson, « du mécanique plaqué sur du vivant », participe pour beaucoup au sentiment d'étrangeté.

> Pour toute créature, y compris les robots, le mouvement est généralement signe de vie. Comme le montre la figure 2, l'ajout de mouvement transforme l'aspect de la vallée de l'étrange en venant en accentuer les pics et les creux. (MORI, 2012, p. 29)

Ce ne sont pas seulement les similarités de formes qui comptent pour accroître le degré d'empathie à l'égard de l'androïde, mais également les rapports de vitesse des gestes. Nous interprétons généralement les petits mouvements des muscles du visage d'une personne comme des signes qui révèlent les intentions de cet individu. Lorsque la programmation du robot ne parvient pas à rendre parfaitement le jeu des contractions, l'interprétation des signes échoue, et l'objet prend alors une allure inquiétante :

> Dans l'Exposition universelle de cette année à Osaka, les robots affichaient un design plus élaboré. Par exemple, l'un de ces robots possédait vingt-neuf muscles artificiels intégrés au

> visage afin d'humaniser ses expressions faciales. D'après son concepteur, le rire correspond à une séquence de distorsions faciales dans laquelle la vitesse est un facteur fondamental. Si la vitesse est réduite de moitié, le rire paraît artificiel. Cela illustre comment de petites variations de mouvement peuvent provoquer la chute d'un robot, d'une poupée ou d'une prothèse de main dans la vallée de l'étrange. (MORI, 2012, p. 29)

Étant donné qu'il n'est pas encore possible de construire de parfaits androïdes, la thèse de la vallée étrange n'est encore qu'une conjecture psychologique qu'il reste à confirmer ou à infirmer. Mais elle n'en demeure pas moins une piste intéressante pour comprendre nos réactions d'empathie. En outre, elle peut aussi fournir une explication au phénomène de détournement que Zazzo observe dans le cas du miroir. Le sentiment d'étrangeté proviendrait, cette fois, non d'un mouvement imparfait, mais au contraire d'un mouvement parfaitement synchronisé de l'image sur les mouvements de l'enfant, ce qui pourrait susciter un sentiment d'étrangeté lorsque le reflet est encore tenu pour un individu bien réel.

Le caractère modal du miroir

Dans la littérature, on trouve bon nombre d'œuvres qui jouent de l'illusion spéculaire. L'altérité du monde du miroir est le thème de contes pour enfants, comme le célèbre *De l'autre côté du miroir* (1872) de Lewis Carroll ou comme le moins connu *L'échiquier devant le miroir* (1921) de Massimo Bontempelli. Dans ce récit, le narrateur décrit une aventure qu'il prétend avoir vécu enfant. Après avoir été puni, il fut enfermé dans une pièce ne comportant pour tout objet qu'un seul miroir qu'il entreprit de traverser. Passé de l'autre côté du miroir, il s'étonne auprès d'un habitant de ce monde de ne pas reconnaître la pièce qu'il vient de quitter : « Mais de ce côté je croyais que c'était comme en face avec une cheminée, un mur bleu. » Son interlocuteur le détrompe immédiatement sur cette illusion : « Oui, dès qu'on franchit le miroir tout est pareil jusqu'au mur bleu. Mais après le mur, tout change. » (BONTEMPELLI, 1990) La moralité de l'histoire est donc que si tout paraît identique de prime abord, un examen plus scrupuleux montre que le monde du miroir diffère bien du monde réel.

Car de fait, dans le miroir, nous avons une image trompeuse, renversée de nous-mêmes, image que nous corrigeons instantanément lorsque nous nous reconnaissons. Dans une série d'expérimentations qui suivirent celles décrites précédemment, Zazzo a placé des enfants devant un moniteur vidéo diffusant directement leur propre image. Le changement de dispositif est motivé par le fait que Zazzo souligne une différence entre l'image vidéo ou photographique d'une part et l'image spéculaire d'autre part :

> c'est l'image du miroir qui est trompeuse mais notre identification à l'image spéculaire est si totale que nous lui attribuons le même espace que nous. Une illusion que nous n'aurions jamais avec une photo. L'image perçue sur l'écran est donc *anti-spéculaire* [...] (ZAZZO, 1993, p. 54)

Les résultats obtenus par Zazzo indiquent que les enfants ne semblent pas du tout surpris que, lorsqu'ils lèvent le bras gauche, ils voient effectivement leur image lever le bras gauche sur l'écran. L'anti-spécularité que leur renvoie l'image projetée par le film n'est pas un objet d'étonnement,

en dépit de remises en question suscitées par l'expérimentateur demandant à l'enfant « es-tu vraiment certain que c'est le même bras qui se lève ? ». Mais lorsque l'image est volontairement désynchronisée, arrêtée ou retardée, l'enfant tend d'abord à nier l'identité, puis, sous l'insistance de la ressemblance de son putatif sosie, finit par admettre que « c'est moi quand-même ».

Prolongeant ses expériences sur des adultes, Zazzo leur pose la question :

> Êtes-vous droitier ? – Oui – Mais votre image dans le miroir est celle d'un gaucher ! (ZAZZO, 1993, p. 75)

et observe que cette simple affirmation provoque une réaction de surprise généralement suivie de l'aveu « C'est vrai ! Je n'y avais jamais pensé ! » Nous avons donc tendance à croire que l'image spéculaire est plus fidèle que l'image filmée, alors que c'est justement le contraire qui est vrai [35].



Le reflet autorise une mise à distance du sujet qui stimule l'apprentissage de l'auto-identification. La construction de notre vision autoscopique, nous l'avons vu, passe par la relation conflictuelle entre deux types de répliques épistémiques : l'image répétée (celle du film, la figure-filtre) et l'image spéculaire (la figure-écho). Nous apprenons à nous projeter dans l'une et dans l'autre. Après quelques transactions liées à notre apprentissage des reflets et compréhension des règles qui gouvernent l'espace spéculaire, nous nous approprions notre image spéculaire comme un double épistémique : nous apprenons de notre reflet. Les expériences de Zazzo montrent en effet que nous élaborons progressivement un sentiment de familiarité avec le double spéculaire, allant de l'empathie (pour quelqu'un) jusqu'à l'identification (de soi) alors même qu'il nous renvoie une image infidèle, l'image d'un monde inversé. Ce n'est d'ailleurs pas la seule trahison du miroir, puisqu'il renvoie une image deux fois plus petite que la réalité [36]. Par l'apprentissage du miroir, l'identité personnelle est élaborée à partir d'un monde irréel, imaginaire. L'apprentissage de l'image spéculaire, qui marque une étape du passage de l'identification à la troisième personne (*de re*) à l'identification à la première personne en tant que première personne (*de se*), indique l'importance de la représentation de soi dans la construction de l'identité personnelle. Les expériences de Zazzo soulignent ainsi que, dans le reflet, l'identité n'est ni reconnue ni découverte de but en blanc, mais qu'elle est progressivement *élaborée*. Le rapport intentionnel *de se* à notre image n'est pas donné d'emblée, il est construit. En outre, son élaboration n'est pas tirée de comparaisons *de re* à partir de ressemblances entre le sujet et son reflet (celles-ci restant hors d'atteinte du sujet), mais à partir de relations basées sur la synchronicité spatio-temporelle, autrement dit, extrinsèques.

[35] À ce titre, dans un article inspiré par les phénomènes d'autoscopie rapportés par Freud, Manuel Rebuschi a fait remarquer le caractère modal des miroirs : « en s'observant dans le miroir, [on] peut avoir des attitudes à propos de [soi]-même si la droite et la gauche étaient inversées, autrement dit dans des mondes possibles qui sont comme le monde actuel à ceci près que la gauche et la droite sont inversées. Le fait que les miroirs (non déformants) représentent un monde très proche du monde actuel, au point que l'on peut faire intellectuellement la correction assez facilement pour reconstituer le monde actuel peut contribuer à expliquer pourquoi la dimension modale du miroir est restée inaperçue » (REBUSCHI, 2012, p. 181-182).

[36] On peut déduite ceci de la simple application du théorème de Thalès aux lois de l'optique. Voir ici une illustration dynamique de ce phénomène : https://mathematiqueslievre.pagesperso-orange.fr/Site/Activites/Thales1.html.

La création de soi procède ainsi d'une imitation de l'autre dont on peut sans doute retracer une partie de la généalogie dans le rapport de l'enfant à l'image dans le miroir, reflet qui, pour l'enfant, apparaît d'abord comme un autre imitant ses propres gestes.

7.5 S'imaginer comme un autre et se croire autre

Nous avons vu que l'apprentissage de la conscience de soi passait par des phénomènes de duplications, la constitution de doubles fonctionnels. Il reste à voir maintenant comment l'on peut encore échapper à une théorie substantialiste tout en admettant que l'on ne duplique rien, c'est-à-dire, en évitant de stipuler des doubles ontologiques du soi. Nous verrons que l'usage du pronom « je » relève d'un acte performatif par l'occupation d'un point de vue plutôt qu'un acte linguistique de référence à un sujet pensant.

Le point d'ancrage de cette section consiste en une analyse d'attitudes d'imagination de duplication à la première personne comme celle qui semble intervenir dans une question comme « et s'y j'étais quelqu'un d'autre ? ».

7.5.1 Imagination et pensées personnelles

Deux manières de s'imaginer autre

Lorsque nous disons, à la troisième personne, qu'un individu est devenu un autre, nous voulons dire par là qu'un trait important de sa personnalité a été affecté. De fait, un accident, une maladie, un événement dramatique peuvent profondément changer une personne, en modifier le comportement et le caractère. Mais que signifie au juste « s'imaginer être un autre » à la première personne ?

Envisager qu'on est un autre consiste, pareillement à ce qui vient d'être dit, à s'auto-attribuer des qualités étrangères. Dans une large mesure, nous pouvons nous imaginer différents de ce que nous sommes. De fait, notre carrière dépend d'un nombre important de facteurs qui auraient pu ne pas se réaliser. Nous pouvons ainsi nous demander ce que nous serions devenus si nos parents avaient choisi d'habiter à tel endroit plutôt qu'à tel autre lorsque nous étions enfants, si nous n'avions pas rencontré notre compagne ou compagnon, ou si nous avions opté pour tel emploi plutôt que tel autre. De telles pensées contrefactuelles sont certainement imaginaires, mais elles n'en demeurent pas moins des pensées plutôt courantes, et contribuent manifestement à notre identité narrative. Et d'ailleurs, s'imaginer différent est moteur pour devenir, effectivement, différent.

Remarquons cependant que nous pouvons aussi nous leurrer sur ce que nous sommes ou sur ce que nous pourrions être afin de ne pas nous confronter à la réalité, et basculer ainsi dans l'illusoire. Si l'identité personnelle est réductible à une identité narrative, il peut alors arriver qu'un sujet perde le récit de lui-même, voire se trompe complètement sur son identité, ou encore qu'il se prenne pour une autre personne (existante). Les amnésies bloquent l'accès aux souvenirs, et par là, à la reconnaissance des relations familiales et sociales. Le sujet amnésique profondément affecté par la maladie parle encore de lui à la première personne, mais il ne sait plus de au juste de qui il s'agit. Dans certaines maladies comme la schizophrénie, l'impossibilité de se constituer un moi implique une forme d'indifférenciation du corps avec ce qui ne relève pas de lui, comme c'est le

cas de Julie, hébéphrènes dont les propos sont rapportés par le psychiatre Ronald D. Laing :

> Ça, c'est la pluie. Je pourrais être la pluie (...) cette chaise, ce mur... Je pourrais être ce mur. C'est une chose terrible, pour une fille, d'être un mur... (LAING, 1979, p. 220)

Quoi qu'il en soit, ces premières remarques indiquent qu'on peut donner deux interprétations possibles à l'idée de « s'imaginer être un autre ». La première consiste à s'imaginer *devenir* un autre, la seconde, à *prendre la place* d'un autre. Christian Dours (DOURS, 2003, p. 53) souligne que la première manière procède par dédoublement d'une des composantes de la personne, soit qu'il s'agisse du corps, de la mémoire ou de la conscience alors que la seconde procède par une permutation d'une partie de la personne avec une partie d'une autre personne, soit, là aussi, qu'il s'agisse du corps, de la mémoire ou de la conscience. Dans le premier cas, un individu compte maintenant pour deux, alors que dans le deuxième cas, les comptes restent un-pour-un.

Comme le souligne Jean-Pierre Vernant (VERNANT, 1989), dans la mythologie, le dédoublement par projection dans autre chose que son propre corps est souvent associé au *démembrement*. Se projeter dans un reflet revient en quelque sorte à se désolidariser de ses parties constitutives. De manière moins métaphorique, nous sommes capables de nous identifier à autrui, de nous imaginer différents, de nous retrouver ailleurs que là où nous sommes réellement, ce qui constitue, en quelque sorte, des dédoublements imaginaires de soi. De fait, la manifestation d'une personne s'étend bien au delà de son corps, et cela, de bien des façons ordinaires, sans que cela nécessite de faire appel à une vision schyzophrénique de l'identité personnelle. Comme nous l'avons déjà signalé à travers l'illustration du film *Plein Soleil*, la manière d'organiser son logement trahit déjà la personnalité, de sorte que la présence d'une personne irradie dans d'autres corps matériels que le sien propre. Nous savons, nous pouvons prendre place « ailleurs », et nous savons aussi identifier autrui à travers autre chose que sa pure présence corporelle. Mais, comme je souhaite le montrer dans ce qui suit, cela n'exige pas du tout de reconnaître en soi ou en autrui, une substance immatérielle comme surajoutée au corps.

Les imaginations *de re* et les imaginations *de se*

Habituellement, l'identification à ce qui n'est pas soi est entravée par différentes barrières immuno-dédoublantes, comme l'aspect du corps, les souvenirs, le caractère, la conscience de soi et les intentions. Caractériser l'imagination à la première personne nécessite déjà de préciser certaines caractéristiques des pensées dites réflexives. Par comparaison aux pensées *de re* et *de dicto*, commençons par rappeler que Willard Van Orman Quine a montré que l'interprétation formelle des propositions intensionnelles est sensible à l'étendue des quantificateurs. Considérons l'expression suivante :

(ex 1) Jean imagine un empereur.

D'un coté, Jean pourrait imaginer n'importe quel empereur, mais pas un empereur en particulier. Nous pourrions ainsi traduire l'expression sous la forme *de dicto* (Ex signifiant « x est empereur ») :

(ex 2) Jean imagine $\exists x$ Ex

D'un autre côté, Jean pourrait tout aussi bien imaginer un empereur particulier, mettons, Napoléon. Dans ce cas, la pensée de Jean se rapporterait à une entité singulière, et l'expression devrait alors être traduite maintenant sous la forme *de re* :

(ex 3) $\exists x$ Jean imagine Ex

Considérons maintenant l'expression suivante :

(ex 4) Jean s'imagine empereur.

Nous avons là une pensée à la première personne, marquée par l'usage d'un pronom réflexif. C'est maintenant Jean qui imagine que *lui-même* est empereur. La question de savoir si une telle pensée à la première personne, ou pensée *de se*, relève du cas d'une pensée générale *de dicto*, d'une pensée particulière *de re*, voire à aucune des deux.

Jean pourrait-il *s'imaginer* être n'importe quel empereur ? Non, il semble que nous puissions éliminer d'emblée la forme *de dicto*, parce que Jean n'imaginerait pas *n'importe quel empereur*, mais bien une personne en particulier. A savoir, lui-même, avec les qualités d'empereur. Sa pensée est donc singulière, elle n'est pas générale. Ceci étant, peut-on alors ranger cette pensée *de se* dans la catégorie des propositions particulières *de re* ?

De fait, une imagination *de se* a ceci en commun avec une pensée *de re* qu'elle concerne bien un individu particulier, Jean. Mais il subsiste toutefois une différence de taille : le particulier en question, Jean, est aussi le sujet de l'attitude propositionnelle. C'est Jean en tant que lui-même qui imagine être empereur. Les pensées réflexives ont ainsi été caractérisées par Frege comme étant un type à part de pensées : « every one is presented to himself in a particular and primitive way, in which he is presented to no one else » (FREGE, 1956, pp. 25-26). Les pensées *de se* sont caractérisées par leur dépendance à un mode spécifique de présentation. Reprenant l'idée de Frege, Hector-Neri Castañeda (CASTAÑEDA, 1968), John Perry (PERRY, 1977) et David Lewis (LEWIS, 1979), ont montré dans les années 70 que les attitudes *de se* constituaient une catégorie irréductible de pensées [37]. Comme l'écrit Dawid Lewis, « self-ascription isn't quite the same thing as ascription *de re* to oneself. » (LEWIS, 1979, p. 543). L'irréductibilité se manifeste par le fait qu'une phrase contenant un pronom personnel (ou toute autre marque de la première personne) ne peut être remplacée par une autre ne contenant pas cette marque sans nécessairement en perdre le sens. Examinons cela de plus près.

Dans un passage souvent cité *Des temps difficiles* de Charles Dickens, alors qu'un personnage demande à sa mère si elle souffre, cette dernière lui répond de la manière suivante :

> Je crois qu'il y a une douleur quelque part dans la chambre [...], mais je ne pourrais pas affirmer que c'est moi qui en souffre. (DICKENS, 1854, p. 106)

[37] Cette thèse reste encore controversée, certains philosophes considérant les pensées *de se* comme un sous-genre de pensée *de re*. François Recanati (RECANATI, 2007a) range ainsi des pensées « *de se* explicites » parmi les pensées *de re* » quand l'objet que vise l'attitude y figure explicitement.

Dickens commente la réponse en la qualifiant d'« étrange ». De fait, il semble effectivement bien étonnant qu'une personne qui dit éprouver de la douleur puisse le faire d'une manière impersonnelle en doutant du sujet atteint par cette sensation.

C'est essentiellement Wittgenstein qui a indiqué la particularité des pensées *de se* : elles sont immunisées contre toute erreur d'identification. Dans le *Cahier bleu*, Wittgenstein souligne en effet que les expressions comportant un pronom personnel s'utilisent de deux manières différentes : selon qu'on les emploie comme objet ou comme sujet. L'emploi du pronom personnel « je » ou « moi » n'est référentiel que dans certaines situations. Il est référentiel quand le mot sert à parler d'un certain objet à propos duquel ont été observés certains faits, cet objet relevant d'un certain organisme humain localisé dans l'espace et le temps. Il est tout à fait primordial que de tels énoncés, supposés qu'ils soient sincères, relèvent d'une observation, et qu'ils soient conséquemment susceptibles d'être faux, comme « ma main est tâchée », ou comme « mon bras est cassé ». Lorsque j'affirme que « mon bras est cassé », il est effectivement possible de faire erreur sur l'objet car il concerne une personne particulière, qui, certes, peut-être moi, mais qui peut aussi ne pas l'être. Après un accident de train ayant agglutiné les corps des passagers les uns sur les autres, je pourrais croire que c'est mon bras qui est cassé quand il s'agit en réalité de celui de mon voisin.

Selon le second emploi des pronoms personnels, qui correspond à des expressions comme « j'ai mal aux dents » ou « j'imagine un rhinocéros », il est en revanche tout à fait impossible de se tromper. Nous n'avons pas ici affaire à des relevés informationnels effectués par observation, il s'agit de connaissances directes. Wittgenstein résume ainsi que

> [...] les cas de la première catégorie impliquent la reconnaissance d'une personne particulière, et dans ce cas il y a possibilité d'erreur, ou, et c'est plutôt ainsi que je le dirais : il est prévu qu'on puisse se tromper. [...] Au contraire, il n'est pas question de reconnaître qui que ce soit lorsque je dis "J'ai mal aux dents". (WITTGENSTEIN, 1965, p. 124-125)

Les pensées *de se* sont des pensées pour lesquelles aucun doute n'est permis, elles sont directement centrées sur la personne qui les pense, elles présupposent que le sujet est inévitablement celui qui emploi l'expression.

Afin de prouver le caractère rigide des indexicaux, David Kaplan a donné un exemple bien connu de pensée faisant apparaître une erreur d'identification de soi. Un homme se voit dans un miroir, sans se reconnaître, et constate : « le pantalon de ce type est en feu ! ». On a affaire à une pensée au sujet d'une personne qui, certes, est bien l'agent de cette pensée, mais pourtant, cette pensée n'est pas *de se*, elle n'est pas une pensée de soi *à la première personne*. Il peut donc arriver accidentellement qu'un contenu de pensée donné soit le contenu d'une pensée *de re* plutôt que d'une pensée *de se*, et comme le constate Kaplan, cela arrive par exemple quand le sujet n'a pas réussi à se localiser dans le monde. Si l'individu dont il est question dans l'exemple de Kaplan réalisait sa méprise, il n'apprendrait rien de plus à propos du monde, et il n'apprendrait rien de nouveau le concernant. Car il sait déjà, quoique de manière impersonnelle, que son pantalon est en feu. Informé de son erreur, il gagnera en revanche une croyance neuve concernant sa localisation dans le monde, en s'identifiant maintenant à l'objet que vise sa pensée.

S'il peut arriver ainsi qu'une pensée soit accidentellement *de re* plutôt que *de se*, en revanche, il est essentiel à une pensée *de se* qu'elle ne soit pas *de re*. Comme le dit Sydney Shoemaker (S.

SHOEMAKER, 1968), un énoncé traduisant une pensée *de se* exprime une proposition complètement « immunisée contre l'erreur d'identification » ;

> First-person statements that are immune to error through mis-identification in the sense just defined, those in which 'I' is used "as subject," could be said to have "absolute immunity" to error through misidentification. A statement like "I am facing a table" does not have this sort of immunity, for we can imagine circumstances in which someone might make this statement on the basis of having misidentified someone else (e.g., the person he sees in a mirror) as himself. But there will be no possibility of such a misidentification if one makes this statement on the basis of seeing a table in front of one in the ordinary way (without aid of mirrors, etc.); let us say that when made in this way the statement has "circumstantial immunity" to error through misidentification relative to 'I'. It would appear that, when a self-ascription is circumstantially immune to error through misidentification, this is always because the speaker knows or believes it to be true as a consequence of some other self-ascription, which the speaker knows or is entitled to believe, that is absolutely immune to error through misidentification; e.g., in the circumstances just imagined the proposition "I am facing a table" would be known or believed as a consequence of the proposition "I see a table in the center of my field of vision." (S. SHOEMAKER, 1968, p. 557)

Dans l'exemple de Shoemaker, l'emploi du pronom n'est plus référentiel, ce qui serait inutile, parce qu'un sujet conscient ne peut pas se prendre pour un autre que lui, ni prendre autrui pour lui. Il est utilisé seulement pour satisfaire le besoin purement grammatical d'un sujet, pour compléter, par défaut, l'usage d'un verbe, mais sans y ajouter de référence, puisque celle-ci est d'emblée donnée par sa focalisation. Il en irait ici de l'expression « je pense » comme il en va de l'expression « il pleut » où le pronom « il » ne renvoie pas à un sujet référentiel auquel le prédicat « pleuvoir » serait applicable, comme s'il y avait quelque chose en plus, mais où il s'agit simplement de satisfaire aux conditions de bonne formulation d'une expression, c'est-à-dire, à la satisfaction d'une règle grammaticale.

Selon Elizabeth Anscombe, qui radicalise le propos de Wittgenstein concernant la signification du pronom personnel, « je » n'est tout simplement pas un sujet grammatical, le mot ne fait pas référence à quelque chose sur lequel s'appliquerait un prédicat. Faire référence à un objet est toujours un acte transitif. Or cela réclame qu'il y ait contact contigu entre l'objet à nommer et le sujet qui le nomme. Mais en revanche, exprimer la conscience de soi demande, comme l'écrit Vincent Descombes, « une coïncidence pure et simple du sujet conscient et de lui-même, et c'est là ce qu'aucune contiguïté, aussi étroite soit elle, ne peut réaliser. » (DESCOMBES, 2010, p. 519)

Dans *Le parler de soi* (DESCOMBES, 2014), Descombes confronte deux manières de penser à soi, l'une, égocentrique, l'autre, égotiste. Une pensée égotiste est une pensée qui décrit un soi à la première personne. Le mot provient bien sûr de la critique littéraire où il sert à qualifier le style d'un écrivain se prenant pour la matière de son livre. Le style égotiste consiste ainsi à parler de soi à la première personne. Une pensée égocentrique fait au contraire intervenir un moi parlant de lui à la troisième personne. Descombes tire la conclusion que les analyse d'Anscombe montrent un inévitable échec des thèses égologiques « qui ont cherché à concevoir la conscience de soi comme le rapport d'un sujet à l'objet qu'il est pour lui-même » (DESCOMBES, 2010, p. 519). Même si l'on peut à juste titre remettre en question, dans l'expression « je pense », la référence à une substance

immatérielle, il s'avère toutefois que l'expression perdrait à être traduite de manière impersonnelle par « ça pense » parce que nous pourrions difficilement encore poser la question « qui pense (ça) ? » là où l'on attend une identification : « tiens, c'est lui (qui pense ça) ». Un nom propre peut être utilisé par méprise, mais pas le pronom « je ». Ainsi, il est possible de souscrire au *cogito* cartésien sous l'idée qu'il affirme qu'une pensée concerne celui qui la pense de telle manière qu'il ne peut en aucun cas ne pas avoir conscience que c'est bien de lui dont il s'agit. En revanche, cela n'engage pas ontologiquement envers une substance du moi.

Se croire un autre

La possibilité de se tromper d'identité à la première personne soulève la question suivante : comment pourrait-on à la fois supposer que les propositions *de se* soient immunisées contre l'erreur et pouvoir faire des erreurs d'identification sur sa propre personne ? Considérons les deux cas suivants :

(ex 5) Jean n'est pas Napoléon mais Jean croit qu'il est Napoléon.

(ex 6) Je ne suis pas Napoléon mais je crois que je suis Napoléon.

La signification de la première phrase, qui utilise la troisième personne, ne paraît pas poser de problème particulier. Jean pourrait croire ou non être Napoléon, de toute façon, il n'est pas Napoléon. En revanche, la deuxième phrase, à la première personne, semble contradictoire. Comment affirmerait-on la positivité d'un état de choses (« Je ne suis pas Napoléon ») en renonçant à y croire (« Je crois que je suis Napoléon ») ? Il semble qu'il y ait, ici, une dissymétrie dans la signification des propositions.

De fait, la contradiction n'en est qu'une qu'à la condition de considérer que l'expression « Je ne suis pas Napoléon » doit être interprétée comme un constat à la première personne. Dans ce cas, il y aurait bien contradiction, puisque constater ne pas être Napoléon impliquerait croire ne pas être Napoléon. Selon Wittgenstein, la contradiction logique n'est, en réalité qu'apparente. Il ne s'agit là que d'une ambiguïté grammaticale. Dans la deuxième expressions, la conjonction des deux propositions peut certainement sembler contradictoire, mais c'est plutôt qu'elle engendre un *trouble grammatical*. Quand je dis que je crois que P, c'est que, écrit Wittgenstein, « au fond, je décris avec ces mots mon propre état mental, – mais cette description est ici, indirectement, une assertion du fait même qui est l'objet de croyance. – Il en irait ici comme lorsque dans certains cas, je décris une photographie afin de décrire ce qui a été photographié. » (WITTGENSTEIN, 1989, § 481) La comparaison à la photographie est destinée à indiquer que la référence passe de manière transparente à travers la représentation. En décrivant mon état mental (je crois que je suis Napoléon), je décris une représentation dirigée vers ce que je crois être un état de fait (je suis Napoléon). Mais tout comme je peux très bien mal décrire une photographie (je me trompe sur la date ou sur le lieu) je peux me tromper sur la description de mon identité. Il n'y a donc pas vraiment de paradoxe ici, mais seulement un emploi ambigu du langage : dans la seconde expression, la première occurrence de « je » relève d'une identification à la troisième personne, contrairement à la seconde qui est véritablement à la première personne.

La contradiction n'est finalement qu'apparente parce que la croyance est d'emblée centrée sur un ensemble de mondes indépendamment des faits actuels. La première occurrence de « Je » parle du monde actuel où tel individu n'est *objectivement* pas Napoléon, et pour la seconde occurrence, il s'agit d'une identification fautive basée sur une croyance. Dans cette seconde situation, les mondes afférents sont ceux des croyances du sujet, mondes dans lesquels son point de vue coïncide effectivement avec l'entité Napoléon.

Le cogito comme acte performatif

Émile Benveniste a souligné que le cogito cartésien consiste sans doute moins en une preuve de l'existence d'une substance immatérielle du moi directement référée par le pronom je, qu'en un acte performatif autorisé par l'emploi du pronom personnel. Selon Émile Benveniste, c'est l'usage même du langage qui élabore la partition du monde entre le *soi* et le *non-soi*. Dans une série de trois articles, à savoir *Structure des relations de personne dans le verbe*, *La nature des pronoms*, *De la subjectivité dans la langue*, Benveniste a décrit la part du pronom personnel « je » dans l'échange verbal. De son point de vue de linguiste, Benveniste constate l'usage universel des pronoms personnels, termes qui existent dans toutes les langues. Il remarque également que les pronoms personnels ne renvoient jamais à un concept ni à un individu. Le pronom « je » ne réfère à aucune entité du monde, car « il n'y a pas de concept de « je » englobant tous les *je* qui s'énoncent à tout instant dans les bouches de tous les locuteurs, au sens où il y a un concept « arbre » auquel se ramènent tous les emplois individuels de *arbre*. Le « je » ne dénomme donc aucune entité lexicale. » (BENVENISTE, 1966, p. 260) Or, puisque le même terme « je » peut être emprunté par toute personne linguistiquement compétente, il n'en désigne aucune en particulier. Ainsi,

> *Je* pose une autre personne, celle qui, tout extérieure qu'elle est à « moi », devient mon écho auquel je dis *tu* et qui me dit *tu*. (BENVENISTE, 1966, p. 260)

L'aptitude à se penser comme sujet relève ainsi d'une capacité qui est de part en part linguistique :

> C'est dans et par le langage que l'homme se constitue comme *sujet* ; parce que le langage seul fonde en réalité, dans sa réalité qui est celle de l'être, le concept d' « ego ». La « subjectivité » [...] est la capacité du locuteur à se poser comme « sujet ». Elle se définit, non par le sentiment que chacun éprouve d'être lui-même (ce sentiment, dans la mesure où l'on peut en faire état, n'est qu'un reflet), mais comme l'unité psychique qui transcende la totalité des expériences vécues qu'elle assemble, et qui assure la permanence de la conscience. Or nous tenons que cette « subjectivité », qu'on la pose en phénoménologie ou en psychologie, comme on voudra, n'est que l'émergence dans l'être d'une propriété fondamentale du langage. Est « ego » qui dit « ego ». Nous trouvons là le fondement de la « subjectivité », qui se détermine par le statut linguistique de la « personne. » (BENVENISTE, 1966, pp. 259-260)

Benveniste fait donc remonter la découverte de la subjectivité à l'usage du langage, et en particulier, à l'usage des pronoms personnels. Cette théorie favorise un abandon de ce qui qualifie de

« vieilles antinomies du langage du « moi » et de l'« autre ». Le cadre de l'intersubjectivité est donc défini par Benveniste par le langage où ce sont les pronoms personnels qui ancrent les personnes.

Nous devrions donc trouver une troisième voie entre théorie substantielle de moi et théorie grammaticale du moi. Plutôt qu'un point de vue substantialiste ou strictement grammatical, on peut aussi faire appel à une théorie pragmatique du sujet. Dans *Cogito, Ergo Sum : Inference or Performance ?* (HINTIKKA, 1968, p. 6), Jaakko Hintikka constate que le *cogito* cartésien peut être discuté de deux manières différentes, premièrement comme inférence logique et deuxièmement de manière pragmatique comme un acte performatif. Hintikka montre dans son article que la certitude d'exister ne relève pas d'une validité logique, mais de la performance consistant à constater l'impossibilité de *se* penser inexistant.

D'abord, d'un point de vue logique, le raisonnement « Je pense, donc je suis » peut ainsi se traduire sous la formulation :

$$\mathrm{P}a \rightarrow \exists x(x = a)$$

Si l'on admet que l'argument du *cogito* est une inférence logique, on a alors l'inférence suivante :

(1)	$\mathrm{P}a$	hypothèse
(2)	$\mathrm{P}a \rightarrow \exists x\ (x = a)$	hypothèse
(3)	$\exists x\ (x = a)$	*Modus ponens* appliqué à (1) et (2).

Pour que le raisonnement soit validé, il faut bien sûr s'assurer des hypothèses, et en particulier de la deuxième. Celle-ci semble relever de l'équivalence suivante :

$$\mathrm{P}a \rightarrow (\exists x)\ \mathrm{P}x \wedge (x = a)$$

d'où l'on peut tirer aussi, sans grande difficulté, que :

$$\mathrm{P}a \leftrightarrow (\exists x)(x = a)$$

L'argument du *cogito* peut alors prendre différentes formes, comme celle que Descartes tire de l'objection de Gassendi à la seconde métaphysique, à savoir, « *ambulo ergo sum* », voire d'autres formes plus ou moins divertissantes comme « j'aime le camembert au lait cru, donc je suis » puisque l'inférence ne dépend finalement pas du prédicat P qui est utilisé dans la prémisse [38].

Descartes se défend de rendre équivalent les deux raisonnements. Contrairement au *cogito*, le fait qu'un corps se promène n'est absolument pas indubitable. Le raisonnement ne peut tenir que du fait d'avoir conscience de se déplacer, mais non du déplacement lui-même : « de ce que je pense me promener, je puis fort bien inférer l'existence de mon esprit, qui a cette pensée, mais non celle de mon corps, lequel se promène » (DESCARTES, 1953c, p. 478).

[38] Contrairement à ce qui est souvent écrit, l'exemple *ambulo ergo cogito* provient de Descartes et non de Gassendi, lequel se contente d'objecter à Descartes, d'une manière plus générale, « Vous n'aviez pas besoin d'un si grand appareil pour prouver que vous êtes ; cela peut se conclure de toute action, car le noeud de la difficulté n'est pas de savoir si l'on existe, mais ce que l'on est » (DESCARTES, 1953c, p. 470). Descartes adapte en fait ici une objection d'Hobbes aux méditations, lequel reprochait au *cogito* d'être un raisonnement spécieux pouvant tout aussi bien légitimer une déduction comme « puisque je suis promenant, donc je suis promenade » (DESCARTES, 1953c, p. 400)

Mais quoi qu'il en soit, et contrairement à la forme inférentielle qui apparaît dans le *Discours de la méthode* (*cogito ergo sum*), Descartes refuse de considérer que la conclusion « je suis » soit tirée de la prémisse « je pense », et c'est d'ailleurs sous la forme « je suis, j'existe » que le *cogito* est exprimé dans les *Méditations Métaphysiques*. Pour Descartes, le *cogito* est certainement moins une inférence logique qu'une expérimentation. Maurice Merleau-Ponty a ainsi souligné le caractère performatif du *cogito* cartésien :

> C'est Moi qui reconstitue le *cogito* historique, c'est moi qui lit le texte de Descartes, c'est moi qui y reconnaît une vérité impérissable et enfin de compte le *cogito* cartésien n'a de sens que par mon propre *cogito*, je n'en penserai rien si je n'avais en moi-même tout ce qu'il faut pour l'inventer. (MERLEAU-PONTY, 1945, p. 425)

Comme le soutient aussi Hintikka, l'affirmation « je suis » n'est pas logiquement incontestable, c'est seulement le cas lorsqu'elle est médiatisée par une performance de pensée qu'elle devient indubitable, plus précisément, à travers l'impossibilité de se penser inexistant :

> Hence the indubitability of this sentence is not strictly speaking perceived by means of thinking (in the way the indubitability of a demonstrable truth may be said to be) ; rather, it is indubitable because and in so far as it is actively thought of. (HINTIKKA, 1968, p. 16)

Dès lors, le *cogito* n'est plus comparable aux raisonnements du type *ambulo ergo sum*, parce qu'il s'en différencie par son aspect performatif. En effet, comme le souligne Hintikka, le *cogito* tire sa force du fait qu'il m'est impossible de concevoir, à travers un acte de penser, ma propre inexistence [39] :

> The performance (act) through which the existential self-verifiability is manifested cannot be any arbitrary human activity, contrary to what Gassendi claimed. It cannot be an act of walking or an act of seeing. It cannot even be an instance of arbitrary mental activity, say of willing or of feeling. It must be just what we said it is: an attend to think in the sense of making myself believe [...] that I do not exist. (HINTIKKA, 1968, p. 17)

Selon Manuel Rebuschi, une conception ontologiquement neutre de l'identité personnelle n'est pas incompatible avec l'expérimentation du *cogito* cartésien. Précisément, on peut faire tenir cet acte performatif avec l'idée de Wittgenstein que le Moi n'est rien d'autre qu'un certain point de vue sur le monde plutôt qu'une partie du monde. De manière plus technique, Rebuschi utilise à cet effet la logique libre et l'interprétation substitutionnelle de la quantification. « Je » n'est pas un nom propre qui référerait à une entité du monde, mais un outil du langage destiné à manifester une position spatio-temporelle, un point de vue, autrement dit, quelque chose qu'il n'est ni illusoire ni irréel, mais qui n'a pas non plus un poids ontologique.

Quel est l'intérêt d'une version neutre de l'argument cartésien ?

> De « je pense » on infère l'existence [...] d'un sujet à la première personne, exprimé par le sujet logico-grammatical « je », qui n'est pas un objet du monde - pas même un objet

[39] Bien qu'il soit concevable que j'aurais pu ne pas exister.

> inexistant. À quoi sert le « je » ? À exprimer un point de vue : le point de vue à la première personne. Le sujet doit être conçu comme un point de vue et rien de plus. Un point de vue, ce n'est pas une substance au sens ontologique, c'est moins qu'un trou ou une ombre. Pour autant, on ne peut pas dire que cela n'existe absolument pas. Il existe des points de vue sur le monde, comme il existe une infinité de chemins entre deux points physiques : les points de vue peuvent exister comme les chemins en un sens neutre, sans venir peupler l'ontologie. (REBUSCHI, 2011)

Le *cogito* est identiquement vrai pour différents individus parce que tout un chacun est capable de l'expérimenter en prenant place dans le "je" ouvert. Mais qu'arriverait-il si, comme le propose Hintikka, Hamlet faisait lui-même l'expérimentation du cogito par le biais de l'acteur ? La performance ne devrait pas être la prérogative de Descartes ou de toute autre personne existante, mais devrait également pouvoir être faite par un acteur jouant un personnage de fiction. Comme le constate Manuel Rebuschi, « il serait en effet aussi absurde d'entendre (un acteur jouant) Hamlet dire "Je n'existe pas" que de l'entendre dire par Descartes » (REBUSCHI, 2011, p. 5) La validité pragmatique du cogito devrait tout aussi bien concerner des personnes réelles que les personnages d'une fiction quelconque, sans quoi celle-ci témoignerait d'une forme d'inconsistance. « Je » est une locution employée par l'auteur d'un roman, mais qui peut tout aussi bien concerner un narrateur imaginaire ou un personnage. Et dans ce cas, nous devrions leur reconnaître un poids ontologique.

Pour éviter de telles situations, Rebuschi propose de neutraliser le contexte d'énonciation du cogito. Habituellement, un contexte est caractérisé par un quadruplet composé d'un agent, d'un lieu, d'un instant, d'un monde possible. A côté de ces contextes, Recanati (RECANATI, 2000) introduit également des contextes imaginaires afin de rendre compte que, dans le roman, ce n'est plus l'auteur, c'est-à-dire le locuteur effectif qui s'exprime à travers l'emploi du pronom personnel « je », mais un locuteur imaginaire. Développant l'idée de Recanati, Voltolini (VOLTOLINI, 2006) parle de contexte fictionnel pour désigner tout contexte dans lequel le locuteur effectif *fait comme si* au moins l'un des paramètres du contexte est autre qu'il n'est. Un contexte fictionnel est donc un contexte dans lequel ou l'agent, ou le lieu, ou l'instant, ou le monde est fictif.

Rebuschi introduit alors les contextes de simulation, contextes fictionnels dont l'unique paramètre modifié est l'agent, quitte à être imaginaire à la manière d'un personnage de fiction. Selon Rebuschi, les discours à la première personne témoigneraient d'une démarcation vague entre contextes de simulation et contextes d'énonciation. Quoiqu'un tel discours serait toujours dû à un individu physique du monde, le locuteur effectif, l'emploi du l'expression « je » pourrait aussi bien renvoyer à l'auteur qu'à un personnage dont il jouerait le rôle.

La question est donc de savoir, quand on emploie « je », qui parle. Est-ce l'auteur ou s'agit-il d'un personnage ? et dans ce cas, quel personnage au juste ? Dans certains contextes, la réponse pourrait être tranchée. C'est par exemple le cas dans la situation décrite par Molière dans *L'Avare*, où Maître Jacques occupe alternativement différentes fonctions, à la fois cuisinier et cocher d'Harpagon. Alors qu'Harpagon s'adresse directement à lui, Maître Jacques demande :

> Maître Jacques : - Est-ce à votre cocher, Monsieur, ou bien à votre cuisinier, que vous voulez parler ; car je suis l'un et l'autre.
> Harpagon : - C'est à tous les deux.

Maître Jacques : - Mais à qui des deux le premier ?
Harpagon : Au cuisinier.
Maître Jacques : - Attendez donc, s'il vous plaît. (Il ôte sa casaque de cocher, et paraît vêtu en cuisinier.)

Ici, les rôles sont bien distingués, aidés en cela par des tenues d'apparat et des corvées bien démarquées. Mais il n'est pas du tout certain qu'une réponse nette puisse être apportée dans tous les cas. En outre, si l'on peut considérer l'identité personnelle comme un faisceau de rôles sociaux successifs, le doute peut alors être porté sur l'existence d'un « vrai moi », le *real me* évoqué par William James, qui subsisterait par deçà ou au delà des personnages incarnés. On peut ainsi mettre en doute l'idée d'un contexte d'assertion exclusivement non-fictionnel. Si nous nous projetons dans des personnages pour parler de nous, pourquoi devrait-on encore faire intervenir, outre ces personnages, une personne plus profonde, plus vraie ou plus réelle ?

Chez Rebuschi, l'importance des contextes de simulation dans la caractérisation ontologiquement neutralisée de l'identité personnelle rejoint la thèse de Mead que le moi est constitué par des jeux de simulation, et contribuent non seulement à rapprocher, mais plus fortement encore à identifier la notion de personne à celle de personnage. Comme l'écrit Rebuschi, « le sujet est dans l'œil de l'interprète, il n'est pas dans le monde ». (REBUSCHI, 2011, p. 19)

Reste à savoir si, entre les rôles imaginaires joués, l'on peut encore effectuer un tri hiérarchique, en admettant que certains rôles sont plus fondamentaux que d'autres. Les rôles occasionnés par nos situations familiale et professionnelle sont sans doute plus importants que d'autres. Si certains rôles nous conviennent, s'ils sont attendus de nous, d'autres sont indus, parfois pour la raison qu'ils sont déjà occupés. Ne nous prenons pas pour Napoléon, c'est-à-dire pour un personnage que nous ne sommes pas.

7.5.2 La duplication de l'imaginant en imaginé dans l'imagination *de se*

Se projeter dans autre chose

De telles réserves n'excluent toutefois pas que nous puissions nous projeter, dans certains contextes, dans un imaginaire « d'être Napoléon », par exemple dans un jeu de rôle ou dans un jeu vidéo. Il existe d'ailleurs bien des raisons d'imaginer être quelqu'un d'autre. Cela peut résulter d'une usurpation d'identité comme celle qui fit l'enjeu de l'affaire *Martin Guerre*. Cela peut aussi provenir de phénomènes beaucoup moins spectaculaires comme pour indiquer une opinion ou donner un conseil : « à ta place, etc. ». En outre , l'identification peut aussi se rapporter à quelque chose qui n'est pas actuellement une personne, comme l'identification empathique qui lie un lecteur d'un roman au protagoniste de son récit.

En ce sens, s'imaginer être un autre est une activité assez ordinaire. Est-ce vrai de toute identité imaginée ? Comparativement, si l'on nous demandait d'imaginer que Néron est Napoléon, nous serions certainement bien plus embarrassés. Comment devrions-nous procéder ? Nous serions en peine de savoir par quel bout prendre la demande. Devrions-nous imaginer un général revêtu d'un bicorne et mettant le feu à Rome, où devrions-nous imaginer un empereur en toge affrontant les troupes austro-russes à l'aide de catapultes ? À première approximation, l'identité imaginaire à la

première personne ne semble pas souffrir de telles difficultés.

Considérons un exemple. Jean pourrait s'imaginer être Napoléon. La question qui sera traitée dans ce qui suit est alors la suivante : dans le monde de son imagination, qu'est-ce qui est retiré ou ajouté à Jean et qu'est-ce qui est ajouté ou retiré à Napoléon ? S'agit-il d'une usurpation laissant la victime ontologiquement indemne du rapt d'identité ? Ou s'agit-il d'une sorte d'annihilation métaphysique ? S'agit-il d'imaginer un changement qualitatif mais pas quantitatif des entités en présence, de sorte que nous aurions affaire au même genre de permutation que dans l'histoire du Prince et du savetier de Locke ?

Imaginations *de re* et imaginations *de se*

En quoi consiste précisément le fait d'imaginer être Napoléon ? Cela ne consiste sans doute pas à imaginer être un personnage ayant une double biographie, celle de Napoléon et la nôtre. Nous ne sommes pas contraints d'imaginer que notre biographie soit mêlée à celle de Napoléon. Lorsque nous imaginons être Napoléon, nous abandonnons probablement nos propres caractéristiques au profit de celles de Napoléon. Du moins, cela est certainement vrai pour nos origines et notre histoire personnelle. Imaginons-nous changer de corps ? Sans doute devrions-nous imaginer avoir les cheveux bruns et souffrir d'un ulcère à l'estomac, même si ce n'est pas actuellement le cas. Mais que penser d'abandonner notre sensibilité ou notre tempérament ? Quelles sont les limites aux altérations que peuvent subir les pensées de soi ?

Considérons d'abord l'identification à la troisième personne. Qu'est-ce que cela signifie alors que Jean s'imagine comme Napoléon ? Dans *Imagination and the Self* Bernard Williams souligne que la compréhension d'une telle attitude nécessite de faire intervenir trois rôles (WILLIAMS, 1973). Il y a bien sûr l'imaginant, c'est-à-dire celui qui imagine, et il y a Napoléon. Mais il faut encore faire intervenir un troisième rôle, celui de l'imaginé, autrement dit celui qui est imaginé être Napoléon. Selon Williams, la question cruciale de nos identifications est alors « qui est l'imaginé ? »

En un sens, nous pourrions vouloir simplement affirmer que « Jean aurait pu être Napoléon », c'est-à-dire que l'imaginant aurait pu être Napoléon. Du point de vue de la sémantique modale, la théorie essentialiste de Kripke affirme que c'est tout à fait impossible. En effet, le principe de nécessité de l'identité stipule que pour tout couple (a,b) de désignateurs rigides, si a est actuellement identique à b, alors a est nécessairement identique à b. Il est donc impossible que deux individus actuellement distincts soient possiblement identiques.

Reste à savoir si imaginer être Napoléon implique nécessairement pour Jean de s'imaginer être identique à Napoléon. Qu'en est-il de l'imaginé, entre Jean et Napoléon ? Cette personne qu'imagine Jean, est-elle Napoléon en plus d'être Jean ? Tout dépend sans doute de ce que l'on entend par « identité de l'imaginant ». Faisons l'hypothèse qu'on a une identité entre l'imaginé et l'imaginant en admettant d'abord un moi « empirique », un moi qui est donné par la connaissance, c'est-à-dire par les qualités et les relations. L'identification serait directement contradictoire : le « moi » empirique de Jean ne contient pas la propriété d'avoir remporté la bataille d'Austerlitz, il ne peut être identique au moi imaginé. Si Jean est imaginé être Napoléon, celui qu'il imagine, dira-t-on, c'est bien Jean lui-même, mais à condition de bien voir que ce Jean n'est pas le Jean habituel, mais un Jean « napoléonisé », c'est-à-dire un « Jean en Napoléon ». Mais à quelle sorte

d'identité aurions-nous affaire ici ?

Imaginer être Napoléon, cela pourrait alors simplement signifier imaginer avoir une carrière différente ou exemplifier d'autres caractéristiques physiques que celles que Jean possède actuellement. De fait, il ne semble pas impossible que celui qui imagine eut pu être différent de ce qu'il est actuellement. Peut-être aurait-il pu avoir des yeux d'une couleur différente. Peut-être aurait-il pu naître plus tôt. Peut-être aurait-il pu participer à des événements qu'il n'a pas connu. Peut-être que... mais au fait, quand devrions-nous au juste cesser nos spéculations contrefactuelles ? Si nous lâchons la bride aux possibilités de changement, nous pourrions bien admettre que l'imaginant eut *toutes* les caractéristiques de Napoléon. Jean peut imaginer porter un bicorne, il peut imaginer être originaire de Corse. Il peut imaginer être né plus tôt. Il peut imaginer avoir fait, à la place de Napoléon, tout ce que Napoléon a fait. Il peut imaginer « être Napoléon » en tant qu'il se rapporte à lui sous certaines descriptions, et tant qu'il ne prétend pas être numériquement identique à Napoléon. Mais qu'est-ce que cela signifierait ? La question n'est rien d'autre que celle de l'individuation des personnes dans les mondes possibles. Si une personne peut, en imagination, se défaire de toutes ses propriétés individuantes, est-ce encore la même personne ? Sinon, à partir de quel stade Jean, s'imaginant qualitativement différent, est-il devenu numériquement différent ? Passant ce seuil, ne dirait-on pas que ce n'est plus *Jean* qui s'imagine Napoléon, mais tout simplement Napoléon ?

En échafaudant une série de mondes possibles où deux individus permutent successivement leurs propriétés respectives, on finit par obtenir deux mondes indiscernables où les individus ont permuté leur rôle. Nous retrouvons bien sûr, appliqué à un cas d'imagination, le paradoxe de Chisholm. Qu'est-ce qui peut alors supporter l'identité de Jean et de Napoléon ? Cette identité devrait alors relever d'un moi complètement affaibli, réduit à n'avoir aucune propriété spécifique. Autrement dit, une conception de l'identité personnelle qui reconnaîtrait des particuliers nus, des egos cartésiens.

Imaginer que nous soyons Napoléon plutôt que nous-mêmes nécessite-t-il de viser un monde dans lequel nous abandonnons notre propre « moi » au profit de celui de Napoléon ? Cela indiquerait alors un monde dans lequel nous n'existerions pas. Il semble certes plausible que l'auteur de ce texte aurait pu ne pas exister. Mais qu'est-ce que cela voudrait dire d'un point de vue cartésien ? Ne pas exister ne signifierait pas ne pas avoir eu de corps, mais plutôt ne pas avoir été une conscience immatérielle, un « ego » cartésien. Or, comme nous l'avons déjà fait remarquer, il semble impossible de distinguer en quoi une conscience cartésienne est différente d'une autre. Si l'explication cartésienne de l'individuation des personnes par des consciences immatérielles était vraie, il pourrait s'avérer que la disparition d'une conscience laisse un monde qualitativement indiscernable du nôtre. C'est ce que souligne aussi Williams :

> But since there is nothing to distinguish my Cartesian consciousness from any other, it is impossible to see any more what would be subtracted from the universe by the removal of me. (WILLIAMS, 1973, p. 42)

L'hypothèse de mon inexistence devrait pourtant faire une différence avec le fait de mon existence, mais sans avec une théorie cartésienne du moi, on ne voit pas au juste ce que serait cette différence. On voit que les difficultés dans lesquelles s'enlisent les conceptions substantialistes de l'identité lorsqu'elles sont confrontées aux identifications imaginaires. Leur issue de secours consiste à supposer des particuliers nus, des moi cartésiens, autrement dit, des doubles ontologiques.

Williams souligne qu'imaginer être Napoléon ne devrait pas consister à concevoir une identité impossible. Il en va seulement ici comme de l'acteur qui s'identifie à son personnage :

> For the fact that I can, in the only way that arouses my interest, imagine being Napoleon has no tendency at all to show that I can conceive, as a logical possibility, that I might have been Napoleon; any more than the fact that Charles Boyer can be Napoleon on the screen enables us to understand (in any serious sense) what it would be for Charles Boyer to have been Napoleon. (WILLIAMS, 1973, p. 45)

La solution que prescrit Williams au problème d'identification imaginaire consiste à supprimer l'imaginé. Le fait que je puisse m'imaginer être Napoléon n'est pas plus problématique que le fait que Charles Boyer puisse jouer le rôle de Napoléon. C'est précisément, pour Williams, l'emploi de la notion de feintise (*pretense*), qui apporterait finalement une solution au problème. Considérons l'acteur qui, engagé dans une narration imaginative où il représente Napoléon, déclame « J'ai conquis et balayé l'ancien monde par les idéaux de la Révolution ». Dans un contexte ordinaire, lorsque des déclarations contiennent le pronom personnel « je », une seule question se pose pour identifier ce « je », à savoir, « qui est celui qui parle ? ». Mais dans le contexte de l'acteur, la question est dédoublée. Car il s'agit maintenant non seulement de savoir « qui est celui qui parle » mais également de savoir « qui croit-il ou prétend-il être ? ». Pour occuper les rôles d'imaginant, d'imaginé et de Napoléon, il n'y aurait donc finalement besoin que deux personnes : l'acteur et Napoléon. Williams ne laisse donc pas de place pour une tiers personne qui incarnerait l'imaginé, et surtout pas une substance cartésienne intermédiaire. Le jeu de fiction de l'acteur n'est pas différent du jeu auquel se prête celui qui imagine être Napoléon. Cela ne consiste pas à faire comme s'il y avait une identité possible entre eux, mais à parler *à la place* de Napoléon. Selon Williams, c'est donc la feinte mimétique qui devait apporter solution aux identifications imaginaires. De ce point de vue, imaginer être Napoléon, c'est s'exclure d'emblée de la scène imaginaire. L'imaginant ne figure tout simplement pas dans le contenu de son imagination. Imaginer être Napoléon, ce n'est donc pas attribuer à l'imaginant les qualités de Napoléon. C'est à Napoléon, plutôt qu'à l'imaginant, qu'est attribuée la propriété imaginée. Dans l'opération, l'imaginé s'est évaporé.

Mais la solution prescrite par Williams reste problématique. Le faire-semblant ne règle pas le problème d'identité qui s'est présenté. Croire, ou prétendre croire, être quelqu'un d'autre, c'est concevoir des mondes où l'on est autre que soi, et non pas faire comme-si l'on ressemblait en tout point à quelqu'un d'autre. On retombe donc sur le même problème d'identité, les mondes fictionnels ne nous épargnent pas les identités impossibles. Comme Walton le souligne, la réponse de Williams est insuffisante parce qu'il s'agit bien de feindre une identité, et l'imaginé ne peut alors pas disparaître complètement du tableau. L'imagination implique toujours l'imaginant, qui ne peut quitter la scène, parce qu'il est un *point de vue sur la scène*. Walton insiste sur le fait que l'imagination est toujours ego-centrée, elle est toujours auto-référentielle.

> Even we are not the central characters, *heroes* of our dreams, daydreams, and games of make-believe, we usually have some role in them - at least that of an observer of other goings-on. Imagining an elephant in Central Park is likely to involve imagining oneself seeing an elephant in Central Park, especially if one visualizes the elephant. If one doesn't visualize it, one still probably imagines knowing about it. [...] I am inclined to think that imagining is essentially self-referential in a certain way, as intending it. (WALTON, 1990, p. 28)

Zeno Vendler déduit de l'impossibilité métaphysique d'être quelqu'un d'autre qu'il n'est tout simplement pas possible d'imaginer être identique à quelqu'un d'autre :

> [...] in imagining being someone else, the object of the exercise cannot be fancying me (Z.V.) being identical with that person. In imagining, for instance, being Ronald Reagan, I cannot be imagining the identity of Z.V. with R.R., for it is patently impossible for these two men to be one and the same, and the patently impossible cannot be imagined. (VENDLER, 1984, p. 105)

En concluant qu'une telle identification serait tout à fait inconcevable, Vendler déduit une impossibilité épistémique d'une impossibilité métaphysique. Or, qu'il s'agisse ou non d'une possibilité métaphysique, il semble néanmoins concevable que Jean puisse être quelqu'un d'autre. Nous pouvons d'ailleurs concevoir bien des impossibilités métaphysiques. Nous pouvons concevoir un scénario dans lequel Lavoisier découvrirait que la composition chimique de l'eau est XYZ plutôt que H_2O. Nous pouvons concevoir que le nez de Kovaliov se soit désolidarisé de son visage et qu'il mène une existence indépendante. Nous pouvons aussi imaginer Gregor Samsa métamorphosé en insecte. Nous pouvons imaginer que Goliadkine est confronté à une réplique pure. Et n'arrive-t-il pas que des individus réels soient pris de délires d'identification ? Dans ses dernières lettres, Nietzsche ne s'est-il pas identifié à Dionysos ?

Une impossibilité métaphysique est-elle nécessairement quelque chose d'inconcevable comme semble le soutenir Vendler ? Selon Kripke (KRIPKE, 1982), la réponse est négative. Rappelons que la distinction qu'établissent Saul Kripke et Hilary Putnam entre identification *a priori* et identité nécessaire : l'eau est nécessairement H_2O, mais la connaissance de cette identité n'est pas *a priori*. On peut, sans contradiction, ignorer certaines vérités nécessaires, donc admettre comme authentiques possibilités épistémiques des impossibilités métaphysiques. Aussi, en dépit de l'impossibilité métaphysique d'être Napoléon, cela n'empêcherait toutefois pas Jean de se représenter comme étant identique à Napoléon.

Mais selon Vendler, imaginer être Napoléon ne signifie pas stipuler une quelconque identité numérique entre ma personne et Napoléon. : « What I can do is imagine being Napoleon, i.e. having the experiences he must have had on the battlefield, or on other occasions »(VENDLER, 1979, pp. 172-173). Imaginer être Napoléon à la première personne, c'est imaginer avoir les expériences de Napoléon, Je ne peux pas imaginer être à la fois moi et Napoléon, ou moi dans Napoléon, je ne peux qu'imaginer être *à la place* de Napoléon. L'imaginé, en ce sens, est bien l'imaginant, occupant la place d'un autre, mais sans admettre une quelconque identité numérique. Il s'agit, écrit Vendler, de considérer un transfert des consciences ((VENDLER, 1979, p. 172) plutôt qu'une identité entre l'imaginé et Napoléon. Cette conception exige cependant de l'identification une reconnaissance de conscience. Si je peux m'imaginer être Napoléon, c'est, dit Vendler en substance, que je reconnais en lui une autre conscience. Cela semble donc interdire l'idée d'imaginer être quelque chose qui n'aurait pas de conscience, comme un œuf poché. Nous pourrions peut-être concevoir ceci, mais pas l'imaginer au sens de Vendler. Si je comprends bien l'idée de Vendler, l'identification correspondrait à un échange de consciences qui rappelle fortement l'expérience de pensée du prince et du savetier.

La tension entre la possibilité de s'imaginer être Napoléon et l'impossibilité métaphysique d'être Napoléon est apaisée lorsqu'on prend en compte la différence entre le fait de s'imaginer *soi-même* être Napoléon et le fait de s'imaginer Napoléon à la troisième personne. C'est, seulement

dans le premier cas que surgit une impossibilité métaphysique.

Que penser d'une situation où l'on s'imagine, à la première personne, être quelqu'un d'autre ? De nombreux auteurs s'accordent à dire que cette sorte d'imagination ne se rapporte pas à un monde dans lequel il y aurait identité numérique entre deux personnes du monde actuel. En revanche, ces mêmes auteurs divergent quant à déterminer par quoi il faut remplacer cette identité. Nous avons vu que chez Williams la notion de rôle s'y substituait. On peut considérer diversement le fait d'imaginer être Napoléon. Soit, comme Williams, on imagine un monde où existe Napoléon mais où celui qui imagine être Napoléon n'existe pas. Soit on considère que les deux y existent, mais que la relation qui se tient entre eux n'est pas la relation d'identité. C'est notamment ce que propose Richard Wollheim dans *The Thread of Life* :

> I can say – there is such an idiom – that I imagine myself being Sultan Mahomet II. But in this idiom, appearances notwithstanding, identity does not occur: I am not saying that I imagine myself being identical with Sultan Mahomet II. And this we can see from the fact that, though identity is symmetrical, 'imagining myself being Sultan Mahomet II' and 'imagining Sultan Mahomet II being me' are not synonyms. They are used to pick out two different imaginative projects. (WOLLHEIM, 1984, p. 75)

L'argument de Wollheim est donc basé sur le fait que la relation entre l'imaginant et le Sultan n'est pas symétrique, contrairement à l'identité. Il y a ici deux projets différents d'imagination. Cela indiquerait, selon Wollheim, que le contenu d'imagination ne relève pas d'une identité. Pour deux individus x et y numériquement différents, on n'imagine jamais que « $x = y$ », on imagine plutôt que « x prenne la place de y ».

Selon Walton, l'argument de Wollheim n'est pas suffisant pour rejeter la thèse que s'imaginer être Napoléon implique bien d'imaginer l'identité avec Napoléon. M'imaginer être Napoléon signifie non seulement imaginer que je suis identique à Napoléon, mais aussi que je possède certaines de ses propriétés historiques, comme avoir vaincu les forces austro-russes lors de la bataille d'Austerlitz. Inversement, imaginer Napoléon être moi signifie imaginer une identité avec Napoléon et imaginer que Napoléon possède certaines de mes propriétés, sans doute moins romanesques, comme s'évertuer à faire comprendre la notion de fraction à des collégiens. Mais dans ce cas, il semble que l'imagination en jeu ne soit plus une attitude *de se* mais *de re*.

La deuxième objection que Walton porte à Wollheim est qu'imaginer être Napoléon n'implique pas nécessairement de prendre la place de Napoléon au sens où l'on instancierait exclusivement les propriétés qui lui sont habituellement reconnues. Comme le remarque Walton, on peut toujours s'imaginer être Napoléon en train d'atterrir sur la Lune, même si cela ne fait évidemment pas partie des propriétés du personnage historique. On peut donc modifier le profil de la personne qu'on incarne en imagination en lui ajoutant certaines caractéristiques, mais aussi en lui en retirant. Mais Walton affirme plus que cela, quand il va jusqu'à dire que l'on pourrait s'imaginer être Napoléon sans instancier *aucune* de ses propriétés :

> To imagine being Napoleon it is not even *necessary* that one imagine having experiences of kinds one believes Napoleon to have had. I might imagine being Napoleon and landing on the moon, without thinking that Napoleon did land on the moon, and without imagining being

> crowed or doing or experiencing *anything else I think Napoleon actually did or experienced* [40]. (WALTON, 1990, p. 33)

La question est alors de savoir si tout ce qui est actuellement connu de Napoléon peut lui être retiré, et si l'on peut inversement lui associer de manière arbitraire quasiment n'importe quelle propriété, il devient difficile de savoir en quoi au juste il s'agit encore de Napoléon. On retombe ici sur la question du paradoxe de Chisholm et sur le recours aux entités exotiques que sont les particuliers nus.

Les modes intérieur et extérieur d'imagination *de se*

Les hypothèses précédentes achoppent sur différentes difficultés qui sont toutes liées au fait que l'identification entre Jean et Napoléon était faite en troisième personne.

Du point de vue à la première personne, imaginer être Napoléon, ce n'est pas imaginer une identité impossible, ce n'est pas feindre de croire être identique à Napoléon. C'est imaginer avoir les expériences et sensations qu'on croit que Napoléon aurait eu à telle occasion ou telle occasion.

Qu'en est-il alors des identifications en première personne ? Et comment les reconnaître ? Pour en distinguer la forme, commençons par analyser les phrases suivantes :

(ex 7) Jean imagine Napoléon.

(ex 8) Jean imagine qu'il est Napoléon.

(ex 9) Jean s'imagine être Napoléon.

(ex 10) Jean imagine être Napoléon.

La phrase **(ex 7)**, qui semble s'interpréter comme une imagination dans laquelle Jean verrait Napoléon en train de faire certaines choses, semble appartenir sans ambages au type *de re*. Tel n'est pourtant pas l'avis de Walton, selon qui toute imagination est nécessairement *de se*, la participation minimale du sujet consistant à être conscient de ce qu'il imagine, de quoi qu'il s'agisse. Pour ma part, je ne suivrai pas cette idée, il semble possible d'imaginer quelque chose sans s'impliquer à la première personne dans l'imagination.

La phrase **(ex 8)** peut s'interpréter de manière transparente, c'est-à-dire de manière *de re*. Le pronom anaphorique « il » indique que Jean imagine qu'un individu particulier, qui n'est autre que lui-même, est Napoléon. De ce point de vue, la phrase n'indique pas la manière dont Jean pense à cet individu, il pourrait aussi bien le faire à la troisième personne. Mais d'un autre côté, cet énoncé pourrait aussi s'interpréter de manière opaque. Il reste que le contexte d'interprétation pourrait s'avérer ambigu : dans une situation, Jean s'imaginerait lui-même en-tant-que-lui-même, être Napoléon, comme s'il disait plus simplement « je suis Napoléon » (imagination *de se*). Mais dans une autre situation, il pourrait aussi s'imaginer être Napoléon parce qu'il se verrait, dans un miroir, habillé d'un chapeau bicorne, sans se reconnaître et dire « ce type est Napoléon ! ». Jean s'imaginerait lui-même être Napoléon mais cette fois-ci, il le ferait de manière impersonnelle (imagination *de*

[40] C'est moi qui souligne.

re). La manière dont Jean imagine être Napoléon est ainsi délivrée par le mode de présentation du contenu de pensée. Selon un certain mode d'identification, qu'on pourrait qualifier d'intérieur, il occupe la place de Napoléon, il voit à travers ses yeux, et dans un second mode d'identification, extérieur, il reste à distance de lui mais s'y identifie par méprise, sans se reconnaître lui-même.

La phrase **(ex 9)** semble avoir le mérite d'éviter de telles ambiguïtés parce qu'elle traduit plus spécifiquement le caractère réflexif de la pensée de Jean, en indiquant que le mode de présentation relève d'une pensée *de se*. Peter Geach décrit de telles phrases réflexives comme « an *oratio obliqua* proxy for the first person pronoun of *oratio recta* » (P. GEACH, 1957, p. 129). Ces « procurations » sont qualifiées quelques années plus tard par Castañeda de « quasi-indicateurs » (*quasi-indicators*) (CASTAÑEDA, 1999), autrement dit de termes qui imposent non seulement une interprétation opaque des expressions où ils apparaissent, mais qui imposent aussi que cette interprétation soit telle que ce à quoi elle est attribuée pense à son objet du point de vue de la première personne. Cependant, il reste que dans certains contextes, le pronom réflexif pourrait cette fois encore ne pas être interprété *de se*, comme dans l'exemple précédent du miroir. Dans ce genre de cas, l'expression « se » ne serait pas utilisée comme un authentique quasi-indicateur logique. Cette fois encore, c'est le mode de présentation du contenu de pensée, et non la seule grammaire, qui permettrait de lever l'ambiguïté.

Il reste à examiner la phrase **(ex 10)**. Bien qu'elle ne contienne aucun pronom réflexif de manière explicite, l'expression mobilise toutefois un élément anaphorique implicite, à savoir, le sujet du verbe à l'infinitif. Le terme, occulté dans la phrase, semble se comporter comme un invisible quasi-indicateur puisque cette fois, la manière dont Jean imagine qu'il est Napoléon est inévitablement présentée comme étant à la première personne. Nous pourrions retraduire ainsi : « Jean s'imagine, en tant que Jean, être Napoléon ». Ici, l'imagination de Jean paraît se rapporter inévitablement à un point de vue personnel et interne.

Jusqu'à présent, nous avons admis qu'une imagination *de se* d'être un autre est toujours intérieure. Mais si je m'imagine, à la première personne, être Napoléon, est-ce alors à dire que je m'imagine nécessairement de manière intérieure ? Puis-je m'imaginer être un autre *en dehors de lui* ? Est-ce bien légitime ?

Après avoir successivement considéré le cas où un agent imagine une scène dans laquelle il n'intervient pas, et le cas où un agent imagine, intérieurement, vivre une scène, Bernard Williams constate qu'il est encore possible d'envisager un troisième cas, celui où l'agent s'imagine lui-même faire des choses, mais de l'extérieur. Mais Williams range cette forme d'imagination dans la première catégorie, celle où il le sujet ne prend pas part aux faits décrits :

> [...] the expression 'imagining myself doing, etc.' could cover also a third possibility for imagery, which constitutes not really a distinct third kind, but a special application of the first: namely that of visualising from the outside a figure who is myself doing the things in question. (WILLIAMS, 1973, p. 38)

L'imagination de se voir agir de l'extérieur est ainsi ramenée par Williams à une forme d'imagination à la troisième personne, et non à un cas d'imagination *de se*. Une imagination *de se* est donc, pour William, toujours descriptible dans les termes suivants :

> I visualise a world in which I am acting, moving around, seeing things, and so forth - a form

> of imagery involving, very often, kinaesthetic imagery of various sorts. (WILLIAMS, 1973, p. 38)

Aussi, une imagination *de se* d'être un autre serait nécessairement une imagination intérieure.

Selon Christopher Peacocke, la caractéristique principale de l'imagination est qu'elle est toujours imagination, d'un point de vue intérieur, d'être dans un certain état de conscience : « To imagine something is always to imagine, *from the inside*, being in some conscious state » (PEACOCKE, 1985, p. 21). Dès lors, l'acte d'imaginer implique toujours un degré minimal de participation perceptive. Nous ne pouvons imaginer voir, entendre, toucher, goûter, ou sentir que quelque chose que nous pourrions voir, entendre, toucher, goûter, ou sentir réellement. Toute imagination se ramènera ainsi à des expériences de sensations ressenties d'un point de vue personnel. En revanche, Walton ne souscrit pas à l'idée qu'une imagination *de se* est nécessairement une imagination d'expérience perceptive. Il donne l'exemple de l'imagination d'un lien de parenté avec Napoléon, qui est bien, d'après lui, une imagination *de se*, qui est interne en ce que l'imaginant a conscience de cette parenté imaginaire, mais par contre, qui n'implique pas, à proprement parler, d'expérience sensitive. En s'éloignant d'une forme de perception affaiblie, la notion d'imagination que mobilise ici Walton se rapproche plus de la conception [41]. Je peux concevoir un lien de parenté sans visualiser les objets afférents.

Est-il vraiment nécessaire qu'une imagination *de se* soit intérieure ? Que serait une identification imaginaire *de se* extérieure ? S'agit-il d'une contradiction dans les termes ? Ou y a-t-il, au contraire, des modes de présentation qui correspondent à de telles pensées ? Que signifierait au juste « s'imaginer faire des choses vu de l'extérieur » ?

Dans « Vicarious Experience », Vendler constate que nous pouvons toujours imaginer des scènes nous concernant de deux façons différentes. Celles-ci sont conduites par deux type d'expressions. Nous voici au sommet d'une falaise qui surplombe l'océan. Vendler nous demande alternativement :

(ex 11) Imaginez nager dans une eau tourmentée et froide.

(ex 12) Imaginez-vous nager dans une eau tourmentée et froide.

Selon Vendler, les deux énoncés expriment deux formes différentes de mise en scène dans l'imagination. La première phrase tend à nous faire *vivre la situation donnée*, à agir dans la situation décrite, autrement dit, à prendre la place du nageur de manière interne. Dans ce cas, nous imaginerions ressentir les percepts qui en dépendent, par exemple la sensation de froid sur notre peau ou le goût salé de l'eau sur nos lèvres. Dans cette forme d'imagination, nous ne jouons jamais plus qu'un seul rôle, celui du nageur. Dans ce type de situation, « s'imaginer être Napoléon » voudrait dire avoir les perceptions qu'aurait Napoléon, comme la douleur lancinante provoquée par un ulcère à l'estomac et comme la variation d'intensité des détonations à l'approche du front de bataille. Nous retrouvons ici le mode de présentation décrit par Williams dans les imaginations *de se*.

[41] Au sujet de l'absence de définition de l'imagination chez Walton, voir la thèse de Guillaume Schuppert (SCHUPPERT, 2019).

L'imagination interne est-elle nécessairement une imagination *de se* ? Selon Walton, c'est bien le cas. Imaginer voir un rhinocéros, c'est imaginer le voir personnellement, et non imaginer une vision « neutre » de rhinocéros. Imaginer voir à travers les yeux d'un autre, ce n'est plus imaginer *de se*. Imaginer Napoléon voir un rhinocéros, ce n'est pas imaginer voir un rhinocéros au sens où on l'entend habituellement. Il y a une participation forte de l'imaginant dans le monde de l'imagination :

> One who imagines Napoleon's seeing a rhinoceros, or imagines a seeing of a rhinoceros without imagining whose seeing it is, does not thereby 'imagine seeing a rhinoceros' as this phrase is ordinarily understood. [...] 'Imagine' fits a common pattern in this regard. To remember giving a speech is to remember oneself giving one, not merely to remember an instance of the speech being given. (WALTON, 1990, p. 31)

Vendler introduit encore un second mode de présentation d'imagination *de se*, donné par la deuxième expression, qui tendrait plutôt, selon lui, à nous faire imaginer la même situation en nous « voyant nous-mêmes nager » de l'extérieur. Elle pourrait se traduire plus proprement par « imaginez-vous en train de vous voir nager ». Un tel mode de présentation d'imagination *de se* est aussi pris en compte par Walton. Admettons que Jean s'imagine jouer un important match de baseball. Il pourrait imaginer cela de l'intérieur, il pourrait imaginer ressentir dans ses mains le choc dû à la frappe de la batte contre la balle, entendre la foule l'acclamer depuis les gradins. Mais il pourrait tout aussi bien s'imaginer de manière externe frapper la balle, du point de vue d'un spectateur assis dans les gradins. S'il imagine la scène de cette manière, son image inclut Jean frappant la balle sur le terrain. Or là où Williams considère ce cas de figure comme une imagination nécessairement *de re*, Walton s'en défend :

> Imaginings *de se* are imaginings about the imaginer himself. But I will suggest in the following section that the imaginer *de se* need not be a *de re* object of his imagining, and that it may not be appropriate to describe him as imagining, "of himself, that..." By "objects" of imaginings I will mean *de se* as well as *de re* objects. (WALTON, 1990, p. 25 n9)

Mais plus encore, Walton considère, non sans une forme de provocation, que c'est plutôt le mode de présentation externe qui convient le mieux pour décrire l'imagination *de se* : « This imagining is, I believe, best classified as *de se*. » (WALTON, 1990, p. 30).

Selon le mode extérieur d'imagination, celui qui s'imagine faire quelque chose le fait en imaginant voir et entendre la situation donnée à imaginer, tandis que dans le mode intérieur, il se figure ce que cela serait d'être dans cette situation et d'avoir les perceptions appropriées. N'en déplaise à Walton, la question subsiste de savoir si l'imagination extérieure de soi est bien une authentique imagination *de se*.

C'est précisément ce que conteste François Recanati. Selon lui, la construction grammaticale « imaginer être Napoléon » ressortit nécessairement à une pensée *de se*. Or comme nous l'avons vu, Walton lui associe les deux perspectives externe et interne. Certes, il y a bien une façon de dire que l'imagination externe est encore à propos de nous. Mais à suivre Recanati (RECANATI, 2007a, p. 13), ces imaginés externes ne sont pas d'authentiques pensées *de se*, mais des pensées personnelles *de re*. La scène de *baseball* décrite par Walton relève pour Recanati d'une pensée *de re* qui ne répond pas à la forme « imaginer jouer au *baseball* ».

C'est cependant contestable, comme l'indique d'ailleurs Recanati lui-même :

> In the example at stake, however, the identification of Gregory as the object his imagining is about is fixed directly by Gregory's imaginative project, in a stipulative manner ; it is not determined by the causal antecedents of the imaginative state. If it were so determined (e.g. if Gregory's project was to imagine that man in the mirror hitting a home run), then it might indeed give rise to Frege cases. I conclude that the sort of *de se* imagining which the Gregory example illustrates is not an implicit *de se* imagining. It is (conditionally) immune to error through misidentification because of a general property of the imagination, but it involves a form of identification nonetheless — a form of identification that is immune to error only because it is stipulative rather than anaphorically inherited. (RECANATI, 2007a, p. 16)

Une imagination extérieure peut ainsi correspondre à une mode de présentation d'une imagination *de se*, en étant immunisée contre l'erreur par stipulation. Comme l'indique aussi Wittgenstein dans *Le Cahier Bleu*, lorsqu'on imagine que la coupole du King's College brûle, il n'y a pas de sens à se demander en quoi l'on sait que c'est cette couple plutôt qu'une autre qui brûle (WITTGENSTEIN, 1965, p. 102).

L'imagination est-elle ainsi parfaitement immunisée contre les erreurs ? Selon Walton, ce n'est pas le cas. Certains doutes concernant la pertinence de l'identification pourraient quand même subsister. Walton considère l'exemple suivant. Admettons qu'une personne se trompe sur l'identification de deux universités, substituant l'une à l'autre. À chaque fois qu'elle aurait affaire au Birbek College, elle croirait à tort qu'il s'agit du King's College. Inversement, elle prendrait le Birbek College pour le King's College. Dans ce cas, à chaque fois que son intention serait d'imaginer le King's College en feu, le contenu de son imagination porterait en réalité sur le Birbek College. Et inversement, si elle avait devant elle le King's College, elle pourrait imaginer cette université en feu sans réaliser que le contenu de son imagination porte précisément, cette fois, sur le King's College (WALTON, 1990, pp. 29-30). Contre la caractérisation de Wittgenstein, cela montrerait que l'imagination n'est pas toujours sauve d'erreurs d'identification.

Mais quel genre d'erreur pourrions-nous commettre dans le cas d'une imagination de *de se* ? Walton souligne qu'il n'y a là pas d'erreur possible : « There is no room for doubt that [Gregory] is himself the player who hits the home run in his imagination » (WALTON, 1990, p. 30). Walton soutient l'immunité contre l'erreur d'identification dans l'imagination *de se* :

> No conceivable evidence about the causal antecedents of [Gregory's] imaginative experiences could make him question that he is himself the player who hits the home run in his imagination. There seems to be no story of the relevant sort to be told about how it got to be Gregory whom he imagines[...] Yet his imagining is not from the inside. It is as though he is watching someone else hit the ball and round the bases, despite his unshakable realization that he is himself that person. (WALTON, 1990, p. 30-31)

Une identification imaginaire *de se* ne peut donc jamais être erronée.

7.5.3 La question du double point de vue

Si l'on peut rejeter les identités imaginaires *de re* comme des impossibilités logiques, la question reste de savoir s'il y a un sens quelconque à admettre une identification imaginaire *de se*.

Suivons ici Walton : admettons que Joyce imagine être Napoléon. Si l'on en ressent le besoin, la meilleure manière d'éviter une pensée métaphysiquement impossible recourant à l'identité *de re* de Joyce et de Napoléon est, selon Walton, une sorte de double processus d'imagination. Admettons que Joyce imagine, en première personne, voir un rhinocéros. Elle pourrait pareillement imaginer Napoléon en train de voir un Rhinocéros :

> Let us say that she illustrates for herself what she imagines Napoleon to experience, by imagining experiencing it herself. (WALTON, 1990, p. 34)

Imaginer être Napoléon, c'est toujours imaginer Napoléon conscient de quelque chose par le fait de parvenir à s'imaginer soi-même le même contenu. Joyce et Napoléon restent donc deux individus distincts, liés par un acte similaire d'imagination, mais pas par une identité numérique. Simplement, tous deux imagineraient la même chose. L'imaginant reste donc toujours présent dans la scène d'imagination, il n'est pas effacé au profit d'un autre individu :

> Joyce does not imagine an identity between herseﬂ and Napoleon. But she does imagine both herself and Napoleon, and these two imaginings, though distinct, are significantly linked. (WALTON, 1990, p. 34)

Walton compare cette stratégie à la mise en scène d'Akira Kurosawa dans *Rashomon* (KUROSAWA, 1950), où plusieurs personnages relatent l'assassinat d'un samouraï par un bandit. Dans son film, Kurosawa représente les événements non pas comme ils sont fictionnellement vus par les personnages, mais comme ceux-ci les décrivent (quitte à mentir). Chacun des témoins commence à raconter son histoire, qui est ensuite mise en image en fonction de ce qu'ils disent. Au final, les témoignages s'avèrent contradictoires : dans son récit, le bandit affirme être le meurtrier du samouraï, dans le récit de la femme, c'est elle qui tue son mari. Mais les séquences qui illustrent ce que racontent les personnages ne sont pas représentées selon le point de vue des personnages, nous ne voyons pas à travers leurs yeux, mais à travers ce que les personnages prescrivent d'imaginer. Le spectateur imagine à travers l'imagination du personnage. Par comparaison, il n'en va pas ici comme dans une scène du film *Halloween* de John Carpenter (CARPENTER, 1978) où un meurtre y est représenté à travers les yeux de l'assassin, la caméra forçant ainsi le spectateur à participer fictionnellement à l'assassinat à travers l'acte de regarder la scène par une identification subjective à l'assassin. Chez Kurosawa au contraire, le point de vue n'est pas ancré au personnage qui raconte la scène, et il ne semble donc pas y avoir de volonté d'identifier fictionnellement le spectateur au personnage. Le spectateur intervient donc dans l'histoire en demeurant distinct, du personnage, mais lié à lui par un acte d'imagination similaire.

A ce sujet, Vendler évoque un réel problème en conclusion de son article : si l'identification imaginaire passe par un point de vue, les imaginations externes demandent un double point de vue :

> For, remember, I can easily imagine myself swimming in the ocean : a scrawny body tossed about by the waves. And I imagined it as seen from the cliff. Seen by whom ? By myself of

> course. But if I imagine myself, complete with eyes of course, in the water, how can I, then, imagine seeing myself from the cliff ? Do I imagine doing it by means of alien eyes, no eyes, or what ? (VENDLER, 1979, p. 173)

Vendler se contente d'affirmer modestement que cette difficulté sort du cadre de son article, et n'y répond donc pas. Quels éléments pouvons-nous, de notre côté, apporter à cette question ?

Constatons que certains subterfuges peuvent être employés pour expliquer des situations qui semblent, du point de vue intérieur, apparemment inimaginables. Comment imaginer de manière interne avoir les yeux verts, si tel n'est pas actuellement le cas ? On peut par exemple imaginer voir son reflet dans un miroir. Les imaginations autoscopiques ne sont pas d'ailleurs pas limités aux seules imaginations de reflets spéculaires. Bernard Williams montre que le recours à la photographie permet d'alterner une imagination sous le point de vue intérieure avec une imagination extérieure, sans quitter pour autant le point de vue à la première personne :

> if I am prone to fantasies of being a world champion racing driver, this could involve kinaesthetic imagery of tension, hands clasped on the steering-wheel, and visualisation of wet tarmac as seen through an oil-spattered windscreen, and so forth; and, also at some different point, some visual image of myself, as though in a newspaper photograph, having a garland hung around my neck. (WILLIAMS, 1973, p. 38)

Le miroir et la photographie permettent ainsi d'alterner imagination extérieure et imagination intérieure de soi sans s'acquitter du mode d'imagination à la première personne.

Walton généralise la thèse d'un double rôle occupé par le récepteur d'une fiction, qui doit toujours adopter un double point de vue *(dual standpoint)* en séparant ce qui relève des faits de ce qui relève de la fiction. S'il n'en était pas ainsi, si le récepteur de la fiction n'avait qu'un seul point de vue, il faudrait alors que ce dernier soit celui de la réalité ou de la fiction. S'il s'agissait du point de vue de la réalité, le récepteur ne devrait alors pas éprouver d'empathie pour les personnages de l'histoire. S'il s'agissait du point de vue de la fiction, il pourrait alors vouloir intervenir dans l'histoire. Le récepteur d'une fiction doit donc à la fois avoir un point de vue sur la réalité en considérant que l'histoire racontée est fictive, et un point de vue sur la fiction en considérant aussi les choses depuis ce monde là. La possibilité d'une double présence acteur-spectateur dans l'imagination *de se* est aussi soulignée par Tzvetan Todorov dans le lien d'identification du lecteur d'un récit fantastique : « Le fantastique [...] ne dure que le temps d'une hésitation commune au lecteur et au personnage, qui doivent décider si ce qu'ils perçoivent relève ou non de la "réalité" telle qu'elle existe pour l'opinion commune [...]. » (DUCROT & TODOROV, 1972, p. 46). Cette caractérisation des récits fantastiques montre que leur construction prévoit une intégration du lecteur à l'histoire, puisqu'il doit lui-même interpréter les événements du monde de la fiction. Quelle est la nature de cette intégration ? Il s'agit d'une identification par empathie selon différents degrés, dont le le plus faible consiste à tenir le rôle de spectateur du récit.

Walton propose encore une autre illustration. Dans le tableau de Rubens (voir figure 7.6), *La Toilette de Vénus*, la déesse de l'amour est vue de dos s'observant dans un miroir que soutient Cupidon. Mais le miroir, qui fait face au spectateur du tableau, ne renvoie pas le reflet de la scène comme il devrait le faire. Il présente en effet le reflet du visage de la déesse de face, c'est-à-dire, tel qu'il apparaîtrait à Vénus depuis son propre point de vue, et à elle seulement. Le miroir, avec



FIGURE 7.6 : Rubens, *La toilette de Vénus*, 1614-1615

sa réflection en anamorphose, pourrait servir de prétexte pour introduire un niveau secondaire, métaleptique, dans la lecture du tableau. Mais ce n'est pas l'opinion de Walton qui préfère dire qu'il y a deux tableaux différents dans le même tout que constitue la toile de Rubens :

> The portrayal of the mirror can be regarded as a separate depiction illustrating what, fictionnaly, in the world of the whole, Venus sees in it. (WALTON, 1990, p. 346)

Walton souligne ainsi que la toile représente, non un seul, mais deux mondes fictionnels, le premier qui est celui où il est fictionnel que Vénus se mire entourée de deux personnages, et le second où il est fictionnel que le spectateur soit identifié à Vénus se voyant dans un miroir.

Walton souligne que dans l'imagination *de se* externe, l'imaginant peut jouer deux rôles à la fois, à la fois comme spectateur et à la fois comme acteur. Une personne pourrait s'imaginer elle-même, de l'extérieur, taper avec une batte dans la balle lors d'un tournoi de *base ball* :

> It is as though he is watching someone else hit the ball and round the bases, despite his unshakable realization that he is himself that person. (He may imagine himself from the inside watching the game from the stands. And of course his imagining, from the stands, hitting the run may alternate with imagining this from the inside.) (WALTON, 1990, p. 31)

Ainsi, le type d'imagination de l'extérieur demande d'imaginer *de se voir* entrain de frapper la batte, alors que le type d'imagination intérieure demande d'imaginer de frapper la batte.

Nous pouvons alors nous demander si cette deuxième manière d'imaginer exige que l'imaginant y joue deux rôles, celui de spectateur, et celui d'acteur. Comme on l'a vu plus haut, la solution consistant à imaginer se voir à travers les yeux d'un autre que soi est clairement écartée par Walton, puisque selon lui l'imaginant est toujours présent d'une manière ou d'une autre dans la scène imaginée. Devrions-nous imaginer être notre propre spectateur et pour cela, avoir un corps dédoublé, dont l'un, tapant une balle ou nageant, est projeté à distance de l'emplacement occupé par l'autre, spectateur de la scène ? Dans une note de bas de page éclairant cette duplication étonnante du personnage à la fois debout sur le terrain de jeu et à la fois assis dans les gradins, Walton déclare :

> How it is that he both spectator and player in the imaginary game ? Perhaps he imagines being the spectator and imagines being the player, without imagining being both at once. (Walton, 1990, p. 31 nb 16)

L'idée générale qu'on peut tirer de Walton est de décentrer la connaissance que p en utilisant autrui comme nouveau point de perspective. Cela consiste à utiliser autrui comme un double

de soi, un corps de remplacement pour une attitude explicitement *de se*. Mon imagination peut alors recréer des états épistémiques, à condition qu'ils soient pensés comme occupés par d'autres individus. Comme l'écrit Jérôme Dokic :

> What I imagine that p, I am in a mental state analogous to that in I which *I or somebedy else* would be if they knew, or a least were in a position to know, that p. (DOKIC, 2008, p. 110)

Pour certains philosophes qui défendent des perspectives simulationnistes, la capacité d'imaginer de l'intérieur n'est pas une perception imaginée mais relève en fait d'une simulation de perception. Quand j'imagine voir une pomme, je n'imagine pas le processus mental de visualisation d'une pomme, mais je simule le déroulement de ce processus. Autrement dit, je me mets dans la position de celui qui ferait l'expérience perceptive de cette pomme. La simulation suppose plus que la simple conception. Alvin Goldman développe la notion cognitive de *enactment imagination* (E-imagination) comme :

> a matter of creating or trying to create in one's own mind a selected mental state, or at least a rough facsimile of such a state, through the faculty of imagination. (GOLDMAN, 2006a, p. 42)

L'E-imagination ne relève pas de la simple conception mais implique un état mental bien plus participatif :

> When I imagine feeling elated, I do not merely suppose that I am elated; rather, I enact, or try to enact, elation itself. Thus, we might call this type of imagination "enactment imagination". (GOLDMAN, 2006b, p. 42)

De ce point de vue, Recanati souligne que l'imagination n'est pas un type particulier de perception doté d'un élément supplémentaire de faire-semblant : « There is a much bigger difference between the basic experiential modes (perception and memory) and the imagination. » (RECANATI, 2007b, p. 206) Alors que la perception est un état nécessairement réflexif, c'est-à-dire un état qui implique d'importer le soi, dans le processus d'imagination, Recanati soutient que l'imagination ne nécessite pas l'importation de l'imaginant dans la représentation. En ce sens, Recanati rapproche les identifications imaginaires des quasi-souvenirs qui sont des souvenirs certes subjectifs, mais impersonnels (au sens où ils ne sont pas nécessairement des souvenirs de la personne). Il appelle ainsi imagination *quasi-de se* le type d'imagination qui est mobilisée par quelqu'un qui s'imagine, à la première personne, être une autre personne

> I will, therefore, coin the term 'quasi-de se' to refer to the first person point of view type of thought one entertains when one imagines, say, being Napoleon. The type of imagining at stake is clearly first personal, yet the imaginer's self is not involved [...]. The properties that are imaginatively represented are not ascribed to the subject who imagines them, but to the person whose point of view she espouses. (RECANATI, 2007b, pp. 206-207)

Néanmoins, je ne pense pas qu'il soit nécessaire de considérer les identifications imaginaires comme de fausses pensées *de se*. La théorie de l'imagination de Recanati repose apparemment sur le présupposé du caractère métaphysique du soi. Mais si nous acceptons l'idée que le soi n'est

qu'épistémique, nous ne sommes pas contraints alors de reconnaître une là impossibilité nécessitant de souscrire à l'idée d'une quasi-imagination. Nous ne déplaçons pas, dans l'imagination, des particuliers nus. Nous déplaçons un point de vue sur le monde.

Retenons que l'idée d'un décentrement du point de vue permet de donner du sens aux identifications imaginaires sans rejeter toutefois l'idée que de telles pensées soient véritablement *de se*. La conception de Walton reste elle aussi, sans doute, tributaire d'un lourd réquisit cartésien, à savoir, l'existence d'egos, comme cela apparaît d'ailleurs clairement dans cette déclaration :

> The notion of the self that figures in imaginings *de se* need not be a very rich or full one [...] The self whom I imagine to be seeing a rhinoceros may be a 'bare Cartesian I'. (WALTON, 1990, p. 31-32)

L'imagination, chez Walton, exige d'implanter les points de vue sur des entités, quoique ces entités soient dévidées de leurs propriétés. Et c'est là qu'apparaissent alors les problèmes de concurrence ontologique. Plutôt que de souscrire à des entités nues, on peut aussi proposer de désancrer les points de vus des entités du domaine.

Peut-être pouvons envisager des solutions n'impliquant pas de duplication corporelle entre le spectateur et l'acteur. Nous pouvons envisager par exemple le cas où nous serions le spectateur d'un film où nous nous verrions nageant, ce qui se rapporte à une forme d'autoscopie sans nécessité de dédoublement ontologique. En outre, les deux modes de représentations d'imagination *de se* intérieur et extérieur sont bien connus des joueurs de jeux vidéos. Dans la vue à la troisième personne, le personnage joué est présenté comme s'il était vu par un autre. Il est manipulé par le joueur comme l'on manipule une marionnette présente devant soi. Pour des raisons techniques, les premiers jeux vidéos ont favorisé ce mode de représentation. Dans la vue à la première personne, le personnage est par contre absent de l'écran et l'environnement du jeu est vu comme le verrait le joueur s'il incarnait le personnage. Certains jeux ont développé des techniques narratives employant des basculements successifs de points de vue. Fanny George constate par exemple que le jeu Silent Hill 4 :

> [...] présente la particularité de forcer l'alternance entre une caméra à la 1re et à la 3e personne, la première étant utilisée dans l'appartement, la seconde dans le reste du jeu. La caméra à la première personne met l'accent sur la subjectivité du personnage et invite le joueur à se projeter en lui, tandis que la caméra à la 3e personne l'en distancie. Au cours du jeu, le retour régulier à la caméra subjective, coïncidant avec le réveil d'Henry dans son lit, les yeux dirigés au plafond, suggère une fusion du joueur avec personnage par le motif du point de vue. L'autoscopie renforce ainsi le sentiment d'immersion. (GEORGES, 2013, p. 55)

Dans les jeux multi-joueurs en ligne, l'avatar participe non seulement à la représentation symbolique en première personne, mais contribue aussi à matérialiser la présence du joueur, en troisième personne, aux autres joueurs. On peut donc, en première approximation, définir l'avatar comme un moyen de représentation et de mise en présence du joueur dans un espace public de communication. L'avatar offre en quelque sorte un point de vue virtuel pour situer l'individu dans l'espace du jeu.

Nos imaginations *de se* semblent bien mobiliser ces deux modes, et cela, parfois de manière alternée. Les modes d'identification intérieur et extérieur peuvent finalement être respectivement



FIGURE 7.7 : Masaccio, *Le paiement du tribut*, 1424-1427

rapprochés des figures-filtres (par incarnation) et des figures-échos (spéculaires). Il pourrait en être comme dans ces représentations picturales médiévales où un même personnage apparaît à différents endroits, comme sur la fresque *Le paiement du tribut* (1424-1427) de Masaccio (voir figure 7.7) que nous avons déjà évoqué. La lecture de cette représentation n'exige pas de croire que le personnage soit simultanément présent à divers emplacements, mais que le tableau représente sur une même surface différents moments du temps. Nous pourrions aussi le décrire comme une manière d'observer différents mondes possibles centrés sur différents points de vue qui sont repérés par le personnage récurrent, mais sans admettre une duplication ontologique. Le dédoublement ne semble donc pas nécessaire.

Aucune duplication ontologique du spectateur et de l'acteur n'est nécessaire, parce que le spectateur peut correspondre à un point de vue sans être assimilé à une entité. Cette idée peut-être modélisée en sémantique formelle à l'aide des mondes centrés où le spectateur est défini non pas comme un objet du domaine, une entité, mais comme le « centre » du monde [42]. Dans la version purement perspectiviste des mondes possibles centrés, le centre est un ensemble ordonné de coordonnées spatio-temporelles [43].

Dans un passage souvent commenté, John Perry (PERRY, 1979) imagine une situation dans laquelle il suit la piste, dans un magasin, d'un client qui s'est muni d'un paquet de sucre troué, pour se rendre finalement compte qu'il s'agit de lui-même. Les mondes épistémiques de Perry avant et après sa découverte se correspondent très exactement, les propositions qui constituent ces mondes étant exactement les mêmes. Il en va ici de même que dans l'exemple de Kaplan du pantalon en feu précédemment cité. S'il y a bien une différence, qui se traduit d'ailleurs dans le fait que Perry adapte un nouveau comportement en abandonnant sa poursuite, elle n'est pas à rechercher quelque part dans les propositions qui décrivent respectivement ces deux ensembles de mondes possibles. La différence est ailleurs, à savoir, dans la localisation du sujet dans les mondes possibles qui décrivent l'état des connaissances du sujet en question.

Dans la structure des mondes centrés, le spectateur d'une scène imaginée peut ou non coïncider avec un acteur vu comme un objet du domaine. Une identification imaginaire *de se* est centrée sur le personnage imaginé, ce qui n'exige pas de véhiculer un ego embarrassant dont on ne sait plus s'il faut ou non l'identifier à Napoléon. Mais elle peut également être centrée sur un point de vue distinct, une autre entité du domaine. « Je » n'est pas un terme qui réfère à une entité, mais un terme qui focalise le monde depuis un point de vue particulier.

On peut alors adapter les relations de croyance-feinte aux mondes possibles centrés. En fai-

[42] Les discussions sur les mondes centrés ont été motivées par les problèmes d'indexicalité remontant aux quasi-indicateurs de Castañeda. Voir (PERRY, 1979), (LEWIS, 1979) et (STALNAKER, 1981).

[43] Dans la version de David Lewis des mondes centrés, le centre est constitué par la donnée d'une personne et d'une coordonnée temporelle.

sant comme si l'on était Napoléon, on ne devient pas Napoléon au sens où l'on deviendrait une autre personne, mais on en occupe, dans les mondes épistémiques afférents, la position spatio-temporelle.

En considérant l'imagination depuis ce point de vue perspectiviste, nous n'avons plus besoin de supposer l'existence de particuliers nus, ni de quelconques doubles ontologiques afin d'expliquer nos identifications imaginaires. Nous pouvons nous passer d'egos cartésiens tout en admettant que l'identité narrative puisse se construire par des jeux de simulation.

7.6 Conclusion

Il est temps de synthétiser nos résultats. Nous avons vu que la théorie substantialiste de l'identité personnelle est confrontée à divers problèmes. D'un côté, elle s'enferre dans des paradoxes qui nécessitent l'introduction de doubles ontologiques comme les particuliers nus, et d'un autre côté, elle ne rend pas compte du caractère construit de l'identité personnelle. En refusant la duplication du corps par la substance inétendue, nous aurions pu rabattre le moi directement sur le corps, mais nous avons vu que cette stratégie menait à une impasse : la personne n'est pas concomitante au corps, elle apparaît après celle du corps, et peut disparaître sans que le corps cesse.

Si aucune des solutions substantialistes ne paraît acceptable, on peut alors se tourner vers la solution fictionnelle. Au terme de chapitre, nous pouvons rendre compte de l'accusation de paradoxe que les substantialistes adressent à Hume quand celui-ci soutient qu'un quelconque examen intérieur n'aboutira jamais à aucun moi, tout en maintenant que nous sommes bien capables de nous identifier, de parler de nous à la première personne. Là où le dualisme cartésien envisage le moi comme un double ontologique du corps, nous avons défendu l'idée que le moi est une construction imaginaire bâtie sur des doubles épistémiques. S'il y a bien un moi qui se rapporte à « je », ce moi est à concevoir plutôt comme récit, comme personnage.

Le récit de soi constitue un rempart contre l'écroulement du moi dans l'indifférenciation objectuelle, dans l'impersonnalité, dans la phagocytose du corps au monde, dans l'hénadisme absolu dont semblent témoigner certaines maladies mentales comme la schizophrénie où l'agent n'est plus capable de démarquer les frontières de son moi. C'est que la personne n'est pas un double ontologique qui nous serait comme donné en supplément du corps, et dont nous serions toujours assurés de profiter sans effort par une forme de contact direct. Le moi nécessite au contraire une performance, la création de doubles épistémiques.

Bien sûr, toute performance narrative ne vaut pas pour un moi, il ne suffit pas de se prendre pour Napoléon pour être Napoléon. La performance narrative est régie par des relations interpersonnelles. Comme nous l'avons dit, les processus d'auto-identification par duplications cognitives sont d'ailleurs bien connus des sociologues et des psychologues depuis le début du XXe siècle.

D'une part, les expériences autour des reflets spéculaires de Zazzo montrent que la construction de l'image de soi précède la reconnaissance de ressemblances qualitatives. L'auto-identification par image spéculaire interposée nécessite donc à l'agent de se reconnaître dans son double virtuel là où il présumait d'emblée deux personnes. On peut d'ailleurs se demander si la thèse dualiste n'est pas sans rapport avec une sorte de réminiscence de la duplication ontologique que l'enfant déduit de son reflet dans le miroir avant de parvenir à s'auto-identifier.

D'autre part, nous avons vu l'importance que les doubles épistémiques jouent dans l'élaboration de l'identité sociologique. Celle-ci est construite par l'assimilation de rôles contribuant à façonner un personnage auquel s'identifier, l'apprentissage des rôles passant par des jeux d'imitation et de simulation comme décrits par Mead.

Nous avons vu que s'imaginer être Napoléon, impose, en terme d'identité personnelle, une concurrence des entités en présence dans l'imagination afférente qui débouche sur une impossibilité métaphysique. On peut toujours prétexter, dira-t-on, que ce qui est métaphysiquement impossible n'est pas logiquement impossible, et que cela n'empêche donc pas cette imagination. Mais ce que les discussions précédentes tendent à indiquer, c'est que nous n'avons pas besoin de cette concurrence ontologique, et que nous pouvons y échapper en ne souscrivant pas à une ontologie substantialiste du moi.

Se concevoir un soi narratif passe ainsi par l'imagination d'être un autre. Cette capacité dépend de deux instances. D'une part l'imitation d'autrui qui fait l'enjeu principal des jeux de rôle décrits par Mead (*play*), dont l'affaire est de *ressembler à quelqu'un d'autre*, de devenir un autre en mimant son comportement, ses paroles, etc., bref, en imitant la personne physique, processus qui fait progressivement émerger la capacité à se localiser dans le monde, et où intervient certainement l'apprentissage de la vision autoscopique, et plus généralement le phénomène de la figure-écho. Mais d'autre part aussi, la simulation qui succède à l'imitation, dans les jeux ludiques (*game*) en ajoutant la contrainte de règles qui nécessitent de se mettre *à la place de l'autre*, mode de la figure-filtre. Nous avons vu que ce type d'imagination, par simulation, nécessitait plus que la simple conception, puisque le changement de point de vue dans l'identification *de se* impose de reconstruire mentalement les sensations, les sentiments et les pensées d'une autre personne en se dégageant cette fois, de la pure réflexivité solipsiste. Nous avons décrit de manière perspectiviste la construction interpersonnelle du soi qui procède par ces changements de point de vue, par la délocalisation imaginaire. L'identification imaginaire *de se* correspond ainsi non pas à une substitution de particuliers nus cartésiens, mais à un décentrement, à un changement de point de vue que nous pouvons interpréter en terme de mondes possibles centrés sans admettre que « je » fonctionne comme un terme référentiel désignant un ego.

Conclusion générale

Da, plötzlich, Freundin ! wurde Eins zu Zwei —

Friedrich Nietzsche, *Sils Maria*.

Parvenu au terme de ce travail sur le double, voici le moment d'en récapituler les jalons et de regrouper nos résultats.

8.1 Rappel de la problématique

Revenons à notre question de départ. Le problème principal que je me suis proposé au début de ce mémoire était de caractériser le problème du double en philosophie.

En particulier, la question fut de déterminer le poids ontologique des doubles, savoir s'il s'agit d'entités réelles ou seulement d'êtres de langage. Rappelons en effet que l'une des sources d'inspiration de ce travail fut l'interrogation induite à partir de l'œuvre philosophique de Clément Rosset, qu'on pourrait exprimer par l'apparent paradoxe suivant : si la réalité est singulière, donc sans double, il est pourtant certains doubles qui témoignent de la réalité. Si les doubles relèvent de l'illusoire, de la représentation nécessairement fausse, et qu'il vaut mieux alors s'en défaire, sommes-nous contraints à ne plus rien dire de vrai que des tautologies du type « le réel, c'est le réel » ?

Depuis la théorie mimétique de Platon, nous savons que toute théorie métaphysique est une forme de duplication du réel. L'appréhension métaphysique du réel est ainsi toujours accompagnée d'une arrière-pensée, la pensée du double. De fait, la question du double siège à l'articulation entre le monde et son interprétation, entre la réalité et le langage, entre la ressemblance et la représentation.

Notre investigation nous aura engagés à parcourir les théories de l'identité des objets concrets, de l'identité des objets fictionnels, de l'identité des objets esthétiques, et enfin de l'identité personnelle, selon deux interrogations : l'une métaphysique : « qu'est-ce qu'un *double* ? », l'autre, pragmatique : « quand y a t-il *double* ? »

8.2 Rétrospective des grandes lignes

Récapitulons les grandes lignes de notre parcours. Dans un premier temps, il s'agissait de montrer que le double est une notion qui n'est pas réductible à l'incohérent, au contradictoire, à l'irrationnel.

Il apparut pourtant d'emblée une première tension dans l'usage du mot « même », comme par exemple lorsque nous disons que « deux personnes ont la même voiture », expression dont l'ambiguïté balance entre unicité ontologique et ressemblances partagées. Nous avons différencié l'identité numérique de l'identité qualitative, la ressemblance, qui n'est rien d'autre, chez le réaliste des universaux, qu'une application de l'identité numérique aux propriétés.

Le double pose ainsi, comme par contraste, la question de l'égalité, ou identité numérique, que nous avons décrite dans les termes mêmes de la tradition analytique comme une relation logique répondant aux principes de réflexivité et d'indiscernabilité des identiques.

Si l'identité peut s'avérer efficace pour le dénombrement de choses qui seraient comme figées dans le temps, qu'en est-il d'individus soumis aux changements ? Nous avons vu que le problème de la persistance temporelle ne nécessitait pas de souscrire à un universalisme ontologique du double, à condition de considérer l'indiscernabilité des identiques comme un principe temporellement relatif.

L'une des acceptions importantes que revêt l'expression « double » est bien sûr celle d'une quantité deux fois plus grande. Tout part donc de notre manière de compter les choses. Caractériser le double nécessitait-il, avant tout, de circonscrire nos manières de dénombrer, d'envisager nos décomptes ordinaires. Le problème était alors de savoir comment l'identité numérique s'applique aux objets concrets.

Nous avons recontextualisé la question de l'application de l'identité numérique en explicitant les considérations de Frege qui ont mené à qualifier l'application de l'identité comme la recherche de critères d'identité. Nous avons ainsi distingué critère métaphysique d'identité et crible épistémique d'identification.

Nous avons vu plus précisément que le dénombrement d'individus implique principalement deux réquisits. La première condition est de saisir une chose comme un tout saillant sur le fond de son environnement, autrement dit, en tant qu'unité, comme quelque chose qui compte pour un. Se pose alors la question des rapports que l'identité noue à l'unité. Les choses étant constituées de parties, elles pourraient apparaître à la fois comme des unités et à la fois comme des multiplicités, en stimulant ainsi une dialectique du multiple. L'unité est-elle un fait métaphysique ou seulement un fait linguistique ? La seconde condition est de saisir la chose en question comme différente de toutes les autres choses de la même sorte, d'appréhender son unicité. En outre, nous avons distingué deux manières de satisfaire de tels réquisits : soit en considérant qu'ils doivent l'être de manière absolu, parce qu'ils relèvent de conditions métaphysiques qui doivent faire l'enjeu de recherches philosophiques, soit en considérant qu'ils relèvent de conditions pratiques d'identifications contextuelles.

Ainsi, nous avons distingué deux conceptions de l'identité, l'une objectuelle, basée sur la recherche de critères d'identité intrinsèques aux objets, et l'autre conventionnaliste. Là où la première admet que l'identité s'applique de manière absolue aux objets, la seconde affirme que c'est

le contexte qui détermine son application.

Nous avons vu que la théorie objectuelle de l'identité menait à une série de paradoxes. Le sortaliste contourne la difficulté liée aux questions de l'unité et de la persistance des choses en stipulant des sortes, qui sont des découpages objectifs du monde. Mais cette solution implique l'émergence de doubles ontologiques.

Finalement, pour répondre aux paradoxes de l'identité, nous avons adopté la manière toute épistémique dont use Quine, définissant les individus comme des regroupements contextuels d'ostensions, conjointement à une théorie de la perdurance qui considère que les individus sont des composés de phases temporelles et spatiales.

Nous avons ainsi favorisé une théorie qui relativise l'application de l'identité logique au contexte, autrement dit, une théorie pragmatique du jugement d'identité. Alors qu'une théorie comme celle de Geach relativise l'identité logique aux propriétés, nous adoptons une théorie qui relativise l'application de l'identité aux contextes. Nous sommes ainsi passés de la question métaphysique d'individuation à celle, pragmatique, du jugement d'identité. Si l'individuation retourne d'un processus cognitif et non de principes métaphysiques, nous sommes néanmoins stimulés à concevoir l'unité à partir de l'observation de la connexité et de la synchronicité motrice des parties constitutives d'un objet.

La question du double passait par une réflexion sur les mondes répétés. Les conceptions mimétiques des mondes répétés sont ancrées sur la notion d'authenticité, qui accorde un primat axiologique à l'une des entités répétée, censée être plus réelle, là où l'autre, basée sur l'originalité (antécédence temporelle).

Lorsque les philosophes caractérisent les mondes possibles de manière métaphysique, l'identité croisée entre monde confronte les individus transmondains à des problèmes d'authenticité, de doubles vies entre mondes possibles, que nous avons appelé les doubles-mêmes. La résolution du paradoxe de Chisholm impliquait alors pour ces philosophes de souscrire à des doubles fantomatiques comme les particuliers fins, ou à des propriétés haeccéitiques. La théorie métaphysique des contreparties prétend quant à elle résoudre le problème de l'identité transmondaine en adoptant une ontologie proliférante de doubles concrets ancrés à leur monde, en refusant toute légitimité d'identité intermondaine.

Après avoir remis en discussion le principe d'identité des indiscernables, nous avons décliné différents types de ce que seraient des indiscernables ontologiques, les répliques pures, les doubles purs et les coïncidents purs. Nous avons indiqué que leur portée ontologique n'est que présumée, que la stipulation de tels doubles n'est relative qu'aux théories de l'individuation qui les font apparaître.

Aussi, la conclusion principale de notre première partie est que le double n'est pas logiquement impossible. Mais si le double est concevable, je me suis efforcé de montrer au fil des lignes qui précèdent que le double n'a aucun poids ontologique. Les présumés doubles ontologiques ne sont finalement que des émanations de conceptions métaphysiques.

L'idée fut donc de passer de la question « quelle dose de ressemblance peut supporter la représentation par rapport aux choses elles-mêmes », question de la mimésis à laquelle notre réponse est « aucune », à la question « quelle dose de ressemblance peuvent supporter les représentations entre elles ? », à laquelle nous répondons : « la dose la plus complète ». Il n'y a pas de doubles

ontologiques du réel, bien que cela ne nous empêche pas de concevoir des doubles parfaits.

Au cours de ce mémoire, nous avons défendu une vision ontologiquement déflationniste du double tout en maintenant une possibilité logique de le concevoir. Le domaine de la fiction s'est imposé comme intéressant particulièrement la question du double, et nous avons vu comment y relever le défi de ce double réquisit.

La sémantique des mondes possibles nous a permis en effet de rendre compte d'une structure métaleptique et non purement mimétique de nos engagements fictionnels. Nous avons ainsi évité la souscription au pan-réalisme et au pan-fictionnalisme. En outre, nous avons montré que la fiction devait s'évaluer sur des niveaux de profondeurs, et non sur un simple clivage réel *vs.* irréel.

Il a fallu d'abord aborder les relations entre le réel, la fiction et l'illusoire. Si les deux premiers domaines peuvent prétendre à l'énonciation de vérités, ce n'est jamais le cas du troisième. Ainsi, l'irréel doit-il être distingué de l'illusoire. Si les doubles ontologiques relèvent de l'illusoire, les doubles épistémiques peuvent quant à eux relever de l'irréel.

Deux écueils menaçaient cette distinction : soit d'étendre de manière illimitée le territoire de l'irréel en révoquant toute objectivité à l'idée de réalité, soit d'absorber entièrement l'irréel en le considérant comme un département exotique de la réalité. Or nous avons rejeté à la fois la conception pan-fictionnaliste, qui dénie toute référence au réel et la conception pan-réaliste, qui admet que tout contenu de représentation possède une valeur ontologique, y compris les personnages de fiction.

L'adaptation par Lewis de la théorie sémantique des mondes possibles aux fictions par l'ajout d'un opérateur de croyance feinte a donné un cadre formel à une possibilité médiane, en admettant la valeur ontologique du réel, tout en récusant les entités fictionnelles. On est ainsi conduit à désolidariser la signification de l'ontologie, et il devient possible d'attribuer le vrai aux propositions fictionnelles sans accroître l'ontologie d'entités extravagantes.

La question du rapport mimétique que la fiction entretient à la réalité nous a conduit à caractériser la fiction comme la production de duplications épistémiques du réel par un processus que nous avons désigné comme métaleptique plutôt que simplement mimétique. La fiction est certes un double, mais relève plus d'une simulation que d'une imitation du réel. Le fictionnel n'est pas comme un plan uniforme de représentations, il ne se résout pas à une démarcation géographique entre deux territoires qui seraient ontologiquement délimités. Les enchâssements fictionnels exigent de considérer la fiction sur différents niveaux de profondeurs, et non plus selon un seul rapport réalité *versus* fiction. Or l'effet de feuilletage par duplications qui en résulte peut là encore être interprété par des relations entre mondes possibles selon le point de vue métaleptique que nous avons adopté en nous appuyant sur la théorie des lignes de mondes.

Après avoir distingué les trois niveaux matériel (textes), culturel (fictions) et sémantique (mondes possibles) qui se rapportent à nos pratiques fictionnelles, nous nous sommes posés la question de la place de l'œuvre textuelle.

De manière générale, les fictions sont supportées physiquement par des enregistrements, qu'il s'agisse de textes sur du papier ou d'électrons dans des transistors. La vie d'une œuvre est tributaire de sa diffusion, autrement dit, de sa reproduction médiatique. De ce point de vue, nous avons donc été amenés à considérer la notion de répétition du point de vue des productions humaines.

Le fait qu'une œuvre donnée ne soit pas associée à un unique enregistrement ou à une copie de cet enregistrement montre que la reproduction d'une œuvre textuelle ne répond pas seulement à un phénomène de reproduction par copies, mais relève de dupliques fonctionnelles. En outre Le cas du *Pierre Ménard* de Borges montre qu'un même texte peut *fonctionner* selon deux œuvres différentes. Ceci semble indiquer que l'œuvre textuelle relève plus d'une performance que d'une substance.

L'authenticité attribuée à une reproduction textuelle est l'objet d'une discrimination qui n'a rien d'ontologique. Il ne s'agit donc plus de stipuler ici des répliques pures, mais d'invoquer le rôle des dupliques fonctionnelles dont le concept a été développé par Athanassios Tzouvaras. Si la copie prétend reproduire, par imitation, le plus grand nombre possible d'aspects d'un texte original, les dupliques procède par la sélection d'affinités fonctionnelles en rapport avec le contexte de reproduction. Une duplique d'un texte lui est interchangeable dans n'importe quel système où l'une et l'autre joue un rôle. L'œuvre textuelle n'est ni l'inscription matérielle, ni une entité immatérielle, mais une performance réalisée par les dupliques du texte. Les dupliques ont un rôle fondamentale pour la réalisation de cette performance : l'œuvre vit de ses doubles.

Nous avons ainsi montré l'impossibilité d'individuer les textes par des critères mimétiques basés sur des qualités intrinsèques, et avons montré le rôle essentiel des dupliques fonctionnelles à l'égard de l'authentification des textes. En outre, les mondes possibles nous permettent de considérer les vérités fictionnelles comme des assertions contextualisées aux mondes possibles décrits par la fiction.

Rappelons-nous le double réquisit de ce travail : rejeter les doubles ontologiques tout en admettant la concevabilité du double, c'est-à-dire en admettant des doubles épistémiques. L'enjeu est particulièrement saillant dans le cas des personnages de fiction. Il s'agissait donc de porter la question du double au niveau des personnages. Bien sûr, nous n'avons plus dans ce cas une ressource matérielle pour asseoir l'individuation des personnages, et les différentes solutions d'individuation proposées par les philosophes passent inévitablement par la stipulation d'entités exotiques, d'objets inexistants, de simples non actualisés ou d'artefacts abstraits.

Insatisfait des typologies prescrites par les analyses allégorique et symbolique du double, j'ai proposé une catégorisation du double selon un clivage entre doubles natifs (indiscernables occultifs et indiscernables ostensif) et doubles importés (personnages récurrents et métaleptiques).

Confrontés à ces cas de personnages, les partisans des entités fictionnelles, ou *ficta*, se heurtent à la difficulté de produire des doubles ontologiques qui ne répondent pas aux attentes des doubles épistémiques. La possibilité de penser des personnages à la manière de doubles épistémiques est, en quelque sorte, déjà un argument contre les doubles ontologiques. Car nous avons vu les difficultés auxquelles sont confrontées les réalistes devant l'identification de doppelgänger : par exemple, pour le meinongien, il s'agit bien d'hypostasier des entités exactement ressemblantes et contradictoires au principe d'identité des indiscernables qu'il retient. Nous avons également constaté que le possibiliste était amené à introduire des haeccéités, propriétés individuantes déjà douteuses dans le cas d'objets matériels, et qui paraissent encore plus interlopes dans le cas d'objets purement fictionnels.

En prenant un peu de hauteur, on peut sans doute reprocher au réaliste des *ficta* qu'il veut faire reposer la signification entièrement sur l'ontologie, qu'il explique le personnage entièrement

par le contenu de pensée, sans tenir compte des véhicules, des modes de présentation qui servent à représenter de tels contenus. La distinction permet pourtant, comme nous l'a indiqué la théorie des dossiers mentaux, de prendre en compte le fait que des concepts et des pensées puissent différer, de nous affecter différemment en tant que récepteur d'une fiction, quand bien même leurs contenus respectifs seraient strictement indiscriminables.

Le descriptiviste frégéen a donc raison de penser que la référence d'une expression ne suffit pas pour rendre compte de la signification, qu'il importe encore de prendre en charge l'importance sémantique de son mode de présentation. Mais il se fourvoie sans doute, succombant par là à une forme de théorie picturale du langage, en affirmant que l'identité de sens implique nécessairement la co-référence, autrement dit, en individuant épistémiquement ou sémantiquement le mode de présentation. Nous pouvons concevoir des choses doubles parce que nous pouvons mobiliser deux concepts distincts dont les contenus sont indiscriminables, et cela ne nécessite nullement d'individuer les concepts par leur contenu. Les processus cognitifs nous permettent de penser le double parce qu'ils dépendent non seulement du contenu de représentation mais aussi du véhicule même de la représentation, des concepts et des pensées.

Si l'identification d'une œuvre textuelle implique un rapport à des doubles épistémiques, les dupliques fonctionnelles, on peut encore mettre en question de manière plus générale le rapport des œuvres esthétiques aux doubles. Se pose la question de la signature comme manifestation de l'individu, comme représentation de sa singularité existentielle.

Comprendre le mécanisme de nos engagements fictionnels à propos de Sherlock Holmes nous permet aussi de mieux saisir nos engagements fictionnels à propos de *la Pietà* de Michel-Ange. Si l'on peut identifier et dénombrer des personnages sans recourir à des entités exotiques, cela offre une voie pour appuyer l'idée que les objets esthétiques ne superposent pas d'entités coïncidentes, qu'une statue et un bloc de marbre ne sont pas *deux* choses mais une seule chose décrite de deux manières différentes.

L'analyse des processus de reproduction des objets, et en particulier des objets esthétiques nous a porté à interroger la question de l'authenticité non seulement au niveau du particulier, mais plus généralement au niveau spécifique. La question du *kitsch*, nous l'avons vu, n'est pas seulement celle de productions matérielles de mauvais goût, mais plus généralement une interrogation sur une authenticité ontologique du sujet. La méthode critique d'authenfication de Morelli nous a incité à déporter la question de l'authenticité esthétique de l'objet au sujet, de l'œuvre à l'artiste. Le modèle artistique idéal est sans doute, *in fine*, le sujet lui-même. L'idéal de l'œuvre d'art totale est, et cela depuis au moins la mythologie grecque, manifestée par le voeux d'une reproduction artificielle d'un sujet, c'est-à-dire par la production d'un moi selon des processus ne relevant pas de la reproduction sexuée. La digression portant sur les objets esthétiques, nous a finalement conduits à poser la question du double dans son rapport à l'identité personnelle.

Dans le dernier chapitre consacré à la question de l'identité personnelle, nous avons indiqué les raisons qui incitent à renoncer à la conception substantialiste de l'identité personnelle, à la fois dans sa version dualiste et dans sa version moniste. Admettre des égos cartésiens, c'est souscrire à des doubles ontologiques fantomatiques, des particuliers nus. Cependant, Hume a sans doute raison d'affirmer que notre identité n'est qu'une illusion. Ce que nous appelons identité personnelle résulte d'identifications imaginaires, de changements de point de vue, de la possibilité de décen-

trer en imagination notre localisation dans le monde. Nous avons décortiqué deux mécanismes de duplication en rapport avec cette conception de l'identité personnelle : la vision autoscopique et les identifications imaginaires en parallèle avec les représentations picturales de la figure écho et de la figure-filtre.

C'est donc la capacité à s'imaginer autre qui aura retenu notre attention dans le dernier chapitre de ce mémoire. Contre les théories substantialistes du Moi, nous avons indiqué un point de vue performatif de l'identité personnelle. Toute performance narrative ne valant pas pour un moi, la performance narrative est, comme toute identification fictionnelle, régie par des relations interpersonnelles.

8.3 Réponse à la problématique

Ce mémoire aura donc pour conclusion deux résultats, la concevabilité du double et l'inexistence de doubles ontologiques.

Nous avons dit les deux hésitations conflictuelles qui sont responsables du jugement de double : de l'identité à la diversité pour l'une, et de la diversité à l'identité pour l'autre. Soit qu'on croyait avoir affaire à une chose, mais qu'on en postule finalement deux, soit qu'on croyait avoir affaire à deux choses, mais qu'on en postule finalement qu'une seule.

De telles croyances auraient pu passer pour des erreurs provisoires, qu'on peut toujours surmonter par un examen plus attentif. D'ailleurs, c'est sans doute ce dont voudraient nous convaincre les philosophes qui prescrivent des doubles ontologiques aux problèmes d'identité des choses matérielles. On croyait avoir affaire à une seule entité, une statue, en voici deux coïncidentes, un objet d'art et un bloc de matière. Nous ne devrions plus faire d'erreur à présent. Pourtant, les paradoxes de l'identité montrent que, sous la prescription de doubles ontologiques, ceux-ci ne font que s'accroître, ce n'est plus, sur ce tapis, *un* chat et *une* masse de tissus félins, mais autant de chats coïncidents que l'autorisent nos manières de découper les entités sur le fond du monde.

Les doubles ontologiques surgissent lorsque l'on affirme que les mots suivent les contours ontologiques des objets, principe premier de la théorie picturale du langage. Abandonner l'idée du caractère ontologique du double, c'est rendre le double à son contexte, à ses motifs humains, à sa fonction cognitive.

Au moment de considérer le chemin parcouru à travers ces pages, il s'agit aussi de revenir sur le sous-titre de cette thèse : « la tension entre l'imitation et la représentation ou les paradoxes de l'identité fictionnelle ». Pour synthétiser les résultats qui précèdent, nous pouvons retenir que cette tension se manifeste dans les théories picturales du langage, théories ancrées sur le caractère mimétique du langage qui admettent qu'un nom puisse référer à un individu indépendamment de son mode de représentation à travers les mondes possibles. De telles théories qui, confrontées aux paradoxes de l'identité induisent des doubles ontologiques dont nous avons spécifié les types.

Nous avons vu que le double nous confronte à la question du caractère spéculaire du langage : dans quelle mesure le langage détermine-t-il notre connaissance du réel et dans quelle mesure le reflète-t-il ? Cette question est sans doute tributaire d'une théorie picturale du langage, et nous mène, lorsqu'on accepte une vision mimétique du langage, à admettre des doubles ontologiques.

Nous parlons du monde à partir d'un schème basé sur des duplications fonctionnelles, qui ne sont pas des duplications par ressemblances naturelles. Nous ne pouvons parler de réel sans lui avoir déjà substitué un double épistémique, ne serait-ce que par le fait d'imposer à notre langage un schème conceptuel destiné à parler du réel. Croire que le langage imite le réel, c'est faire proliférer des doubles ontologiques, et se perdre dans le non-sens d'entités coïncidentes et fantomatiques, d'un monde peuplé de coïncidents purs et de particuliers fins. Il suffit sans doute de renoncer au caractère spéculaire du langage, et d'admettre l'importance fondamentale des doubles épistémiques dans notre rapport pragmatique au réel.

Je terminerai en répondant donc ainsi à la question de Rosset : les doubles ontologiques peuvent être associés aux doubles illusoires, alors que les doubles épistémiques se rapprochent de ce que Rosset appelle « doubles de proximité ».

8.4 Quelques pistes de réflexion

Le plaidoyer exprimé dans ce mémoire est que l'économie de l'individuation doit être analysée en vertu des transactions du double. Selon les circonstances, notre rapport aux choses nous fait parfois renoncer à certaines unités au profit d'autres, aussi, le présent travail fut essentiellement consacré à l'aspect conceptuel du double. Nous avons ainsi balisé différents types de doubles ontologiques et épistémiques. Mais faute de temps et de place, nous avons dû renoncer à d'autres perspectives. Des travaux ultérieurs mériteraient néanmoins de compléter l'aspect conceptuel du double par une évaluation des enjeux qu'il occasionne dans la perception, dans la croyance et dans les comportements.

On peut penser, en esthétique, à la question du double point de vue. A cet égard, Maurice Denis adressait à ses amis Nabis le conseil suivant : « se rappeler qu'un tableau – avant d'être un cheval de bataille, une femme nue, ou une quelconque anecdote – est essentiellement une surface plane recouverte de couleurs en un certain ordre assemblées ». Contemplant une toile, nous voyons à la fois une surface peinte et une représentation. En outre, les nouvelles technologies produisent des changements qui affectent nos objets de perception, mais sans doute aussi, la perception elle-même par une sorte de double perspective. Or si nous pouvons voir une même chose de différentes façons, nous pouvons aussi et plus généralement concevoir notre réalité selon des points de vue contradictoires et simultanés, dont, là encore, l'interprétation en mondes possibles peut apporter un intérêt explicatif, à la manière dont nous avons indiqué dans ce qui précède comment la nouvelle de Maupassant pouvait être lue simultanément de deux façons incompatibles.

Concernant le sentiment du double, mentionnons par exemple la question de la sensation du *déjà vu*, dont nombre de philosophes et de psychologues ont tenté d'expliquer le mécanisme qui produit en celui qui en fait l'expérience, une forme intense de désorientation temporelle. Il serait intéressant de mobiliser, quant à la description de ce phénomène qui semble procéder d'une sorte de court-circuit entre la perception et la mémoire, la ressource des mondes possibles pour expliquer cette sorte de perspective ancrée sur une temporalité dédoublée.

Mais plus généralement encore, nous pourrions aussi aborder ce que j'aimerai appeler l'éthologie humaine du double, autrement dit, le rapport au double dans les comportements que nous adoptons. De faits, différents objectifs sont assignés aux doubles, qui servent autant à représenter

le réel, qu'à s'en détourner. Les fictions et les sciences dupliquent le monde par des représentations selon des buts qui peuvent converger ou non, soit qu'elles tentent de le comprendre, d'en projeter une ombre dans l'ordre de la connaissance, soit qu'elles tentent, chacune à sa manière, d'en fuir le caractère complexe et totalitaire. Le rapport de croyance au double est autant oraculaire qu'hallucinatoire, il est lui-même double, il tient autant de la recherche d'une reproduction fidèle que d'un désaveu pur et simple nous incitant à « aller voir ailleurs si nous y sommes ».

Index des noms

Épicure, 103, 104, 106

Adams, Robert, 167, 196, 197
Alexandre d'Aphrodise, 100, 101
Allais, Alphonse, 502
Amblard, Maxime, 583
Amiel, Henri-Frédéric, 570
Anaximène, 99
Anaximandre, 102
Andersen, Hans Christian, 60, 412
Anouilh, Jean, 366
Anscombe, Elizabeth, 606
Apollinaire, 333, 334
Apollonius de Thyane, 355
Arctinos de Milet, 354
Arendt, Hannah, 566
Aristote, 38, 105, 246–248, 250, 256, 281, 295, 298, 355, 479, 537, 570
Armstrong, David, 87
Artaud, Antonin, 297
Aucouturier, Gustave, 589
Auerbach, Erich, 298, 429
Augustin, 355, 570
Aurèle, Marc, 570
Auster, Paul, 507, 508
Ayer, Alfred, 564
Ayoade, Richard, 222

Bachelard, Gaston, 227, 232, 298
Baker, David, 95
Bakhtine, Mikhaïl, 409
Ballard, J.G., 295
Balzac, Honoré de, 229, 309, 312, 415, 416
Barnard, Christiaan , 557
Barrie, James Matthew, 535
Barthes, Roland, 282–285, 487
Bateson, Gregory, 583
Baud, Jean-Pierre, 86
Baudelaire, Charles, 301, 382
Baudrillard, Jean, 282, 293–296
Baxter, Donald, 65
Bay, Michael, 222, 403
Bayard, Pierre, 381
Bazin, André, 365, 487
Beckett, Samuel, 407
Bellosi, Luciano, 496
Benjamin, Walter, 486, 492, 493, 503, 508
Benovski, Jiri, 488
Benveniste, Émile, 608
Bergson, Henri, 599
Berkeley, George, 22, 64, 65
Berthoz, Alain, 588
Berto, Francesco, 258, 389, 430, 431, 434, 456, 460
Bessière, Jean, 226, 455
Beziau, Jean-Yves, 460, 461
Black, Max, 191, 192, 195–198, 208, 213
Blanc, Michel, 222
Blanqui, Auguste, 105–109
Bloch, Robert, 412
Boèce, 38
Bontempelli, Massimo, 600
Borel, Émile, 379
Borges, Jorge Luis, 326, 334, 371–376, 378–384
Bower, T.G.R., 81
Bradbury, Ray, 354, 355
Broch, Hermann, 508, 509
Brun, Jean, 100

Brunelleschi, 496
Brunschwig, Jacques, 570
Buechler, John Carl, 222
Burke, Michael, 40
Burnet, Frank, 549
Burton, Tim, 446
Butler, Joseph, 554, 556
Byrne, Alex, 460
Byrne, Ruth, 460

Céline, Louis-Ferdinand, 381
Cabrol, Christian, 557
Caillois, Roger, 455, 478
Calle, Sophie, 507, 508, 515
Campbell, John, 80
Campbell, Keith, 548
Camus, Albert, 365, 412, 422
Cantor, Georg, 63
Capgras, Joseph, 411
Carnap, Rudolf, 240–242, 245
Carpenter, John, 416, 623
Carrara, Massimiliano, 496
Carroll, Lewis, 414, 600
Casati, Roberto, 61
Cassirer, Ernst, 63, 64
Castañeda, Hector-Neri, 310, 604, 619, 628
Cervantes, Miguel de, 371–373, 378–381
Chalmers, David, 296, 460
Chamisso, Adalbert von, 392, 393, 578
Chamisso, Aldabert von, 60, 412
Chastel, André, 499
Chauvier, Stéphane, 61, 79
Cherubini, 501
Chisholm, Roderick, 41, 146, 147, 149, 150, 159–161, 163–167, 169, 261
Christie, Agatha, 315, 388
Chrysippe, 39, 100, 101
Churchill, Winston, 229
Cimabue, Cenni di Pepo dit, 485
Cléanthe, 100
Clément, René, 222, 405
Clair, Jean, 496, 504
Clarke, Samuel, 180
Claudel, Paul, 366
Coates, Paul, 390
Collodi, Carlo, 510, 534
Cometti, Jean-Pierre, 486
Conan Doyle, Arthur, 233, 241, 262, 267, 270, 302, 305–309, 311, 312, 314–316, 388, 412
Contim, Filipe Drapeau Vieira, 19, 148, 158, 160, 203, 205–208
Contin, Filipe Drapeau Vieira, 155
Conty, Évrard de, 588
Copernic, 287
Cortázar, Julio, 415
Costa-Leite, Alexandre, 459
Crane, Tim, 445, 446, 452
Cronenberg, David, 222, 404, 412
Cukor, George, 222
Currie, Gregory, 369–371

Dällenbach, Lucien, 414
Démocrite, 102
Danto, Arthur, 317, 477, 501–503
Darwin, Charles, 594
Daumier, Honoré, 291
Davidson, Donald, 330
Dawkins, Richard, 479
De Kooning, Willem, 507
Dedekind, Richard, 63
Deleuze, Gilles, 102
Delevsky, Jacques, 106
Dennett, Daniel, 446, 564
Descartes, René, 460, 510–514, 519, 521–523, 525–534, 538, 541, 551, 573, 574, 584, 593, 609, 610
Descombes, Vincent, 179, 416, 530, 532, 539, 562, 566, 579, 584, 606
Desnos, Robert, 405
Deutsch, David, 114
DeWitt, Bryce, 114
Dick, Philip Kindred, 229–231, 295, 403, 509, 510, 534, 555, 565
Dickens, Charles, 229, 604, 605
Dickie, George, 503
Diderot, Denis, 415, 592
Dokic, Jérôme, 626

Doležel, Lobomír, 395
Doležel, Lubomír, 317–319, 429
Donnellan, Keith, 274
Dostoïevski, Fiodor, 222, 322, 360, 365, 399, 408–410, 567, 578, 583, 589
Doumenc, Philippe, 315, 413
Dours, Christian, 518, 603
Dryden, John, 398
Duchamp, Marcel, 347, 501, 502
Ducrot, Oswald, 390
Durand, Gilbert, 391

Eco, Umberto, 229, 296, 311, 314, 366, 378, 413, 422, 477, 490, 491, 598
Eddington, Arthur, 111
Eklund, Mati, 75
Elgin, Catherine, 356, 364, 365, 374–376, 378–382
Eliade, Mircea, 96, 97
Ellis, Bret Easton, 230, 585
Elsig, Frédéric, 499
Empédocle, 99, 104, 247
Engel, Pascal, 58, 94
Ensor, James, 406
Ephèse, Michel d', 570
Epicharme, 29
Erikson, Erik, 579, 580
Escher, Maurits Cornelis, 459
Euclide, 18
Evans, Gareth, 58, 426
Evans, Walker, 489, 490
Everett, Anthony, 321
Everett, Hugh, 113, 114

Ferraris, Maurizio, 300
Ferrer, Daniel, 359
Ferret, Stéphane, 36, 45
Ficher, David, 222
Fichte, Johann Gottlieb, 539–542
Fielding, Henry, 383
Fine, Kit, 463
Finney, Jack, 411
Fiocco, Oreste, 460
Flaubert, Gustave, 315, 413, 567, 568
Fleming, Victor, 222
Fontenelle, 105
Forbes, Graham, 163
Foror, Jerry, 448
Foucault, Michel, 568
Fraenkel, Abraham, 461
Frazer, James George, 392, 534
Frege, Gootlob, 8
Frege, Gottlob, 14–19, 22–25, 37, 62–65, 71, 72, 79, 141, 172, 237–240, 243, 244, 248, 249, 251, 255, 258, 261, 305, 447, 448, 458, 604
Freud, Sigmund, 280, 456, 505, 506, 533, 535–538, 587
Friedrich, Caspar David, 542
Friend, Stacie, 349
Frontisi-Ducroux, Françoise, 401

Gabriel, Markus, 93, 94
Gandhi, 229
Garnier, Robert, 366
Gary, Romain, 449
Gassendi, 609
Geach, Peter, 39, 72, 74, 418, 445, 619
Genette, Gérard, 5, 247, 312, 381, 383, 384, 415, 428
George, Fanny, 627
Gesell, Arnold, 548
Gibbard, Allan, 36, 53
Giesz, Ludwig, 493
Gilson, Étienne, 501
Girard, René, 541, 581–584
Giraudoux, Jean, 398
Goethe, Johann Wolfgang von, 572, 589
Goffman, Ervin, 580
Gogol, Nicholas, 222, 408, 412, 559, 567, 578
Goldman, Alvin, 626
Gombrowicz, Witold, 570
Goodman, Nelson, 94, 125, 222, 232, 288–292, 294, 312, 325, 326, 330, 332, 338, 339, 341, 342, 356, 357, 364, 365, 367, 371, 374–376, 378–382, 477, 482, 486, 490, 497, 498, 501, 502

Gorgias, 279–282
Granger, Gilles Gaston, 481
Greimas, Algirdas, 565
Grimm, Jacob, 413
Grimm, Wilhelm, 413
Gueroult, Martial, 531
Guichard, Charlotte, 499
Gusdorf, Georges, 571, 572

Héraclite, 96, 99, 100, 288
Héraclite d'Éphèse, 29
Hésiode, 98
Haüy, René Just, 393
Haber, Stuart, 345
Hacking, Ian, 194–196, 213
Hannah, Sophie, 388
Harvey, Steve, 502
Hawley, Katherine, 197
Hegel, Georg Wilhelm Friedrich, 300
Heidegger, Martin, 68, 69, 491, 492, 503
Heil, John, 7, 92, 351
Heinlich, Nathalie, 506
Heintz, John, 309
Herdman, John, 578
Hergé, 21, 404
Highsmith, Patricia, 405
Hilbert, David, 330
Hintikka, Jaakko, 303, 384, 450, 451, 454, 609–611
Hirsch, Eli, 75, 76, 82, 83
Hirst, R.J., 586
Hitchcock, Alfred, 222, 412
Hobbes, Thomas, 513, 551, 609
Hodges, Wilfrid, 9
Hoffmann, Ernst Theodor Amadeus, 403, 534, 536, 541
Hofmannsthal, Hugo von, 60
Homère, 248
Horace, 298
Houdé, Olivier, 81
Howell, Robert, 311
Hume, David, 24, 29, 263, 459, 545, 562, 629
Husserl, Edmund, 522, 529, 530
Huysmans, Joris-Karl, 322, 353

Ingarden, Roman, 258, 267–269, 271, 276, 388
Ingres, Jean-Auguste-Dominique, 291, 501
Iser, Wolfgang, 319
Iñárritu, Alejandro González, 222

Jacob, François, 546
Jago, Mark, 460
James, Henry, 407
James, Montague Rhodes, 463
James, William, 578
Janet, Pierre, 555
Janzen, Daniel H., 546
Jaśkowski, Stanislaw, 464
Jean-Paul, 222, 390, 541
Jentsch, Ernst, 533, 536–538
Jeunet, Jean-Pierre, 403, 404
Johnston, Mark, 42
Jourde, Pierre, 221, 392, 394, 539
Joyce, James, 381
Jung, Carl Gustav, 393
Jung, Johann, 462

Köhler, Wolfgang, 84, 85
Kôbô, Abe, 60
Kafka, Franz, 45, 384, 403, 583
Kant, Immanuel, 255, 299, 300, 310, 475, 476, 539
Kantorowicz, Ernst, 402
Kaplan, David, 137, 167, 605
Kieslowski, Krzysztof, 222
Kimber, Susan, 548
Kishono, Fumio, 297
Klein, Felix, 461
Kleist, Heinrich von, 398
Kovacshazy, Cécile, 390, 391, 538, 539
Kraus, Wilhelmine, 541
Kripke, Saul, 67, 68, 122, 125, 133–135, 137, 140–145, 153–155, 164, 168–172, 174–176, 206, 212, 213, 271–274, 306–308, 418, 430, 435, 444, 456, 460, 545, 616
Kuhn, Thomas, 287, 288

Kundera, Milan, 487, 509
Kung, Peter, 460
Kurosawa, Akira, 222, 402, 623

L'Isle-Adam, Villiers, 510
Laing, Ronald D., 603
Lapovsky, Ben, 344
Lavoisier, 106
Le Cœur, Marc, 498
Le Douarin, Nicole, 549
Leibniz, Gottfried Wilhelm, 10, 105, 118, 122–124, 144, 160, 294
Leinbiz, Gottfried Wilhelm, 180, 187–189, 191
Lenain, Thierry, 500
Leopardi, Giacomo, 110
Lepeltier, Thomas, 115
Leucippe, 99, 102
Levine, Sherrie, 489, 490
Lewis, David, 48, 50–53, 57, 58, 87, 124–128, 131–133, 135, 145, 146, 151–160, 164–167, 170, 182, 183, 202, 203, 206–211, 213, 232, 258, 279, 305–309, 311–313, 315–318, 382, 450, 451, 543–545, 604, 628
Lienhardt, Godfrey, 582
Locke, John, 20, 22, 45, 463, 550–552, 554, 555, 562
Long, Anthony Arthur, 100
Lord, Albert B., 355
Lorrain, Jean, 406
Losey, Joseph, 407
Lovecraft, Howard Phillips, 79
Lovelace, Ada, 513
Lucrèce, 102, 103, 106, 593
Lynch, David, 222, 531

Möbius, August Ferdinand, 416
Méredieu, Florence de, 344
MacDonald, Margaret, 233
Mach, Ernst, 340, 587
Magritte, René, 591
Maitte, Corinne, 326, 494
Mally, Ernst, 262, 265
Malraux, André, 485
Manet, Édouard, 591
Margitte, René, 456
Martinière, Nathalie, 578
Marx, Karl, 103, 104
Masaccio, 465, 628
Maupassant, Guy de, 221, 360, 410, 454–456, 458, 461–465, 467, 593
McLuhan, Marshall, 354, 355
Mead, George Herbert, 573–576, 584, 630
Meinong, Alexius, 256–264, 266, 269
Merleau-Ponty, Maurice, 593, 610
Michel-Ange, 291, 484, 490
Milgram, Paul, 297
Mill, John Stuart, 62, 105, 142, 418
Miller, Carl, 578
Millikan, Ruth, 449, 450
Minsky, Marvin, 482
Miyoshi, Masao, 578
Molière, 354, 398–400, 611
Molnar, Véra, 345
Mondrian, Pietr, 344
Monod, Jacques, 546
Montaigne, Michel de, 518
Montesquieu, 105
Morelli, Giovanni, 505, 506
Mori, Masahiro, 599
Morizot, Jacques, 334, 477
Morris, Robert, 507
Musiol, Michel, 583
Myro, George, 48

Némésius, 101
Nabokov, Vladimir, 222, 542
Nakamoto, Satoshi, 346
Neumann, Kurt, 403
Nietzsche, Friedrich, 98, 99, 104, 110, 232, 299–302, 331, 530, 533, 572, 616
Nieztsche, Friedrich, 300
Niiniluoto, Ilkka, 459
Ninan, Dilip, 453
Nitezsche, Friedrich, 99
Noll, Michael, 344
Nougé, Paul, 334
Nyberg, Henrik Samuel, 96

Offenbach, Jacques, 365
Origène, 101
Orlan, 510
Ovide, 587, 596

Péguy, Charles, 102
Pagin, Peter, 448
Pagnol, Marcel, 403
Paquet-Brenner, Gilles, 222
Parfit, Derek, 518, 550, 560–562
Parménide, 28, 29, 92, 93, 288
Parry, Milman, 355
Parsons, Terence, 257, 264–266, 419–421, 423–427
Pascal, Blaise, 488, 584, 585
Pastoureau, Michel, 496
Peacocke, Christopher, 620
Peele, Jordan, 222
Peirce, Charles Sanders, 391
Penrose, Roger, 459
Penzo, Giorgio, 300
Perrault, Charles, 413
Perry, John, 604, 628
Phidias, 485
Philon d'Alexandrie, 40
Piaget, Jean, 63, 81
Pinard, Ernest, 567
Pindare, 247
Pirandello, Luigi, 577
Plantinga, Alvin, 125, 167
Platon, 70, 96, 100, 148, 149, 177–180, 234–239, 246, 247, 284, 300, 355, 521, 569, 570, 580
Plaute, 398–402, 409, 542
Plessner, Helmuth, 576, 577
Plotin, 92
Poe, Edgar, 511, 512, 590
Poe, Edgard, 450
Porandello, Luigi, 577, 578
Pouivet, Roger, 349, 350, 362, 364
Pradeu, Thomas, 549
Price, Marjorie, 46
Priest, Graham, 258, 264, 389, 430–434, 456, 460
Prieto, Arthur, 483
Prieto, Luis, 476
Prieto, Luis Jorge, 214
Propp, Vladimir, 565
Proust, Joëlle, 562
Proust, Marcel, 563, 567
Pruszkowski, Krzysztof, 406
Pryor, James, 448
Ptolémée, 287
Putnam, Hilary, 174, 616
Pyrrhon, 496
Pythagore, 17

Quine, Willard Van Orman, 10, 31, 51, 53–58, 76, 83, 125, 151, 154, 182, 193, 239, 242–246, 248, 249, 263–266, 303, 330, 364, 371, 423, 424, 603

Racine, Jean, 382
Rahman, Shahid, 450
Ramsey, Frank, 65
Rank, Otto, 60, 392, 393, 401, 533–536, 538
Rantala, Veikko, 304
Rauschenberg, Robert, 507
Rebuschi, Manuel, 16, 234, 246, 583, 601, 610–612
Recanati, François, 447–450, 454, 604, 611, 621, 622, 626
Reid, Thomas, 554, 556
Reitman, Ivan, 401
Rembrandt, 499
Renauld, Marion, 233, 246, 336
Renoir, Auguste, 497–500
Rey, Abel, 110
Richardson, Samuel, 383
Ricœur, Paul, 473, 532, 565, 566
Riffaterre, Michel, 285, 286
Rizzolatti, Giacomo, 480
Rogers, Robert, 538, 578
Romero, George A., 222
Rosenberg, Jay, 495
Rosset, Clément, 18, 27, 28, 70, 78, 79, 92, 104, 116, 117, 135, 179, 180, 227,

228, 231, 232, 234, 248, 279, 300, 404, 461, 486, 520, 563
Rotrou, Jean, 398, 399
Rouault, Georges, 291
Rousseau, Jean-Jacques, 570
Routley, Richard, 263
Rubens, Pierre Paul, 624
Rudder Baker, Lynne, 42
Rudder, Orlando de, 355
Russell, Bertrand, 17, 67, 80, 93, 119, 140, 168, 184, 189–191, 198, 233, 243, 244, 248, 249, 256, 260, 261, 263, 264, 267, 269, 271, 305, 315, 419, 424, 431, 458
Ryan, Marie-Laure, 314, 456
Ryle, Gilbert, 495

Sagoff, Mark, 484
Sainsbury, Richard Mark, 449
Salmon, Nathan, 165
Sapho, 247
Sartre, Jean-Paul, 93, 227, 234, 496, 580
Saussure, Fredinand, 282
Scarbrough, Will, 222
Scarron, Paul, 415
Schaeffer, Jean-Marie, 339, 340, 373, 374, 479
Scheffler, Israël, 291, 292, 330
Schelling, Friedrich Wilhelm Joseph von, 300, 541
Schoemaker, Sydney, 557
Schopenhauer, Arthur, 300
Schroeder, Barbet, 222
Schuppert, Guillaume, 620
Scorsese, Martin, 222
Scott, Ridley, 361, 403, 534, 557, 565
Searle, John, 388
Sedley, David, 29, 100
Serres, Michel, 338
Sextus Empiricus, 280
Shakespeare, William, 401, 475
Shelley, Mary, 403, 510, 534
Shoemaker, Sydney, 556, 558, 605, 606
Shyamalan, Night, 222
Sider, Theodore, 87
Smith, Barry, 58
Soavi, Marzia, 496
Socrate, 100, 101, 570
Sophocle, 366
Spinoza, Baruch, 92
Stéphan, Lucien, 500
Stalnaker, Robert, 125, 128, 628
Stelarc, 510
Stendhal, 229
Stevenson, Charles, 364
Stevenson, Robert Louis, 222, 412
Stobée, 100
Stoichita, Victor, 590
Stornetta, W. Scott, 345
Stougatski, Arcadi, 406
Stougatski, Boris, 406
Strauss, Richard, 60
Stravinsky, Igor, 365
Strawson, Peter, 66, 449, 450
Sturtevant, Elaine, 489, 515
Swift, Jonathan, 380
Szeminska, Alina, 81

Telenheimo, Tero, 451
Terashima, Nobuyoshi, 296
Thalès, 99, 288
Thomason, Judith, 42
Thomasson, Amie, 258, 275–278, 349, 383, 388, 428, 435–439, 441, 442
Tiffin, John, 296
Tintoret, 501
Todorov, Tzvetan, 390, 435, 456, 461, 466, 624
Tolkien, John Ronald Reuel, 424
Tolstoï, Léon, 229, 311, 429, 458
Tortonese, Paolo, 221, 392, 394, 539
Torza, Alessandro, 208
Tournier, Michel, 366
Troubetzkoy, Wladimir, 392, 393, 410, 541, 585
Tulenheimo, Tero, 450–452, 454
Turing, Alan, 344, 510, 512–514
Turner, Victor, 582

Twain, Mark, 453
Tye, Michael, 449
Tymms, Ralph, 541
Tzouvaras, Athanassios, 327, 328

Uexküll, Jakob von, 95
Unger, Peter, 41, 56, 57

Valéry, Paul, 313, 382
Van Meegeren, Hans, 499, 500
Vanden Eeckhoudt, J.P., 478
Varzi, Achille, 58
Vaucanson, Jacques, 475
Vautier, Ben, 507
Vax, Louis, 462, 466
Vega, Lope de, 402
Vendler, Zeno, 616, 620, 621, 623, 624
Vermeer, Johannes, 500
Vernant, Jean-Pierre, 400, 568–570, 603
Veyne, Paul, 288
Vian, Boris, 246
Villeneuve, Denis, 222
Viollet-le-Duc, Eugène, 483
Voltaire, 105
Voltolini, Alberto, 263, 440–442, 611

Wölfflin, Heinrich, 499
Wachowski, Lanna, 411
Wachowski, Lilly, 411
Wallon, Henri, 594
Walton, Kendall, 351, 487, 615, 617, 618, 620–625, 627
Wansing, Heinrich, 459, 465
Warhol, Andy, 347, 501, 502
Wehmeier, Kai, 71
Wiggins, David, 37, 38, 42, 46, 47, 197, 463, 559
Wilde, Oscar, 412
Williams, Bernard, 554, 613, 615, 617, 619–621, 624
Wilmut, Ian, 548
Wittgenstein, Ludwig, 68, 69, 289, 450, 605, 607, 622
Wollheim, Richard, 353, 355, 617
Woo, John, 222
Woods, John, 458, 459
Woolf, Virginia, 411
Wreen, Michael, 495
Wynn, Karen, 82

Xu, Fei, 80

Zénon, 100
Zénon d'Élée, 28
Zalta, Edward, 257, 265, 266, 419, 421
Zazzo, René, 402, 548, 591, 592, 594–601, 629
Zemach, Eddy, 381
Zemeckis, Robert, 411
Zermelo, Ernst, 461
Zeuxis, 236

BIBLIOGRAPHIE

Œuvres de fiction

ANDERSEN, H. C. (1876). L'Ombre (1847). In *Contes d'Andersen* (p. 172–190). Traduction par David Soldi. Paris : Librairie Hachette et Cie.

ANONYME. (1998). *Divines insanités de Nasr Eddin Hodja*. Domaine turc. Traduction par J.L. Maunoury. Paris : Phébus.

APOLLINAIRE. (1965a). Calligrammes, poèmes de la paix et de la guerre (1913-1916). In *Œuvres poétiques*. Bibliothèque de la Pléiade. Paris : Gallimard.

APOLLINAIRE. (1965b). Le bestiaire. In *Œuvres poétiques*. Bibliothèque de la Pléiade. Paris : Gallimard.

AUSTER, P. (1992). *Léviathan* (T. de C. LEBOEUF, Éd.). Paris : Actes Sud.

BARIE, J. (1902). *The Little White Bird*. United Kingdom : Hodder et Stoughton.

BAUDELAIRE, C. (1961). Spleen de Paris. In *Œuvres complètes*. Bibliothèque de la Pléiade. Paris : Gallimard.

BAY, M. (2005). *The Island*. Produit par Michael Bay, Ian Bryce, Laurie MacDonald et Walter Parkes. Scénario de A. Kurtzman, R. Orci & C. Tredwell-Owen. USA : DreamWorks SKG et Warner Bros. Pictures.

BONTEMPELLI, M. (1990). *L'échiquier devant le miroir (1921)*. L'Arpenteur, Domaine italien. traduit de l'italien par Jean-Baptiste Para. Gallimard.

BORGES, J. L. (1957). Le jardin aux sentiers qui bifurquent. In *Fictions* (p. 91–104). Paris : Gallimard.

BORGES, J. L. (1983). Pierre Ménard, auteur du Quichotte. In *Fictions* (p. 41–52). Paris : Gallimard.

BRADBURY, R. (1985). *Fahrenheit 451*. Présence du Futur. Paris : Denoël.

BRAHM, J. (p.d.). *The Twilight Zone, épisode 21, saison 1 mirror image*. Diffusé aux États-Unis le 26 février 1960 sur CBS. Rediffusion sur TF1 le 4 avril 1984. Scénario : Rod Serling.

CARPENTER, J. (1978). *Halloween*. Scénario de John Carpenter et Debra Hill. USA : produit par Compass International Pictures, Falcon International Productions.

CARPENTER, J. (1995). *In the Mouth of Madness*. Traduction française : *L'Antre de la folie*, scénario de Michael De Luca. Production : New Line Cinema. États-Unis.

CARROLL, L. (2010). *La traversée du miroir et ce qu'Alice trouva de l'autre côté* (Traduction de L. Bury). Calmann-Lévy.

CHAMISSO, A. v. (1874). Préface à l'édition française. In *Merveilleuse histoire de Pierre Schlémihl ou l'homme qui a perdu son ombre (traduction française de 1838)*. Paris : Librairie Hachette et Cie.

CHAMISSO, A. v. (1973). L'étrange histoire de Peter Schlemihl ou l'homme qui a vendu son ombre (1814). In *Romantiques allemands tome 2*. Bibliothèque de la Pléiade. Paris : Gallimard.

COLLODI, C. (2003). *Les aventures de Pinocchio*. Traduction de la Comtesse de Gencé. Paris : Le livre de poche.

de CONTY, É. (1990). *Le livre des échecs amoureux moralisés*. Bibliothèque du Moyen Français. Édition de B. Roy et F. Guichard-Tesson. Montréal : CERES.

DICK, P. K. (1968). *Do Androids Dream of Electric Sheep ?* Doubleday.
DICK, P. K. (1987). Comment construire un univers qui ne s'effondre pas deux jours plus tard. In *Le crâne*. Paris : Denoël.
DICKENS, C. (1854). *Des temps difficiles*. Folio Classique. Trad. de l'anglais par Andhrée Vaillant. Préface de Pierre Gascar. Gallimard.
DOSTOÏEVSKI, F. (1969). *Le Double (1865)*. Folio. Traduction de Gustave Aucouturier. Paris : Gallimard.
DOUMENC, P. (2007). *Contre-enquête sur la mort d'Emma Bovary*. Paris : Actes Sud.
DOYLE, C. (1979a). Le signe des quatre. In *Sherlock holmes* (T. 1). Bouquins. Paris : Robert Laffont.
DOYLE, C. (1979b). Une étude en rouge. In *Sherlock holmes* (T. 1). Bouquins. Paris : Robert Laffont.
GOETHE, J. W. v. (1941). *Poésie et vérité, souvenirs de ma vie*. Traduction française de Pierre du Colombier. Paris : Aubier.
GOGOL, N. (1966). Le nez (1836). In *Œuvres complètes* (p. 597–621). Bibliothèque de la Pléiade. Paris : Gallimard.
HANNAH, S. (2014). *Meurtres en majuscules*. Paris : Éditions du Masque.
HERGÉ. (1954). *On a marché sur la Lune*. Tournai : Casterman.
HOFFMANN, E. T. A. (1979). L'homme au sable. In *Contes nocturnes*. Libretto. Paris : Éditions Phébus.
HOFMANNSTHAL, H. v. (1992). *La Femme sans Ombre (1919)*. Der Doppelgänger. traduit de l'allemand par Jean-Yves Masson. Paris : Verdier.
HOMÈRE. (1955b). Odyssée. In *Illyade, Odyssée*. Bibliothèque de la pléiade. Paris : Gallimard.
JEUNET, J.-P. (1997). *Alien : Resurrection*. Produit par : Badalato, Carroll, Giler and Hill. Scénario de J. Whedon. USA : Production 20th Century Fox.
KAFKA, F. (1980). La métamorphose. In *Kafka : œuvres complètes, tome 2*. Bibliothèque de la Pléiade. Traduction française de *Die Verwandlung*, Kurt Wolff Verlag, Leipzig : 1917. Paris : Gallimard.
KÔBBÔ, A. (1994). *Les murs*. Poches. Récits traduits du japonais par Marc Mécréant. Picquier.
KUNDERA, M. (1985). *Le livre du rire et de l'oubli (1979)*. Collection Du monde entier. Traduction française par François Kérel. Paris : Gallimard.
KUROSAWA, A. (1950). *Rashomoon*. Scénario de Akira Kurosawa, Shinobu Hashimoto et Ryunosuke Akutagawa. Japon : Production : Daiei.
L'ISLE-ADAM, V. (1993). *L'ève future (1886)*. Folio classique. Gallimard.
LOVECRAFT, H. P. (2000). Le cauchemar d'Innsmouth. In *La couleur tombée du ciel*. Folio SF. Nouvelles traduites de l'Américain par Jacques Papy et Simone Lamblin. Gallimard.
LYNCH, D. & FROST, M. (2017). *Twin peaks : the return, épisode 18 : what is your name ?* Scénario de Daniel Mainwaring et Richard Collins, série diffusée pour la première fois entre le 21 mai 2017 et le 3 septembre 2017 sur la chaîne Showtime. États-Unis : Lynch/Frost Productions, Propaganda fictions, Spelling Television.
MAUPASSANT, G. (1979a). Le Horla (1886). In *Contes et nouvelles, tome 2*. Bibliothèque de la Pléiade. Paris : Gallimard.
MAUPASSANT, G. (1979b). Lui ! (1883). In *Contes et nouvelles, tome 2*. Bibliothèque de la Pléiade. Paris : Gallimard.

MOLIÈRE. (1947). Amphitryon. In *Œuvres complètes*. bibliothèque de la pléiade. Gallimard.
O'BRIEN, F. J. (1995). The Lost Room (1858). In R. DALBY (Éd.), *The mammoth book of victorian and edwardian ghost stories* (p. 55–70). London.
ODAR, B. b. & FRIESE, J. (2017). *Dark*. USA : Production : Wiedemann et Berg Television.
OVIDE. (1992). *Les métamorphoses*. Folio classique. Édition de Jean-Pierre Néraudau, Traduction française de Georges Lafaye. Gallimard.
PAUL, J. (1963). *Siebenkäs (1796)*. Collection bilingue des classiques étrangers. introduction, traduction, Note par Pierre Jalabert. Paris : Aubier Éditions Montaigne.
PEELE, J. (2019). *Us*. USA : Universal Pictures.
POE, E. (1932). William Wilson. In *Œuvres en prose*. Bibliothèque de la Pléiade. Traduit de l'anglais (États-Unis) par Charles Baudelaire. Paris : Gallimard.
PROUST, M. (1987). Le Côté de Guermantes. In *À la recherche du temps perdu* (T. 2). Bouquins. Robert Laffont.
SARTRE, J.-P. (1981). La nausée (1938). In *Œuvres romanesques*. Bibliothèque de la Pléiade. Paris : Gallimard.
SCOTT, R. (1982). *Blade runner*. Scénario : Hampton Fancher et David Webb Peoples, d'après le roman de Philip Kindred Dick. USA : The Ladd Company.
SCOTT, R. (2012). *Prometheus*. Produit par R. Scott. Scénario de Damon Lindelof et Jon Spaihts. USA : production Brandywine Productions, Scott Free Productions, Dune Entertainment.
SHELLEY, M. (2008). *Frankenstein ou le Promethée moderne (1818)*. FolioPlus. Traduit de l'anglais par Paul Couturiau. Paris : Gallimard.
STROUGATSKI, A. & STROUGATSKI, B. (1974). *Le lundi commence un samedi*. Présence du Futur. Paris : Denoël.
SWIFT, J. (1965). Voyages de gulliver. In *Œuvres*. Bibliothèque de la Pléiade. Paris : Gallimard.
TOLKIEN, J. R. R. (1992). *Le seigneur des anneaux*. Traduction de Francis Ledoux. Paris : Christian Bourgois.
VIAN, B. (1963). *L'écume des jours*. Paris : Pauvert.
WACHOWSKI, L. [Léna] & WACHOWSKI, L. [Lilly]. (1999). *Cloud Atlas*. Adaptation du roman *Cartographie des Nuages* de David Mitchell. États-Unis, Allemagne : production Anarchos Pictures, Cloud Atlas Productions, X-Filme Creative Pool.
WOOLF, V. (1928). *Orlando : a biography*. Londres : Hogarth Press.

Essais et articles

ADAMS, R. M. (1974). Theories of Actuality. *Noûs*, *8*, 211–231.
ADAMS, R. M. (1979). Primitive Thiness and Primitive Identity. *The Journal of Philosophy*, *76*(1), 5–26.
ADAMS, R. M. (2007). Les théories de l'actualité. In *Métaphysique contemporaine, propriétés, mondes possibles et personnes*. Paris : Librairie philosophique Vrin.
ARISTOTE. (1996). *Poétique*. Paris : Gallimard.
ARISTOTE. (2001). *Catégories*. Belles Lettres.
ARISTOTE. (2004). *Traité du ciel*. GF. Traductionpar C. Dalimier et P. Pellegrin. Paris.

ARMSTRONG, D. M. [D. M.]. (1978). *A theory of universals (volume 2 : universals and scientific realism)*. Cambridge : Cambridge University Press.
ARMSTRONG, D. M. [David M.]. (2010). *Les universaux, une introduction partisane.* Paris : Ithaque.
ASHE, G. (1985). *The discovery of king arthur*. New York : Doubleday.
AUERBACH, E. (1968). *Mimésis. La représentation de la réalité dans la littérature occidentale. (1946)*. Bibliothèque des Idées. Paris : Gallimard.
AYER, A. J. (1952). *Language, Truth and Logic (1946)*. New York : Dover.
AYER, A. J. (1954). *Philosophical Essays*. London : Palgrave Macmillan.
BACHELARD, G. (1943). *L'air et les songes. Essai sur l'imagination du mouvement*. Paris : José Corti.
BAKER, L. R. (1997). Why constitution is not identity. *The Journal of Philosophy*, *94*, 599–621.
BAKER, L. R. (2000). *Persons and bodies : a constitution view*. Cambridge Studies in Philosophy. Cambridge University Press.
BAKHTINE, M. (1970). *Problèmes de la poétique de Dostoïevski (1927)*. Slavica. Lausanne : L'Âge d'Homme.
BARCAN, M. (1947). The Identity of Individuals in a Strict Functional Calculs of Second Order. *Journal of Symbolic Logic*, *12*, 12–15.
BARTHES, R. (1970). *S/Z*. Paris : Éditions du Seuil.
BARTHES, R. (1980). *La chambre claire*. Les cahiers du cinéma. Paris : Éditions de l'Étoile, Gallimard, Le Seuil.
BARTHES, R. (1984). La mort de l'auteur. In *Essais critiques*. Points. Paris : Éditions du Seuil.
BARTHES, R. (1985). *L'aventure sémiologique*. Points. Paris : Éditions du Seuil.
BARTHES, R. (2000). Le discours de l'histoire (1967). In *Le bruissement de la langue : essais critiques iv*. Collection points : Séries essais. Paris : Éditions du Seuil.
BATESON, G. (1972). *Steps to an Ecology of Mind : Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology. Traduit en français chez Seuil, collection « point essais » sous le titre* Vers une écologie de l'esprit. University Of Chicago Press.
BAUD, J.-P. (1993). *L'affaire de la main volée. Une histoire juridique du corps*. Des Travaux. Paris : Éditions du Seuil.
BAUDRILLARD, J. (1976). *L'échange symbolique et la mort*. Bibliothèque des Sciences Humaines. Paris : Gallimard.
BAUDRILLARD, J. (1979). *De la séduction*. Collection L'espace critique. Paris : Galilée.
BAUDRILLARD, J. (1981). *Simulacre et simulation*. Paris : Galilée.
BAUDRILLARD, J. (1991). *La guerre du golfe n'a pas lieu*. Paris : Galilée.
BAUDRILLARD, J. (2000). *Mots de passe*. Folio Essais. Paris : Grasset.
BAXTER, D. L. M. (1991, mars). Berkeley, Perception, and Identity. *Philosophy and Phenomenological Research*, *51*(1), 85. doi :10.2307/2107821
BAYARD, P. (2010). *Et si les œuvres changeaient d'auteur ?* Paris : Éditions de Minuit.
BAYARD, P. (2014). *Il existe d'autres mondes*. Paradoxe. Paris : Éditions de Minuit.
BAZIN, A. (1958). *Qu'est-ce que le cinéma ? i. ontologie et langage*. Collection "7e Art". Paris : Éditions du Cerf.
BECKER, H. S. (1963). *Outsiders. Studies in the Sociology of Deviance*. New York : The Free Press, a Division of Macmillan Publishing Co., Inc.

BENJAMIN, W. (2000). Kitsch onirique (1927). In *Œuvres ii*. folio essais. Paris : Gallimard.

BENJAMIN, W. (2005). *L'œuvre d'art à l'époque de sa reproductibilité technique (1972)*. Éditions Allia.

BENOIST, J. (2001). *Représentations sans objet aux origines de la phénoménologie et de la philosophie analytique*. Paris : Presses Universitaires de France.

BENOVSKY, J. (2012). Le réalisme modal de David Lewis nous condamne-t-il à la souffrance éternelle ? *Revue Klesis*, *La philosophie de David Lewis*(24), 36–55.

BENVENISTE, É. (1966). De la subjectivité dans le langage. In *Problèmes de linguistique générale, tome 1*. Paris : Gallimard.

BERGMANN, G. (1967). *Realism*. Madison.

BERGSON, H. (1946). *Matière et mémoire (1896)*. Quadrige. Paris : Presses Universitaires de France.

BERGSON, H. (1955). *Le rire (1900)*. Genève : Albert Skira.

BERKELEY, G. (2009a). A Treatise Concerning the Principles of Human Knowledge. In M. C. DESMOND (Éd.), *Philosophical Writings*. Cambridge : Cambridge University Press.

BERKELEY, G. (2009b). An Essay Towards a New Theory of Vision. In M. C. DESMOND (Éd.), *Philosophical Writings*. Cambridge : Cambridge University Press.

BERKELEY, G. (2009c). Three Dialogues Between Hylas and Philonous. In M. C. DESMOND (Éd.), *Philosophical Writings*. Cambridge : Cambridge University Press.

BERTHOZ, A. (2014). Le cerveau créateur de mondes : données récentes sur les bases neurales de « l'imagerie mentale ». *2014*, Institut des Hautes Etudes pour la Science et la Technologie.

BERTO, F. (2013). *Existence as a real property. the ontology of meinongianism*. Synthese library. Studies in Epistemology, logic, methodology ans philosophy of science. Springer.

BERTO, F. (2014). On conceiving the inconsistent. *Proceedings of the Aristotelian Society*, *114*, 103–121.

BERTO, F. (2017). Counting the particles : entity and indentity in the philosophy of physics. *Metaphysica*, *18*(1).

BERTO, F. & SCHOONEN, T. (2018). Conceivability and possibility : some dilemmas for humeans. *Synthese*, *195*(6), 2697–2715.

BESSIÈRE, J. (2018). Le double. Chamisso, Dostoïevski, Maupasant, Nabokov. In E. DURANTE & D. AMAURY (Éds.), *Le double, littérature, arts, cinéma. Nouvelles approches*. Collection Colloques, congrès et conférences sur la littérature comparée. Paris : Honoré Champion.

BÉZIAU, J.-Y. (2016). Possibility, imagination and conception. *Principios*, *23*(40).

BIASI, P.-M. d. (2011). *Génétique des textes*. Biblis. Paris : CNRS Éditions.

BIDEAUD, J. (2001). *Les chemins du nombre*. PU du Septentrion.

BLACK, M. (2012). L'identité des indiscernables. *Philosophia Scientiæ*, *16*(3), 121–132. Traduction par Sébastien Motta de *The Identity of Indiscernibles*, Mind, vol. 61, no 242, avril 1952, 153-164.

BLANCHOT, M. (1959). *Le livre à venir*. Folio Essais. Paris : Gallimard.

BLANQUI, A. (1872). *L'éternité par les astres*. Paris : Librairie Germer Baillière.

BOLZANO, B. (1843). Aufsatz, worin eine von Herrn. Exner in seiner Abhandlung, 'Uber Nominalismus und Realismus' angeregte logische Frage beant wortet wird (1843). In E. WINTER, J. BERG, F. KAMBARTEL, L. J. & van ROOTSELAAR B. (Éds.), *Bernard Bolzano Gesamtausgabe*. Stuttgart : Bad Cannstatt.

BOREL, É. (1972). Comptes rendus de l'académie des sciences de 1930 t.190. In *Œuvres d'émile borel* (T. 2, p. 1139–1142). Paris : C.N.R.S.

BOUVERESSE, R. (1987). *La force de la règle. Wittgenstein et l'invention de la nécessité*. Critique. Paris : Éditions de Minuit.

BOWER, T. (1974). *Development in Infancy*. W. H. Friedmand.

BRENTANO, F. (1944). *Psychology from an Empirical Standpoint (1944)*. International Library of Philosophy. London : Routledge.

BROCH, H. (2001). *Quelques remarques à propos du kitsch (1955)*. Paris : Éditions Allia.

BRUN, J. (1958). *Le stoïcisme*. Que sais-je ? Presses universitaires de France.

BRUNSCHWIG, J. (1983). Aristote et l'effet Perrichon (1987). In J.-L. MARION (Éd.), *Hommage à ferdinand alquié : la passion de la raison* (p. 361–377). Épiméthée. Paris : Presses Universitaire de France.

BURKE, M. B. (1994). Dion and Theon : an essentialist solution to an ancient puzzle. *Journal of Philosophy*, *3*(91), 129–139.

BURNET, F. M. (1969). *Self and Not-Self*. Cambridge : Cambridge University Press.

BUTLER, J. (1975). Of Personal Identity, The Analogy of Religion (1736). In J. PERRY (Éd.), *Personal Identity*. Berkeley : University of California Press.

BYRNE, A. (2007). Possibility and imagination. *Philosophical Perspectives*, *21*(1), 125–144.

BYRNE, R. (2005). *The rational imagination : how people create alternatives to reality* (1re éd.). Bradford Books. The MIT Press.

CAILLOIS, R. (1938). *Le mythe et l'homme*. Folio. Paris : Gallimard.

CAILLOIS, R. (1966). *Images, images... essais sur le rôle et les pouvoirs de l'imagination*. Paris : José Corti.

CAMPBELL, J. (2002). *Reference and Consciousness*. Oxford Cognitive Science Series. Oxford University Press, USA.

CAPGRAS, J. & REBOUL-LACHAUX, J. (1923). L'illusion des sosies dans un délire systématisé chronique. *Bulletin de la Société Clinique de Médecine Mentale*, *11*, 6–16.

CARNAP, R. (1950). Empiricism, Semantics, and Ontology. In *Meaning and Necessity*. Chicago : University of Chicago Press.

CARNAP, R. (1997). *Signification et nécessité*. Collection Bibliothèque de Philosophie. trad. de F. Rivenc et de P. de Rouilhan. Paris : Gallimard.

CARRARA, M. & SOAVI, M. (2010). Copies, Replicas, and Counterfeits of Artworks and Artifacts. *The Monist*, *93*(3), 414–432.

CASATI, R. (2002). *La découverte de l'ombre. De Platon à Galilée, l'histoire d'une énigme qui a fasciné les grands esprits de l'humanité.* Le livre de poche.

CASSIRER, E. (1953). *Substance and Function. Substance et fonction. Éléments pour une théorie du concept*, Traduction de l'allemand par Pierre Caussat aux Éditions de Minuit, 1977 : Paris. Dover Pubns.

CASTAÑEDA, H.-N. (1968). On the Logic of Attributions of Self-Knowledge to Others. *The Journal of Philosophy*, *68*(15), 439–456.

CASTAÑEDA, H.-N. (1979). Fiction and Reality : their Fundamental Connections. An Essay on the Ontology of Total Experience. *Poetics*, *8*, 31–62.

CASTAÑEDA, H.-N. (1999). *The Phenomeno-Logic of the I : Essays on Self-Consciousness* (J. HART & T. KAPITAN, Éds.). Bloomington : Indiana University Press.

CHALMERS, D. (2002). Does conceivability entail possibility ? In T. GENDLER & J. HAWTHORNE (Éds.), *Conceivability and Possibility* (p. 145–200). Oxford University Press.
CHALMERS, D. J. (2017). The Virtual and the Real. *Disputatio*, *9*(46), 309–352. Récupérée à partir de https://content.sciendo.com/view/journals/disp/9/46/article-p309.xml
CHANDLER, H. S. (1976). Plantinga and the Contingently Possible. *Analysis*, *36*(2), 106–109.
CHASTEL, A. (1974). Signature et signe. *Revue de l'art*, *26*.
CHAUVIER, S. (2010). L'unique en son genre. *Philosophie*, *3*(106), 3–22.
CHISHOLM, R. (1963). The Logic of Knowing. *The Journal of Philosophy*, *60*(25), 773–795.
CHISHOLM, R. (1967). Identity Through Possible Worlds : Some Questions. *Noûs*, *1*.
CHISHOLM, R. (1973a). Homeless Object. *Revue Internationale de Philosophie*, *2-3*, 207–223.
CHISHOLM, R. (1973b). Parts as Essential to Their Wholes. *The Review of Metaphysics*, *26*(4), 581–603.
CHISHOLM, R. (1976). *Person and Object : A Metaphysical Study*. Open Court Publishing Company.
CHISHOLM, R. (1986). *Brentano and Intrinsic Value*. Cambridge : Cambridge University Press.
CLAIR, J. (2012). Vrais faux et faux vrais. *Médium*, *3-4*(32 - 33), 146–168.
COATES, P. (1988). *The Double and the Other : Identity as Ideology in Post-Romantic Fiction*. Palgrave Macmillan UK.
COLLECTIF. (1988). *Vrai ou faux ? Copier, imiter, falsifier*. Paris : Bibliothèque nationale de France.
COLLECTIF. (2001). *Les présocratiques*. Bibliothèque de la Pléiade. Paris : Gallimard.
COLONNA, V. (2004). *Auto-fiction et autres mythomanies littéraires*. Auch : Tristram.
COMETTI, J.-P. (2015). *Conserver/Restaurer. L'œuvre d'art à l'époque de sa préservation technique*. nrf Essais. Paris : Gallimard.
CONCHE, M. (1986). *Héraclite - fragments*. Presses Universitaires de France.
COSTA-LEITE, A. (2010). Logical properties of imagination. *Abstracta*, *6*(1), 103–116.
COUDERC, P. (1946). *Le calendrier*. Que sais-je ? Paris : Presses Universitaires de France.
CRANE, T. (2013). *Objects of thought*. Oxford : Oxford University Press.
CRESSWELL, J. (2006). From Modal Discourse to Possible Worlds. *Studia Logica : An International Journal for Symbolic Logic*, *82*(3), 307–327.
CRESSWELL, M. J. (1979). The World Is Everything That Is the Case. Ithaca : Cornell University Press.
CROSS, R. (2014). Medieval theories of haecceity. In E. ZALTA (Éd.), *The Stanford Encyclopedia of Philosophy* (Summer 2014). Metaphysics Research Lab, Stanford University.
CURRIE, G. (1989). *An Ontology of Art*. Studies in Contemporary Philosophy. Palgrave Macmillan UK.
CURRIE, G. (1991, juillet). Work and Text. *Mind*, *100*(3), 325–340.
DALIBARD, J. (2016). Statistique de Bose–Einstein et condensation. Collège de France. Dernière lecture le 24/07/2021. Récupérée à partir de https://www.college-de-france.fr/media/jean-dalibard/UPL8114772573593160085_cours_2016_____Chapitre_1.pdf
DÄLLENBACH, L. (1977). *Le récit spéculaire. Essai sur la mise en abyme*. Poétique. Paris : Seuil.
DAMER, B. & HINRICHS, R. (2014). The Virtuality and Reality of Avatar Cyberspace. In M. GRIMSHAW (Éd.), *The Oxford Handbook of Virtuality*. New York : Oxford University Press.

DANTO, A. (1981). *The Transfiguration of the Commonplace : A Philosophy of Art*. London : Harvard University Press.
DANTO, A. (1993). *L'assujettissement philosophique de l'art*. Paris : Éditions du Seuil.
DANTO, G. (1964). The Artworld. *The Journal of Philosophy*, *61*(19), 571–584.
DANTO, G. (2016). Réponse de danto. *Cahiers philosophiques*, *1*(144), 130–133. doi :10.3917/caph.144.0130
DAUSSET, J. (1990). La définition biologique du soi. In J. BERNARD, M. BESSIS & C. DEBRU (Éds.), *Soi et non-soi. des biologistes, médecins, philosophes et théologiens s'interrogent*. Collection science ouverte. Paris : Seuil.
DAVIDSON, D. (1968). On Saying That. *Synthèse*, (19), 130–146. repris dans *Words and Objections*, sous la dir. de D. Davidson et H. Hintikka, Dordrecht, Reidel, 1969, pp. 158-174.
DAWKINS, R. (1989). *The Selfish Gene (1976)*. Oxford Landmark Science. Oxford : Oxford University Press.
DE RUDDER, O. (1983). Pour une histoire de la lecture. *Médiévales*, *3*, 97–110.
DECLOS, A. (2017). *La métaphysique de Nelson Goodman* (thèse de doct.). Thèse de doctorat de philosophie dirigée par Roger Pouivet et Isabelle Thomas-Fogiel, Université de Lorraine 2017. Récupérée à partir de http://www.theses.fr/2017LORR0238/document
DELEUZE, G. (1968). *Différence et répétition*. Epiméthée. Paris : Presses Universitaires de France.
DELEUZE, G. (1969). *Nietzsche et la philosophie*. Quadrige. Paris : Presses Universitaires de France.
DELEVSKY, J. (1945). Note sur la possibilité des répétitions cosmologiques. *Isis*, *36*(1), 19–21.
DELLA ROCCA, M. (2005). Two Spheres Twenty Spheres and the Identity of Indiscernibles. *Pacific Philosophical Quarterly*, *86*.
DENNETT, D. [Daniel]. (1992, janvier). The Self as a Center of Narrative Gravity. In F. KESSEL, P. COLE & D. JOHNSON (Éds.), *Self and Consciousness : Multiple Perspectives*. Hillsdale : NJ : Erlbaum.
DENNETT, D. [D.C.]. (1968). Geach on Intentional Identity. *Journal of Philosophy*, *65*, 335–341.
DESCARTES, R. (1953a). Discours de la méthode. In *Œuvres*. Bibliothèque de la Pléiade. Paris : Gallimard.
DESCARTES, R. (1953b). Les principes de la philosophie. In *Œuvres et lettres*. Bibliothèque de la Pléiade. Paris : Gallimard.
DESCARTES, R. (1953c). Méditations métaphysiques. In *Œuvres et lettres*. Bibliothèque de la Pléiade. Paris : Gallimard.
DESCOMBES, V. (1971). *Le platonisme*. Collection sup le philosophe. Paris : Presses Universitaires de France.
DESCOMBES, V. (1983). *Grammaire d'objets en tous genres*. Critique. Paris : Les Éditions de Minuit.
DESCOMBES, V. (2004). *Le complément du sujet*. nrf essais. Paris : Gallimard.
DESCOMBES, V. (2010). Le marteau, le maillet et le clou. *Revue de Métaphysique et de Morale*, *4, Le Moi/The Self/Le Soi*, 495–519.
DESCOMBES, V. (2013). *Les embarras de l'identité*. Paris : Gallimard.
DESCOMBES, V. (2014). *Le parler de soi*. Folio Essais. Paris : Gallimard.
DESCOMBES, V. (2019). Avoir une identité ? Récupérée à partir de https://www.youtube.com/watch?v=JY4am8oIKEc

DESCOMBES, V. & LARMORE, C. (2009). *Dernières nouvelles du moi*. Quadrige Essais débats. Paris : Presses Universitaires de France.
DEUTSCH, D. (1997). *The Fabric of Reality*. Viking Adult.
DEWITT, B. (1973). The Many Universes Interpretation of Quantum Mechanics. In B. DEWITT & N. GRAHAM (Éds.), *The Many Worlds Interpretation of Quantum Mechanics*. New Jersey : Princeton University Press.
DIDEROT. (1956). *Lettre sur les aveugles à l'usage de ceux qui voient (1749)*. Classiques Garnier. Garnier.
DIDEROT, D. (2006). *Jacques le fataliste et son maître (1785)*. Folio. Gallimard.
DIVERS, J. (2002). *Possible Worlds*. Routledge.
DOKIC, J. (2008). Epistemic Perspectives on Imagination. *Association Revue Internationale de Philosophie*, *1*(243), 99–118.
DOLEŽEL, L. (1985). Le triangle du double, un champ thématique. *Poétique*, *64*, 463–471.
DOLEŽEL, L. (1988). Mimesis and Possible Worlds. *Poétique. Aspects of Literary Theory*, *9*(3), 475–496.
DOLEŽEL, L. (1998). *Heterocosmica. Fiction and Possible Worlds*. Parallax. Londres : Johns Hopkins.
DONNELLAN, K. S. (1966). Reference and Definite Descriptions. *The Philosophical Review*, *75*(3), 281–304. doi :10.2307/2183143
DONNELLAN, K. S. (1974). Speaking of Nothing. *The Philosophical Review*, *83*(1), 3–31.
DOUBROVSKY, S. (2001 [1977]). *Fils*. Filio. Paris : Gallimard.
DOURS, C. (2003). *Personne, personnage. Les fictions de l'identité personnelle*. Æsthetica. Rennes : Presses Universitaires de Rennes.
DRAPEAU VIEIRA CONTIM, F. (2010). *Qu'est-ce que l'identité ?* Paris : Librairie philosophique J. Vrin.
DRAPEAU VIEIRA CONTIM, F. (2012). L'essence en contexte. *Klesis – Revue philosophique. La philosophie de David Lewis*, *24*.
DRAPEAU VIEIRA CONTIM, F. (2016). Identité. In KRISTANEK (Éd.), *L'encyclopédie philosophique* (Édition académique).
DUCHÊNE, H. (2006). La guerre du faux n'aura pas lieu. *Dossiers Archéologie et sciences des origines*, *312*.
DUCROT, O. & TODOROV, T. (1972). Personnage. In O. DUCROT & T. TODOROV (Éds.), *Dictionnaire encyclopédique des sciences du langage* (p. 286–292). Collection Points Essais. Paris : Éditions du Seuil.
DURAND, G. (1989). *L'imaginaire symbolique (1964)*. Quadridge. Paris : Presses Universitaires de France.
EARMAN, J. (1989). *World Enough and Space-Time. Absolute versus Relational Theories of Space and Time*. Cambridge : MIT Press.
ECO, U. (1978). Pour une reformulation du concept de signe iconique. doi :10.3406/comm.1978.1438
ECO, U. (1979). *Lector in fabula. Le rôle du lecteur ou la coopération interprétative dans les textes narratifs*. Collection biblio essais. Le livre de poche.
ECO, U. (1985). *La guerre du faux*. Paris : Grasset.
ECO, U. (1992). *Les limites de l'interprétation (1990)*. Paris : Grasset.

ECO, U. (2010). Quelques commentaires sur les personnages de fiction. *SociologieS*. Récupérée à partir de https://sociologies.revues.org/3141
ECO, U. (2013). *Confessions d'un jeune romancier*. Paris : Le Livre de poche.
EDDINGTON, A. S. (1927). *The Physical Nature of the World. Gifford Lectures of 1927*. Editions H. G. Callaway.
EKLUND, M. (2009, janvier). Carnap and Ontological Pluralism.
ELIADE, M. (1969). *Le mythe de l'éternel retour*. Folio Essais. Paris : Éditions Gallimard.
ÉLIADE, M. (1957). *Le sacré et le profane*. Folio Essais. Paris : Éditions Gallimard.
ÉLIADE, M. (1963). *Aspects du mythe*. Folio Essais. Paris : Éditions Gallimard.
ELLIS, B. E. (2019). *White*. Pavillon. Traduction française de Pierre Guglielmina. Paris : Robert Laffont.
ELSIG, F., GUICHARD, C., PARSHALL, P., SÉNÉCHAL, P. & PHILIPPE, B. (2009). Le connoisseurship et ses révisions méthodologiques. *Perspective (revue en ligne)*, *3*. Récupérée à partir de http://perspective.revues.org/1303
ENGEL, P. (1985). *Identité et référence*. Paris : Presses de l'École Normale Supérieure.
ENGEL, P. (2004). Les objets vagues le sont-ils vraiment ? *Cahiers de philosophie de l'Université de Caen*, *40-41*, 3–19.
ENGEL, P. (2015). Le réalisme kitsch. Récupérée à partir de https://zilsel.hypotheses.org/2103
ENGEL, P. & NEF, F. (1988). Identité, vague et essences. *Les Études philosophiques n°4. Logique et philosophie de la connaissance*, 475–494.
ERDRICH, N. (2018). Pourquoi Hercule Poirot n'est-il pas un objet abstrait ? In C. BOURIAU & G. SCHUPPERT (Éds.), *Perspectives philosophiques sur les fictions* (p. 3–22). Kimé.
EVANS, G. (1978). Can There be Vague Objects ? *Analysis*, *38*(4), 208.
EVANS, G. (1982). *The Varieties of Reference*. New York : Clarendon Press Oxford.
EVEN-ZOHAR, I. (1985). Les règles d'insertion des "réalèmes"dans la narration. *Littérature*, *57*, 109–118. Traduction française par Ruth Amossy.
EVERETT, A. (2013). *The Nonexistent*. Oxford University Press.
EVERETT, H. (1973). The Theory of the Universal Wave Function (1957). In DEWITT, B. S., GRAHAM & NEILL (Éds.), *The Many-Worlds Interpretation of Quantum Mechanics*. Princeton : Princeton University Press.
FERRER, D. (2011). *Logiques du brouillon. Modèle pour une critique génétique*. Poétique. Paris : Éditions du Seuil.
FERRET, S. (1996). *Le bateau de Thésée. le problème de l'identité à travers le temps*. Paris : Les Éditions de Minuit.
FERRET, S. (1998). *L'identité*. Corpus. Paris : GF-Flammarion.
FICHTE, J. G. (2000). *La doctrine de la science nova methodo*. Le livre de poche.
FICKERT, K. (1979). *Kafka's Doubles*. Berne : Lang.
FINE, K. (1994). Essence and modality : the second philosophical perspectives lecture. *Philosophical Perspectives*, *8, Logic and Language*, 1–16.
FINE, K. (2000). A counter-example to locke's thesis. *The Monist*, *83*(3), 357–361.
FINE, K. (2003). The non-identity of a material thing and its matter. *Mind*, *112*(446), 195–234.
FIOCCO, M. O. (2007). Conceivability, imagination and modal knowledge. *Philosophy and Phenomenological Research*, *74*(2), 364–380.
FODOR, J. (1987). Psychosemantics : The Problem of Meaning in the Philosophy of Mind.

FONTENELLE. (1852). Entretiens sur la pluralité des mondes (1686). In *Œuvres de Fontenelle. étude sur sa vie et son esprit par Voltaire, la Marquise de Lambert, Grimm, Garat, Sainte-Beuve, Arsène Houssaye*. Paris : Eugène Didier.

FORBES, G. (1980). Origin and Identity. *Philosophical Studies : An International Journal for Philosophy in the Analytic Tradition*, *37*(4), 353–362.

FORBES, G. (1984). Two Solutions to Chisholm's Paradox. *Philosophical Studies : An International Journal for Philosophy in the Analytic Tradition*, *46*(2), 171–187.

FRANÇOIS, R. (1983). La sémantique des noms propres : remarques sur la notion de « désignateur rigide ». *Langue française*, *57*, 106–118. doi :https://doi.org/10.3406/lfr.1983.5159

FRAZER, J. G. (1981). *Le rameau d'or. Le roi magicien dans la société primitive. Tabou et les périls de l'âme (1890)*. Bouquins. Paris : Robert Laffont.

FREGE, G. (1956). *The Thought : A Logical Inquiry*.

FREGE, G. (1969a). *Les fondements de l'arithmétique (1884)*. L'ordre philosophique. Traduction française de Claude Imbert. Paris : Éditions du Seuil.

FREGE, G. (1969b). Recherches logiques. In *Les fondements de l'arithmétique*. L'ordre philosophique. Traduction de Claude Imbert. Éditions du Seuil.

FREGE, G. (1971a). Écrits logiques et philosophiques. traduction de claude imbert. Éditions du Seuil.

FREGE, G. (1971b). Sens et dénotation. In *Écrits logiques et philosophiques*. Traduction française de Claude Imbert. Paris : Éditions du Seuil.

FRENCH, S. (1989). Why the Principle of the Identity of Indiscernibles is not Contingently True Either. *Synthese*, *78*(2), 141–166.

FRENCH, S. & KRAUSE, D. (2006). *Identity in Physics. A Historical, Philosophical and Formal Analysis*. Oxford : Oxford University Press.

FREUD, S. (1938). Résultats, idées, problèmes. In *Résultats, idées, problèmes* (T. 2). Paris : Presses universitaires de France.

FREUD, S. (1984). Traitement psychique (traitement d'âme) (1890). In *Résultats, idées, problèmes* (T. 1). Paris : Presses universitaires de France.

FREUD, S. (1985a). Le créateur littéraire et la fantaisie. In *L'inquiétante étrangeté et autres essais (1919)*. Coll. « Folio Essais ». Paris : Gallimard.

FREUD, S. (1985b). Le Moïse de Michel-Ange. In *L'inquiétante étrangeté et autres essais (1919)* (p. 85–125). Coll. « Folio Essais ». Paris : Gallimard.

FREUD, S. (1985c). L'inquiétante étrangeté. In *L'inquiétante étrangeté et autres essais*. Folio Essais. Paris : Gallimard.

FREUD, S. (1992). *Le trait d'esprit et sa relation à l'inconscient (1905)*. Folio Essais. Paris : Gallimard.

FREUD, S. (2010). *Névrose, psychose et perversion*. Presses Universitaires de France.

FRIEND, S. (2007). Fictional Characters. *Philosophy Compass*, *2*.

FRONTISI-DUCROUX, F. (2007, janvier). Les Grecs, le double et les jumeaux. *Enfances & Psy*, *34*, 118. doi :10.3917/ep.034.0118

GABRIEL, M. (2013). *Warum es Die Welt Nicht Gibt*. Traduction française : *Pourquoi le monde n'existe pas*, traduit par Georges Sturm, éditions Jean-Claude Lattès, 2014, Paris. Berlin : Ullstein Buchverlage GmbH.

GEACH, P. (1957). On Beliefs about Oneself. *Analysis*, *18*, 23–24.

GEACH, P. T. (1962). *Reference and Generality. An Examination of Some Medieval and Modern Theories*. London : Cornell University Press.
GEACH, P. T. (1967). *Identity*. Cornell University Press.
GEACH, P. T. (1987, janvier). Reference and Buridan's Law. *Philosophy*, *62*(239), 7–15. doi :10.1017/S0031819100038547
GENETTE, G. (1966). Frontières du récit. In *L'analyse structurale du récit*. Collection Points. Éditions du Seuil.
GENETTE, G. (1972). *Figure III*. Poétique. Paris : Seuil.
GENETTE, G. (1994). *L'œuvre de l'art*. Poétique. Paris : Éditions du Seuil.
GENETTE, G. (2004). *Fiction et diction (1991)*. Collection Poétique. Paris : Éditions du Seuil.
GENETTE, G. (2016, juillet). Ménard et nous. *Fabula-LhT*, *n° 17, Pierre Ménard, notre ami et ses confrères*. Récupérée à partir de http://www.fabula.org/lht/17/genette.html
GEORGES, F. (2013). L'immersion fictionnelle dans le jeu vidéo. Le cas de Silent Hill. *Presses Universitaires de France | « Nouvelle revue d'esthétique*, *11*, 51–61.
GIBBARD, A. (1975). Contingent Identity. *Journal of Philosophical Logic*, (4), 187–221.
GIBBS, J. W. (1902). *Elementary Principles in Statistical Mechanics*. New york : Charles Scribner's sons.
GIESZ, L. (1960). *Phänomenologie des kitsches*. Heidelberg : Rothe.
GILSON, É. (1972). *Peinture et réalité*. Paris : Vrin.
GINZBURG, C. (1980). Signes, traces, pistes. Racines d'un paradigme de l'indice. *Le Débat*, (6), 3–44.
GIRARD, R. (1983a). *Critiques dans un souterrain (1976)*. biblio essais. Paris : Le livre de Poche.
GIRARD, R. (1983b). *Mensonge romantique et vérité romanesque*. Pluriel. Paris : Grasset.
GIRARD, R. (1993). *La violence et le sacré (1972)*. Pluriel. Paris : Grasset.
GOLDMAN, A. (2006a). Imagination and Simulation in Audience Responses to Fiction. In S. NICHOLS (Éd.), *The Architecture of the Imagination : New Essays on Pretence, Possibility, and Fiction*. Clarendon Press.
GOLDMAN, A. (2006b). Simulating Minds : The Philosophy, Psychology, and Neuroscience of Mindreading.
GOODMAN, N. (1949). On likeness of meaning. *Analysis*, *10*(1), 1–7.
GOODMAN, N. (1972). Seven Strictures on Similarity. In *Problems and Projects :* (p. 437–450). Idianapolis : Bobbs-Merrill.
GOODMAN, N. (1973). *Fact, Fiction, and Forecast* (4th). Cambridge : Indianapolis : Bobbs-Merrill.
GOODMAN, N. (1976). *Languages of Art : An Approach to a Theory of Symbols*. Traduction française *Langages de l'art*, Collection Pluriel Hachette, Paris : 1990. Hackett Pub. Co. Inc.
GOODMAN, N. (1993). On Some Worldly Worries. *Synthese*, *95*(1), 9–12.
GOODMAN, N. (2004). *La structure des apparences*. Analyse et philosophie. Traduction française coordonnée par Jean-Baptiste Rauzy de Neslon Goodman, *The Structure of Appearance*, D. Reidel, 1977. Paris : Vrin.
GOODMAN, N. (2006). *Manières de faire des mondes (1978)*. Folio Essais. trad. fr. de M.-D. Popelard de *Ways of Worldmaking*, Hackett Publishing, Indianapolis : 1978. Paris : Gallimard.
GOODMAN, N. (2009). *L'art en théorie et en action*. Folio Essais. Paris : Gallimard.
GOODMAN, N. (2017). *The Structure of Appearance*. New Problems of Philosophy. Routledge.

GOODMAN, N. & ELGIN, C. (1994). *Reconceptions en philosophie dans d'autres arts et d'autres sciences [1988]*. L'interrogation philosophique. Traduction française de J.-P. Cometti et R. Pouivet. Paris : Presses Universitaires de France.

GOODMAN, N. & QUINE, W. v. O. (1947, décembre). Steps Toward a Constructive Nominalism. *The Journal of Symbolic Logic*, *12*(4), 105–122.

GOURINAT, J.-B. (2002). Éternel retour et temps périodique dans la philosophie stoïcienne. *Revue philosophique de la France et de l'étranger*, *127*(2), 213. doi :10.3917/rphi.022.0213

GRAILLOT, L. & BADOT, O. (2005, avril). Marketing méditerranéen ou marketing hyperréel : tentative d'éclaircissement à partir du cas de l'Occitane. In *Journée d'Etude " Marketing Méditerranéen : Pensée Méditerranéenne Appliquée au Marketing et/ou Marketing Appliquée dans l'Espace Méditerranéen ? "* Milan, Italie. Récupérée à partir de https://hal.archives-ouvertes.fr/hal-00482155

GRANGER, G. G. (1995). *Le probable, le possible et le virtuel*. Paris : Odile Jacob.

GRAUKROGER, S. (1995). *Descartes : an Intellectual Biography*. Oxford : Clarendon Press.

GREEN, A. (1980). *Le double double. Ceci et cela.* Préface au *Double* de Dostoïevski. Folio. Paris : Gallimard.

GRODDECK, G. (1976). *Le livre du ça*. Tel. Paris : Gallimard.

GUEROULT, M. (1968). *Descartes selon l'ordre des raisons. 1 l'âme et dieu*. Éditions Montaigne. Paris : Aubier.

GUICHARD, C. (2008). La signature dans le tableau aux xviie et xviiie siècles : identité, réputation et marché de l'art. *Sociétés et Représentations*, *1*(25), 47–77.

GUSDORF, G. (1991). *Les écritures du moi. Lignes de vie 1*. Paris : Éditions Odile Jacob.

HABER, S. & STORNETTA, W. (1991). How to Time-Stamp a Digital Document. *Journal of Cryptology*, *3*, 99–111. doi :https://doi.org/10.1007/BF00196791

HACKING, I. (1975). The Identity of Indiscernables. *Journal of Philosophy*, *9*(72), 249–256.

HARTSHORNE, C. (1965). *Anselm's Discovery*. La Salle Open Court Publishing Company.

HAWLEY, K. (2009). Identity and Indiscernibility. *Mind*, (118), 101–119.

HEIDEGGER, M. (1957). Le principe d'identité. In *Question i et ii*. tel. Gallimard.

HEIDEGGER, M. (1962). L'origine de l'œuvre d'art. In *Chemins qui ne mènent nulle part (1949)*. Tel. Paris : Gallimard.

HEIL, J. (2011). *Du point de vue ontologique*. Traduction française de Dominique Berlioz et François Loth. Paris : Ithaque.

HEINICH, N. (1998). *Le triple jeu de l'art contemporain*. Paris : Éditions de Minuit.

HEINICH, N. (1999). Art contemporain et fabrication de l'inauthentique. *Terrain*, (33).

HEINICH, N. (2008). La signature comme indicateur d'artification. *Sociétés et représentations*, (25), 97–106.

HEINTZ, J. (1979). On Reference and Inference in Fiction. *Poetics*, *8*, 85–99.

HERDMAN, J. (1990). *The Double in Nineteeth-Century Fiction*. Basingstoke : Macmillan.

HESIODE. (1928). *Théogonie, Les Travaux et les Jours, Bouclier*. Texte établi par Paul Mazon. Paris : Les Belles Lettres.

HINTIKKA, J. (1962). *Knowledge and Belief. An Introduction to the Logic of the Two Notions*. Ithaca : Comell.

HINTIKKA, J. (1968). Cogito, Ergo Sum : Inference or Performance ? *The Philosophical Review*, *71*(1), 3–32.

HINTIKKA, J. (1969a). *Models for Modalities*. Dordrecht : D. Reidel.
HINTIKKA, J. (1969b). Semantics for Propositional Attitudes. In J. W. DAVIS, D. J. HOCKNEY & W. K. WILSON (Éds.), *Philosophical logic*. Dordrecht : Reidel.
HINTIKKA, J. (1970). The Semantics of Modal Notions and the Indeterminacy of Ontology. *Synthese*, *21*, 408–424.
HIRSCH, E. (1992). *The Concept of Identity*. Oxford : Oxford University Press.
HIRSCH, E. (2011). *Quantifier variance and realism. essays in metaontology*. Oxford : Oxford University Press.
HIRST, R. (1959). *The Problems of Perception*. London : Allen et Unwin.
HOBBES, T. (1655). *Elementorum philosophiae sectio prima de corpore*. Londres : excusum sumptibus Andrea Crook.
HODGES, W. (2004). Elementary Predicate Logic. In F. G. (DOV M. GABBAY (Éd.), *Handbook of philosophical logic* (2nd, T. Volume 1). Springer.
HOMÈRE. (1955a). *Iliade*. bibliothèque de la Pléiade. Gallimard.
HORACE. (1944). Art poétique. In *Horace. œuvres complètes. satires, épitres, art poétique* (T. 2). Traduction de François Richard. Paris : Librairie Garnier Frères.
HOUDÉ, O. (2004). *La psychologie de l'enfant*. Que sais-je ? Paris : Presses Universitaires de France.
HOWELL, R. (1979). Fictional Objects : How They are and how They aren't. *Poetics*, *8*, 129–177.
HUME, D. (1968a). *A treatise of human nature* (2nd). Oxford : Clarendon Press.
HUME, D. (1968b). *Traité de la nature humaine tome i*. trad. A. Leroy. Paris : Aubier-Montaigne.
HUNT, A. E. (1899). Ethnographical note on the Murray Island, Torres Straits. *The Journal of the Anthropological Institute of Great Britain and Ireland*, *28*(1/2), 5–19.
HUSSERL, E. (1996). *Méditations cartésiennes. Introduction à la phénoménologie*. Traduction de l'allemand par Gabrielle Peiffer, et Emmanuel Lévinas. Première édition 1947. Paris : Librairie philosophique Vrin.
IFRAH, G. (1981). *Histoire universelle des chiffres*. Paris : Robert Laffont.
INGARDEN, R. (2013). *The Controversy over the Existence of the World*. Traductionanglaise d'Arthur Szylewicz. Frankfurt am Main : Peter Lang GmbH.
INWAGEN, P. (1981). The Doctrine of Arbitrary Undetached Parts. *Pacific Philosophical Quarterly*, *62*, 123–137.
INWAGEN, P. (1990). *Material beings*. New York : Cornell University Press.
INWAGEN, P. V. (1986). Two Concepts of Possible Worlds. *Midwest Studies In Philosophy*, *11*(1), 185–213. doi :10.1111/j.1475-4975.1986.tb00494.x.
ISER, W. (1993). *The Fictive and the Imaginary : Charting Literary*. Baltimore : Johns Hopkins University Press.
JACKOBSON, R. (1963). *Essai de linguistique générale*. Paris : Éditions de Minuit.
JACOB, F. (1981). *Le jeu des possibles*. Paris : Fayard.
JACQUES, F. (1973). Référence et description chez Meinong. De la phénoménologie à l'analyse. *Revue internationale de philosophie. Meinong*(104-105), 266–287.
JAGO, M. (2014). *The Impossible. An Essay on Hyperintensionality*. Oxford University Press.
JAMES, M. R. (1924). Introduction. In *Ghosts and Marvels. A Selection of Uncanny Tales from Daniel Defoe to Algernon Blackwood Selected by V. H. Collins World's Classics*. Oxford University Press.

JAMES, W. (1920). *The Letters of William James*. Boston : The Atlantic Monthly Press.
JANZEN, D. H. (1977). What Are Dandelions and Aphids ? *The American Naturalist*, *111*(979), 586–589.
JAŚKOWSKI, S. (1948). Rachunek zdan dla systemow dedukcyjnych sprzecznychi. *Studia Societatis Scientiarum Torunensis*, 55–77.
JEFFREY, R. C. (1990). *The Logic of Decision*. University Of Chicago Press.
JENTSCH, E. (1906). Zur Psychologie des Unheimlichen. *Psychiatrisch-Neurologische Wochenschrift*, *8*(22-23), 195–198. trad. en français par Bertrand Ferron, *L'inquiétante étrangeté et autres essais*, éditions Gallimard, Paris : 1985.
JOHNSTON, M. (1992). Constitution is not identity. *Mind*, *101*(401), 89–105.
JOURDE, P. & TORTONESE, P. (1996). *Visages du double, un thème littéraire*. Collection « fac » série « littérature ». Paris : Nathan.
JUNG, J. (2010). Du paradoxe identitaire au double transitionnel : le Horla de Guy de Maupassant. *Revue française de psychanalyse*, *74*(2), 507–519.
KANT, I. (1970). Du premier fondement de la différence des régions dans l'espace. In *Quelques opuscules précritiques*. Traduction française de Sylvain Zac. Vrin.
KANT, I. (1986). *Prolégomènes à toute métaphysique future qui pourra se présenter comme science (1783)*. Bibliothèque des textes philosophiques. Traduction française de Louis Guillermit. Paris : Librairie philosophique Vrin.
KANT, I. (2012). *Critique de la raison pure*. Traduction de A. Tremesaygues et B. Pacaud. Préface de C. Serrus. Presses Universitaires de France.
KAPLAN, D. (1975). How to Russell a Frege-Church. *Journal of Philosophy*, *72*, 716–729.
KAPLAN, D. (1979). Transworld Heirlines. In *The Possible and the Actual, Readings in the Metaphysics of Modality*. Cornell University Press.
KIRK, R. & ROGER, S. (1974). Zombies v. Materialists. *Proceedings of the Aristotelian Society*, *48*, 135–163.
KÖHLER, W. (1964). *Psychologie de la forme. Introduction à de nouveaux concepts en psychologie*. Folio essais. trad. de Serge Bricianer. Gallimard.
KOŁAKOVSKI, L. (1969). *Toward a marxist humanism*. New York : Grove Press.
KOVACSHAZY, C. (2012). *Simplement double. Le personnage double, une obsession du roman au xxe siècle*. Perspectives Comparatistes n°14. Paris : Classiques Garnier.
KRAUS, J. (1927). *Die Lehre des Johannes Duns Skotus O.F.M., von der Natura Communis ; Ein Beitrag zum Universalienproblem in der Scholastik*. Studia Friburgensia.
KRIPKE, S. (1963). Semantical Considerations on Modal Logic i Normal Modal Propositional Calculi. *Acta Philosophica Fennica*, *16*, 83–94.
KRIPKE, S. (1965). Semantical analysis of modal logic ii : non-normal modal propositional calculi. In *The theory of models* (p. 206–20). Amsterdam, North-Holland : Addison et al.
KRIPKE, S. (1982). *La logique des noms propres*. Trad. fr. de Pierre Jacob et François Recanati de *Naming and Necessity*, D. Reidel Publishing Co., USA, 1972. Paris : Éditions de Minuit.
KRIPKE, S. (2011). Vacuous Names and Fictional Entities. In *Philosophical Troubles, Collected Papers* (T. 1). Oxford University Press.
KRIPKE, S. (2015). *Reference and Existence. The John Locke Lectures*. Oxford University Press.
KUHN, T. (1983). *La structure des révolutions scientifiques (1962)*. Champs. Paris : Flammarion.
KUNDERA, M. (1987). *L'insoutenable légèreté de l'être*. Du monde entier. Paris : Gallimard.

KUNG, P. (2010). Imagining as a guide to possibility. *Philosophy and Phenomenological Research*, *81*(3), 620–663.
LAËRCE, D. (1999). *Vies, doctrines et sentences des philosophes illustres*. Paris : Le livre de poche.
LAING, R. D. (1979). *Le moi divisé*. Pluriel. Traduit de l'anglais par Claude Elsen. Paris : Stock.
LEIBNIZ, G. W. [G. W.]. (1903). *Opuscules et fragments inédits de Leibniz*. Édité par L. Couturat. Félix Alcan.
LEIBNIZ, G. W. [Gottfried Wilhelm]. (1839). Lettre du 2 juin 1716 á clarke. In *Opera philosophica*. Berolini.
LEIBNIZ, G. W. [Gottfried Wilhelm]. (1969). *Essais de théodicée*. Paris : Garnier Flammarion.
LEIBNIZ, G. W. [Gottfried Wilhelm]. (1989). Lettre à arnauld, 14 juillet 1686. In *Discours de métaphysiques et correspondance avec arnauld*. Paris : Librairie Philosophique Vrin.
LEIBNIZ, G. W. [Gottfried Wilhelm]. (1993). *Confessio philosophi la profession de foi du philosophe*. Paris : Librairie philosophique Vrin.
LEIBNIZ, G. W. [Gottfried Wilhelm]. (1995). Monadologie. In *Discours de métaphysique suivi de monadologie*. Tel. Paris : Gallimard.
LEIBNIZ, G. W. v. & CLARKE, S. (2000). *G. W. Leibniz and Samuel Clarke : Correspondence*. Hackett Publishing Company.
LENAIN, T. (2013). Le faux en art. *CeROArt*. Récupérée à partir de http://ceroart.revues.org/2947
LEOPARDI, G. (2000). *La théorie du plaisir*. Édition établie par Giorgio Panizza, traduit de l'italien par Joël Gayraud. Paris : Alia.
LEPELTIER, T. (2010). *Univers parallèles*. Science Ouverte. Paris : Éditions du Seuil.
LEWIS, D. (1968). Counterpart Theory and Quantified Modal Logic. *The Journal of Philosophy*, *65*, 113–126.
LEWIS, D. (1970). Anselm and Actuality. *Noûs*, *4*(2), 175–188.
LEWIS, D. (1971). Counterparts of Persons and Their Bodies. *The Journal of Philosophy*, *68*(7), 203–211.
LEWIS, D. (1973). *Counterfactuals*. Cambridge : Havard University Press.
LEWIS, D. (1979). Attitudes De Dicto and De Se. *Philosophical Review*, *88*(4), 513–543.
LEWIS, D. (1983a). Extrinsic Properties. *Philosophical Studies*, *44*, 197–200.
LEWIS, D. (1983b). Individuation by Acquaintance and by Stipulation. *The Philosophical Review*, *92*(1).
LEWIS, D. (1983c). New work for a theory of universals. *Australasian Journal of Philosophy*, *61*, 343–377.
LEWIS, D. (1988). Vague Identity : Evans Misunderstood. *Analysis*, *48*, 128–130.
LEWIS, D. (1991). *Parts of Classes*. Oxford : Blackwell.
LEWIS, D. (1993). Many, but Almost One. In K. CAMBELL, J. BACON & L. REINHARDT (Éds.), *Ontology, Causality and Mind : Essays on the Philosophy of D. M. Armstrong* (p. 23–38). Cambridge University Press.
LEWIS, D. (2007). *De la pluralité des mondes*. Traduction française de M. Caveribère et J.-P. Cometti de *On the plurality of Worlds*, Basil Blackwell Publishing, Oxford : 1986. Éditions de l'Éclat.
LEWIS, D. (2012). La vérité dans la fiction (1978). *Revue Klesis*, *La philosophie de David Lewis*(24), 36–55. Traduction de Yann Schmidt de « Truth in Fiction », American Philosophi-

cal Quarterly, 1978, 15 : 37–46 ; texte repris in D. Lewis, Philosophical Papers, Volume I, Oxford, Oxford University Press, 1983.

LIENHARDT, G. (1961). *Divinity and Experience : The Religion of the Dinka*. Oxford University Press.

LOCKE, J. (1998a). *An Essay Concerning Human Understanding (1689)* (2nd). Penguin Classics.

LOCKE, J. (1998b). *An Essay Concerning Human Understanding (1689). livre ii, chapitre xxvii de Identity and Diversity*. traduction française *Identité et Différence. L'invention de la conscience*, traduit par Etienne Balibar. Seuil.

LOCKE, J. (1998c). *Identité et différence* (TraductionÉtienne Balibar). Collection points. Éditions du Seuil.

LOCKE, J. (2001). *Essai concernant l'entendement humain (1694). livres i et ii*. Traduction française de J.-M. Vienne. Paris : Vrin.

LONG, A. & SEDLEY, D. [D.N.]. (1987). *The Hellenistic Philosophers, vol. 1 : Translations of the Principal Sources, with Philosophical Commentary*. Textes grecs et latins avec note et bibliographie, 2 volumes. Cambridge : Cambridge University Press.

LORD, A. B. (1971). *The Singer of Tales*. New York : Atheneum.

LORRAIN, J. (1900). *Histoires de masques*. Librairie Paul Ollendorff.

LOUX, J. M. (1978). *Substance and attribute. a study in ontology*. Philosophical Studies Series in Philosophy 14. Netherlands : Springer.

LOWE, E. J. (2011). Vagueness and Metaphysics. In G. RONZITTI (Éd.), *Vagueness : A Guide*. Logic, Epistemology, and the Unity of Science 19. Springer.

LOWE, J. (2003). Individuation. In M. J. LOUX & Z. D. W. (Éds.), *The Oxford handbook of metaphysics*. Oxford : Oxford University Press.

LUCRÈCE. (1978). De la nature. (T. 1). Collection des universitaires de France publiée sous le patronage de l'Association Guillaume Budé. Texte établi et traduit par Alfred Ernout. Les Belles Lettres.

MACH, E. (1991a). *Analyse des sensations*. Éditions Jacqueline Chambon.

MACH, E. (1991b). *L'analyse des sensations*. trad. F Eggers and Monnoyer, J.-M. Nîmes : J. Chambon.

MACKIE, P. (1987). *How Things Might Have Been : Individuals, Kinds and Essential Properties*. Oxford : Oxford University Press.

MACKIE, P. & JAGO, M. (2017). Transworld Identity. In E. ZALTA (Éd.), *The Stanford Encyclopedia of Philosophy* (Winter 2017). Metaphysics Research Lab, Stanford University.

MACNAMARA, J., REYES, G. & LA PALME REYES, M. (1994). Reference, Kinds and Predicates.

MAITTE, C. (2015). Imitation, copie, contrefaçon, faux : définitions et pratiques sous l'ancien régime. *ESKA, Entreprises et histoire*, *1*(78), 13–26.

MALLY, E. (1912). *Gegenstands theoretische Grundlagender Logik und Logistik*. Leipzig : Barth, J.-A.

MALRAUX, A. (1996). *Le musée imaginaire (1965)*. Folio Essais. Paris : Gallimard.

MARSHALL, D. & WEATHERSON, B. (2018). Intrinsic vs. Extrinsic Properties. In E. ZALTA (Éd.), *The Stanford Encyclopedia of Philosophy* (Spring 2018). Metaphysics Research Lab, Stanford University.

MARTINIÈRE, N. (2008). *Figures du double. Du personnage au texte*. Rennes : Presses Universitaires de Rennes.

MARX, K. (2007). *Le 18 brumaire de louis bonaparte*. GF Philosophie. Paris : Flammarion.

MEAD, G. H. (1925). The Genesis of the Self and Social Control. *International Journal of Ethics*, *35*(3), 251–277.

MEAD, G. H. (1934). *Mind, Self and Society. From the Standpoint of a Social Behaviorist*. Chicago : The University of Chicago Press.

MEINONG, A. (1899). *Über Gegenstände höherer Ordnung und deren Verhältnis zur inneren Wahrnehmung, Zeitschrift für Psychologie und Physiologie der Sinnesorgane xxi : 182-272*. Réimprimé in Meinong 1968, Vol. II : 377-480. Traduction anglaise 'On Objects of Higher Order and Their Relationship to Internal Perception' par Schubert Kalsi, 1978, 137-208.

MEINONG, A. (1910). *Über Annahmen (1902)* (2e éd.). Edition revisée de celle de 1902. Leipzig : J. A. Barth.

MEINONG, A. (1915). *Über Möglichkeit und Wahrscheinlichkeit. Beiträge zur Gegenstandstheorie und Erkenntnistheorie*. trad. anglaise 'On Possibility and Probability. Contributions to Object Theory and Epistemology'. Leipzig : J. A. Barth.

MEINONG, A. (2000). *Théorie de l'objet et présentation personnelle*. Paris : Librairie philosophique Vrin.

MÉREDIEU, F. d. (2003). *Arts et nouvelles technologies, art vidéo et art numérique*. In Extenso. Larousse.

MERLEAU-PONTY, M. (1945). *Phénoménologie de la perception*. Collection tel. Paris : Gallimard.

MERLEAU-PONTY, M. (1964). *L'œil et l'Esprit*. Folio essais. Paris : Gallimard.

MILGRAM, P. & KISHINO, F. (1994, décembre). A taxonomy of mixed reality visual displays. *IEICE Trans. Information Systems*, *vol. E77-D, no. 12*, 1321–1329.

MILL, J. S. (1904). *A System of Logic*. New York : Harper et Row.

MILL, J. S. M. (1866). *Système de logique déductive et inductive. exposé des principes de la preuve et des méthodes de recherche scientifique (1843)*. Traduit de la sixième édition anglaise (1865) par Louis Peisse. Librairie philosophique de Ladrange.

MILLER, K. (1985). *Double. Studies in Literary History*. Oxford : Oxford University Press.

MILLIKAN, R. (1997). Images of Identity : In Search of Modes of Presentation. *Mind*, *106*(423), 499–519.

MINSKY, M. (1965). Matter, Mind and Models. *Proc. of the IFIP Congress*, 45–49.

MIYOSHI, M. (1969). *The Divided Self : A Perspective on the Literature of the Victorians*. New York : New York University Press.

MONOD, J. (1970). *Le hasard et la nécessité*. Éditions du Seuil.

MONTAIGNE, M. d. (1928a). Les essais. (T. 3). Publiés sur l'exemplaire de Bordeaux avec une introduction de Fortunat Strowski. Paris : Chronique des lettres françaises.

MONTAIGNE, M. d. (1928b). Les essais. (T. 6). Publiés sur l'exemplaire de Bordeaux avec une introduction de Fortunat Strowski. Paris : Chronique des lettres françaises.

MOORE, G. (1922). The Conception of Intrinsic Value. In *Philosophical studies* (p. 256–275). Traduction française dans G.E. Moore, *Principia Ethica*, Paris, PUF, 1998, p. 307-328. Londres : Routledge.

MORI, M. (2012). La vallée de l'étrange. *Gradhiva, revue d'anthropologie et d'histoire des arts*, *15 : Robots étrangement humains*, 27–33.

MORIZOT, J.-P. (1999). *Sur le problème de Borges. Sémiotique, ontologie, signature*. Éditions Kimé.

MYRO, G. (1997). Identity and time. In M. C. REA (Éd.), *Material constitution*. Rowman & Littlefield.

NEF, F. (2006). *Les propriétés des choses expérience et logique*. Paris : Librairie philosophique Vrin.

NIETZSCHE, F. (1965). *La volonté de puissance*. NRF. Traduction française de Geneviève Blanquis. Paris : Gallimard.

NIETZSCHE, F. (1968). *Par-delà bien et mal*. Traduction de Conélius Ham. Paris : Gallimard.

NIETZSCHE, F. (1974a). L'antéchrist. In *Friedrich nietzsche œuvres philosophiques complètes tome vii*. nrf. Textes et variantes établis par G. Colli et M. Montinari. Traduit de l'allemand par Jean-Claude Hémery. Paris : Gallimard.

NIETZSCHE, F. (1974b). Le crépuscule des idoles. In *Friedrich Nietzsche œuvres philosophiques complètes tome vii*. nrf. Textes et variantes établis par G. Colli et M. Montinari. Traduit de l'allemand par Jean-Claude Hémery. Paris : Gallimard.

NIETZSCHE, F. (1975). *La philosophie tragique à l'époque des Grecs*. Folio. Gallimard.

NIETZSCHE, F. (1982). Le gai savoir « la gaya scienza » - fragments posthumes été 1881 – été 1882. In G. COLLI & M. MONTINARI (Éds.), *Œuvres philosophiques complètes v*. Traduction de Pierre Klossowski. Édition revue, corrigée et augmentée par Marc B. de Launay. Paris : Gallimard.

NIETZSCHE, F. (1988). Humain, trop humain. un livre pour esprits libres. In G. COLLI & M. MONTINARI (Éds.), *Œuvres philosophiques complètes iii*. Traductionpar Robert Rovini. Édition revue par Marc B. de Launay. Paris : Gallimard.

NIINILUOTO, I. (1985). Imagination and fiction. *Journal of semantics*, *4*(209-222).

NINAN, D. (2015). On recanati's mental files. *Inquiry*, *58*(4), 368–377. doi :10.1080/0020174X.2014.883751

NOLL, M. (1967). The Digital Computer as a Creative Medium. *IEEE SPECTRUM*, *4*(10), 89–95.

NOLT, J. E. (1986). What Are Possible Worlds ? *Mind*, *95*(380), 432–445.

NYBERG, H. S. (1931). Questions de cosmogonie et de cosmologie mazdéennes. In *Journal asiatique. receuil trimestriel de mémoire et de notices relatifs aux études orientales, tome 219*. Société asiatique.

O'LEARY HAWTHORNE, J. (1995). *The bundle theory of substance and the identity of indiscernibles*.

OLSON, E. T. (2017). *Que sommes-nous ? Sur la nature métaphysique des personnes*. Paris : Ithaque.

OWEN, R. D. (1945). Immunogenetic Consequences of Vascular Anastomoses between Bovine Twins. *Science*, *102*, 400–405.

PAGIN, P. (2013). The cognitive Significance of Mental files. *Disputatio*, *5*(36).

PARFIT, D. (1984). *Reasons and Persons*. Oxford : Clarendon Press.

PARFIT, D. (2015). L'identité personnelle. In *Identité et survie*. Traduit de l'anglais par Rémi Clot-Goudard, Vanessa Di Marino, Stéphane Dunand et Mathieu Mulcey. Ithaque.

PARMÉNIDE. (1986). *Le poéme*. Collection épiméthée. PUF.

PARSONS, T. (1980). *Nonexistent Objects*. New Heaven, Yale University Press.

PARSONS, T. (1982). Are There Nonexistent Objects ? *American Philosophical Quarterly*, *19*(4), 365–371.
PASCAL. (1950). Pensées. In *L'œuvre de Pascal*. Bibliothèque de la pléiade. Texte établi et annoté par Jacques Chevalier. Paris : Gallimard.
PASCAL, B. (1950). Les Pensées. In *Œuvres*. Bibliothèque de la Pléiade. Paris : Gallimard.
PASTOUREAU, M. (1988). Le faux n'existe plus. In *Vrai ou faux ? copier, imiter, falsifier*. Paris : Bibliothèque nationale de France.
PAVEL, T. (1988). *Univers de la fiction*. Paris : Éditions du Seuil.
PEACOCKE, C. (1985). Imagination, Experience, and Possibility. In J. FOSTER & H. ROBINSON (Éds.), *Essays on Berkeley : A Tercentennial Celebration*. Oxford University Press.
PÉGUY, C. (1916). Œuvres complètes de Charles Péguy. (T. 8). Paris : Nouvelle Revue Française.
PERRY, J. (1977). Frege on Demonstratives. *Philosophical Review*, *86*(4), 474–497.
PERRY, J. (1979). The Problem of the Essential Indexical. *Nous*, *13*(1), 3–21.
PIAGET, J. (1968). *Le structuralisme*. Paris : PUF.
PIAGET, J. (1970). *L'épistémologie génétique*. Que sais-je ? Presses Universitaire de France.
PIAGET, J. & SZEMINSKA, A. (1997). *La genèse du nombre chez l'enfant (1941)*. Delachaux et Niestlé.
PIRANDELLO, L. (1982). *Un, personne et cent mille (1927)*. L'imaginaire. Paris : Gallimard.
PLANTINGA, A. (1974). *The Nature of Necessity*. Oxford : Oxford University Press.
PLANTINGA, A. (1976). Actualism and Possible Worlds. *Theoria*, 139–160.
PLATON. (1950a). Alcibiade. In *Œuvres complètes tome i*. Bibliothèque de la Pléiade. Paris : Gallimard.
PLATON. (1950b). La république. In *Œuvres complètes tome i*. Bibliothèque de la Pléiade. Paris : Gallimard.
PLATON. (1950c). Le timée. In *Œuvres complètes ii*. Bibliothèque de la Pléiade. Paris : Gallimard.
PLATON. (1955). *Œuvres complètes*. Bibliothèque de la Pléiade. Paris : Gallimard.
PLATON. (1959). Cratyle. In *Œuvres complètes*. Bibliothèque de la Pléiade. Paris : Gallimard.
PLESSNER, H. (1960). Soziale Rolle und menschliche Natur. In J. DERBOLAV & F. NICOLIN (Éds.), *Erkenntnis und Verantwortung : Festschrift für Theodor litt. Herausgegeben*. Düsseldorf : Pädagogischer Verlag Schwann.
PLESSNER, H. (1974). Soziale Rolle und menschliche Natur. In *Diesseits der utopie* (p. 23–35). Frankfurt a/M : Suhrkamp.
PLESSNER, H. (2003). Die Grenzen der Gemeinschaft. Eine Kritik des sozialen Radikalismus. In *Gesammelte Schriften. Band v*. Frankfurt am Main : Suhrkamp.
PLUTARQUE. (1937). Vie de Thésée. In trad. JACQUES AMYOT (Éd.), *Les vies des hommes illustres*. Bibliothèque de la Pléiade. Paris : Gallimard.
POE, E. A. (1932). Eureka (1848). In *Œuvres en prose. traduction de charles baudelaire*. Bibliothèque de la Pléiade. Paris : Gallimard.
POUIVET, R. (2010). *Philosophie du rock*. L'interrogation philosophique. Paris : Presses Universitaires de France.
POUIVET, R. (2012). L'ontologie de régime de Gérard Genette. *Aisthesis. Pratiche, linguaggi e saperi dell'estetico*, *4*(1). Récupérée à partir de http://www.fupress.net/index.php/aisthesis/article/view/10981

POUIVET, R. (2016, octobre 6). Les œuvres d'art existent-elles ? *Rencontres philosophiques de langres 2016*, Langres : Canal U. Récupérée à partir de https://www.canal-u.tv/video/eduscol/les_uvres_d_art_existent_elles_rpl.46839

PRADEU, T. (2008). Qu'est-ce qu'un individu biologique ? In P. LUDWIG & T. PRADEU (Éds.), *L'individu. Perspectives contemporaines*. Paris : Vrin.

PRADEU, T. (2009). La mosaïque du soi : les chimères en immunologie. *Bulletin d'histoire et d'épistémologie des sciences de la vie*, *16*(1), 19–27.

PRICE, M. S. (1977, avril). Identity Through Time. *The Journal of Philosophy*, *74*(4). doi :10.2307/2025438

PRIEST, G. (1995). *Beyond the limits of thought*. Cambridge University Press.

PRIEST, G. (2005). *Towards Non-Being : The Logic and Metaphysics of Intentionality*. Oxford : Oxford University Press.

PRIETO, L. J. (1975). *Pertinence et pratique*. Le sens commun. Paris : Editions de Minuit.

PRIETO, L. J. (1992). Le mythe de l'original. In *Esthétique et poétique* (p. 131–156). Coll. « Points Essais ». Paris : Éditions du Seuil.

PRIOR, A. N. (1971). *Objects of Thought*. Oxford : Oxford : Clarendon Press.

PROUST, J. (1996). Identité personnelle et pathologie de l'action. In I. JOSEPH & J. PROUST (Éds.), *La folie dans la place* (T. 7, p. 155–176). Raisons pratiques. Paris : Éditions de l'École des hautes études en sciences sociales.

PRYOR, J. (2016). Mental Graphs. *Review of Philosophy and Psychology*, *7*(2), 309–341. doi :10.1007/s13164-015-0280-1

PUTNAM, H. (1975). The meaning of meaning. In *Mind, Langage and Reality. Philosophical Papers vo. ii* (p. 215–271). Gunderson, Keith.

PUTNAM, H. (1984). *Raison, vérité et histoire (1981)*. Propositions. Paris : Editions de Minuit.

PUTNAM, H. (1985). Signification, référence et stéréotypes. *Philosophie*, (5), 21–44.

PUTNAM, H. (1987). Truth and Convention : On Davidson's Refutation of Conceptual Relativism. *Dialectica*, *41*, 69–77.

QUINE, W. V. O. (1951a). On Carnap's View on Ontology. *Philosophical Studies*, *2*(5), 65–72.

QUINE, W. V. O. (1951b). Two Dogmas of Empiricism. In *From a Logical Point of View*. Cambridge : Havard University Press.

QUINE, W. V. O. (1977). *Relativité de l'ontologie et autres essais*. Philosophie. Paris : Aubier.

QUINE, W. V. O. (1992). *Quiddités dictionnaire philosophique par intermittence*. L'ordre philosophique. Paris : Éditions du Seuil.

QUINE, W. V. O. (2003a). De ce qui est. In *Identité, ostension et hypostase*. Traduction sous la dir. de Sandra Laugier. Paris.

QUINE, W. V. O. (2003b). De ce qui est (1948). In *Du point de vue logique. Neuf essais logico-philosophiques*. Traduction française de J. Vidal-Rosset, S. Laugier et Ph. de Rouilhan de « On what there is », in Willard van Orman Quine, *From a Logical Point of View. Nine Logico-Philosophical Essays*, Harper and Row Publishers, New York, 1953, p. 8-9. Paris : Vrin.

QUINE, W. v. O. (1976a). Grades of Discrimination. *The Journal of Philosophy*, *73*(5).

QUINE, W. v. O. (1976b). Worlds Away. *The Journal of Philosophy*, *73*, 859–863.

QUINE, W. v. O. (1980). *From a Logical Point of View, Logico-Philosophical Essays (first ed. 1953)*. New York : Harper Torchbooks.

RABACHOU, J. (2016). *Qu'est-ce qu'un monde ?* Chemins philosophiques. Paris : Vrin.
RAMSEY, F. P. (1990). Universals (1925). In D. MELLOR (Éd.), *Philosophical papers (1990)* (p. 31–33). Cambridge : Cambridge University Press.
RANK, O. (1973). *Don Juan et le Double (1925)*. Petite Bibliothèque Payot. Paris : Payot.
RANTALA, V. (1989). Impossible World Semantics and Logical Omniscience. *Acta Philosophica Fennica*, *35*, 106–115.
REBUSCHI, M. (2000). *Peut-on dire ce qui n'est pas ? Objets mathématiques et autres fictions : sémantique et ontologie* (thèse de doct., Université de Nancy).
REBUSCHI, M. (2008). *Qu'est-ce que la signification ?* Chemins Philosophiques. Paris : Librairie Philosophique J. Vrin.
REBUSCHI, M. (2011). Le cogito sans engagement. *Igitur - Arguments philosophiques*, *3*(2), 1–25.
REBUSCHI, M. (2012). Représentations de soi et modalités. *Philosophia Scientiæ*, *16*(2).
REBUSCHI, M. (2018). Hintikka, la perception et ses objets. *Klesis*, *39*.
REBUSCHI, M., AMBLARD, M. & MUSIOL, M. (2013). Schizophrénie, logicité et perspective en première personne. *L'évolution psychiatrique*, *78*, 127–141.
REBUSCHI, M. & RENAULD, M. (2010). Fiction, Indispensability and Truths. In *Truth in fiction* (p. 245–286). Ontos Verlag.
RECANATI, F. (2000). *Oratio Obliqua, Oratio Recta : An Essay on Metarepresentation*. Cambridge : MIT Press/Bradford Books.
RECANATI, F. (2007a). Imagining de Se. *Presented at 'mimesis, metaphysics and make-believe' conference held in honour of Kendall Walton*, the University of Leeds.
RECANATI, F. (2007b). *Perspectival Thought. A Plea for (Moderate) Relativism*. New York : Oxford University Press.
RECANATI, F. (2015). Replies. *Inquiry*, *58*(4), 408–437. doi :10.1080/0020174X.2015.1033868
RECANATI, F. (2018). Qu'est-ce qu'un personnage de fiction ? Récupérée à partir de http://savoirs.ens.fr/expose.php?id=3354
REID, T. (2011). *Essays on the Intellectual Powers of Man (1785)*. Cambridge University Press.
RENAULD, M. (2014). *Philosophie de la fiction. vers une approche pragmatiste du roman*. Æsthetica. Rennes : Presses Universitaires de Rennes.
RESCHER, N. & BRANDOM, R. (1980). *The logic of inconsistency : a study in non-standard possible worlds semantics and ontology*. Oxford : Basil Blackwell.
REY, A. (1927). *Le retour éternel et la philosophie de la physique*. Paris : Librairie Flammarion.
RICŒUR, P. (1954). Étude sur les *Méditations Cartésiennes* de husserl. *Revue Philosophique de Louvain*, *52*(33), 75–109.
RICŒUR, P. (1985). *Temps et récit III. le temps raconté*. Collection Points Essais. Paris : Éditions du Seuil.
RICŒUR, P. (1990). *Soi-même comme un autre*. Collection Points Essais. Paris : Éditions du Seuil.
RIFFATERRE, M. (1978). *La sémiotique de la poésie*. Paris : Éditions du Seuil.
RIFFATERRE, M. (1982). L'illusion référentielle. In *Littérature et réalité* (p. 91–118). Paris : Éditions du Seuil.
RODRIGUEZ-PEREYRA, G. (2004). The bundle theory is compatible with distinct but indiscernible particulars. *Analysis*, *64*(1), 72–81.
RODRIGUEZ-PEREYRA, G. (2014). *Leibniz's Principle of Identity of Indiscernibles*. Oxford : Oxford University Press.

ROGERS, R. (1970). *A Psychoanalitic Study of the Double in Literature*. Détroit : Wayne State University Press.
ROSEN, G. (1990). Modal Fictionalism. *Mind*, *99*(395), 327–354.
ROSEN, G. & DORR, C. (2002). Composition as a Fiction. In R. GALE (Éd.), *The blackwell companion to metaphysics* (p. 151–174). Blackwell.
ROSENBERG, J. (1978). *The Practice of Philosophy*. Englewood Cliffs : Prentice Hall.
ROSKILL, M. W. (1989). *What is art History ?* (2nd). University of Massachusetts Press.
ROSSET, C. (1973). *L'anti-nature*. Paris : Presses Universitaires de France.
ROSSET, C. (1978). *Le réel, traité de l'idiotie*. Paris : Éditions de Minuit.
ROSSET, C. (1979). *L'objet singulier*. Paris : Éditions de Minuit.
ROSSET, C. (1983). *La force majeure*. Collection « Critique ». Paris : Éditions de Minuit.
ROSSET, C. (1993). *Logique du pire : éléments pour une philosophie tragique (1971)*. Quadrige. Presses Universitaires de France-PUF.
ROSSET, C. (1997). *Le démon de la tautologie, suivi de cinq pièces morales*. Paris : Éditions de Minuit.
ROSSET, C. (1999). *Le réel, l'imaginaire et l'illusoire*. Biarritz : Distance.
ROSSET, C. (2004). *Impressions fugitives : l'ombre, le reflet, l'écho*. Paris : Minuit.
ROSSET, C. (2006). *Fantasmagories, suivi de le réel, l'imaginaire et l'illusoire*. Paris : Minuit.
ROSSET, C. & BAYARD, P. (2015). Pierre Bayard, Clément Rosset. Les visiteurs du soi. *Philosophie magazine*, (91).
ROUSSEAU. (1973). *Les confessions*. Folio Classique. Paris : Gallimard.
ROUTLEY, R. (1980). Exploring Meinong's Jungle and Beyond : An Investigation of Noneism and the Theory of Items, *Departmental Monograph Canberra : Australian National University*, *3*.
RUSSELL, B. (1905). On Denoting. *Mind*, *14*(4), 479–493.
RUSSELL, B. (1910). Knowledge by Acquaintance and Knowledge by Description. *Proceedings of the Aristotelian Society*, *11*, 108–128.
RUSSELL, B. (1912). On the Relation of Particulars and Universals. *Proceedings of the Aristotelian Society, reprinting in Russell 1956, pp 105-124*.
RUSSELL, B. (1940). *An Inquiry into Meaning and Truth*. Trad. fr. de Philippe Devaux : *Signification et vérité*, Paris : Flammarion, 1993. London : Routledge.
RUSSELL, B. (1965). *Problèmes de philosophie*. Paris : Payot.
RUSSELL, B. (2010). *Principles of Mathematics (1903)*. Routledge classics. London : Routledge.
RUSSELL, B. & WHITEHEAD, A. N. (1997). *Principia Mathematica to 56, Second Edition* (2e éd.). Cambridge University Press.
RYAN, M.-L. (2006). Des mondes possibles aux univers parallèles. *Fabula, la recherche en littérature*.
RYLE, G. (1954). *Dilemmas*. Cambridge : Cambridge University Press.
SAGOFF, M. (1978). On Restoring and Reproducing Art. *The Journal of Philosophy*, *75*(9), 452–460.
SAINSBURY, R. M. & TYE, M. (2013). *Seven Puzzles of Thought and How to Solve Them an Originalist Theory of Concepts*. USA : Oxford University Press.
SAINSBURY, R. (2010). *Fiction and Fictionalism*. New Problems of Philosophy. Routledge.
SALMON, N. (1981). *Reference and Essence*. Princeton : Princeton University Press.

SARTRE, J.-P. (1940). *L'imaginaire*. Paris : Gallimard.
SARTRE, J.-P. (1976). *L'être et le néant. Essai d'ontologie phénoménologique*. Collection tel. Paris : Gallimard.
SARTRE, J.-P. (2012). *L'imagination [1936]*. Quadrige. Paris : Presses Universitaires de France.
SAUNDERS, S. (2006). Are Quantum Particles Objects ? *Analysis*, *66*, 52–63.
SCARRON. (1981). *Le roman comique*. Texte établi, présenté et annoté par Yves Giraud. Paris : GF-Flammarion.
SCHAEFFER, J.-M. (1989). *Qu'est-ce qu'un genre littéraire ?* Poétique. Paris : Éditions du Seuil.
SCHAEFFER, J.-M. (1999). *Pourquoi la fiction ?* Paris : Éditions du Seuil.
SCHEFFLER, I. (1954). An Inscriptional Approach to Indirect Quotation. *Analysis*, *14*(4), 83–90.
SCHMITT, A. (2010). *Je réel, je fictif. au delà d'une confusion postmoderne*. Essais de littérature Cribles. Toulouse : Presses universitaires du Mirail.
SCHNIEDER, B. & SOLODKOFF, T. v. (2009, janvier). In defence of fictional realism. *The Philosophical Quarterly*, *59*(234).
SCHOPENHAUER, A. (1966). *De la quadruple racine du principe de raison suffisante (1847)*. Bibliothèque des textes philosophiques. Traduction française de F.-X. Chenet. Librairie philosophique J. Vrin.
SCHUPPERT, G. (2019). *Le « faire-semblant » en procès. Examen et défense de la philosophie de Kendall L. Walton* (thèse de doct.). Thèse de doctorat de philosophie dirigée par Christophe Bouriau, Université de Lorraine. Récupérée à partir de http://www.theses.fr/2017LORR0238/document
SEARLE. (1982). *Sens et expression (1975)*. Paris : Éditions de Minuit.
SEARLE, J. (1958). Proper Names. *Mind*, *67*(266), 16–173.
SEARLE, J. (1998). *La construction de la réalité sociale*. nrf essais. Traduction française de *The Construction of Social Reality*, Free Press, New York, 1995. Paris : Gallimard.
SEDLEY, D. [David] & BRUNSCHWIG, J. (1989). Le critère d'identité chez les stoïciens (1982). *Revue de Métaphysique et de Morale*, *94 Recherches sur les stoïciens*(4), 513–533.
SELLARS, W. (1952). Particulars. *Philosophy and Phenomenological Research*, *13*(2), 184–199.
SERRES, M. (1968). *Hermès 1 la communication*. Paris : Éditions de Minuit.
SHOEMAKER, D. (1984). Personal Identity : A Materialist's Account. In S. SHOEMAKER & R. SWINBURNE (Éds.), *Personal Identity*. Oxford, Blackwell Publishing.
SHOEMAKER, S. (1963). *Self-Knowledge and Self-Identity*. Cornell University Press.
SHOEMAKER, S. (1968). Self-Reference and Self-Awareness. *The Journal of Philosophy*, *65*(19).
SHOEMAKER, S. (1970). *Persons and Their Past*. Reproduit in Identity, Cause and Mind, Cambridge, Cambridge University Press, 1984, p. 19–47.
SIDER, T. (2006). Bare Particulars. *Philosophical Perspectives*, *20*, 387–397.
SIDER, T. (2011). *Writing the Book of the World*. Oxford : Clarendon Press.
SMITH, B. & VARZI, A. C. (2000). Fiat and Bona Fide Boundaries. *Philosophy and Phenomenological Research*, *60*(2), 401–420. doi :10.2307/2653492
SOLLERS, P. (1968). *Le roman et l'expérience des limites*. Logiques. Paris : Éditions du Seuil.
SOMMEFELT, A. (1938). *La langue et la société. caractéres sociaux d'une langue de type archaïque*. Oslo : H. Aschehoug et Cie.
STALNAKER, R. (1976). Possible Worlds. *Noûs*, *1*(10), 65–75.
STALNAKER, R. (1981). Indexical belief. *Synthese*, *49*, 129–151.

STALNAKER, R. (2012). *Mere Possibilities : Metaphysical Foundations of Modal Semantics*. Princeton : Princeton University Press.

STEINER, G. (1998). *Après Babel. Une poétique du dire et de la traduction*. Bibliothèque de l'évolution de l'humanité. Trad. Lucienne Lotringer et Pierre-Emmanuel Dauzat. Paris : Albin Michel.

STÉPHAN, L. (1991). *Le vrai, l'authentique et le faux*. Éditions du Musée national d'art moderne, Centre Georges Pompidou.

STEVENSON, C. L. (1957). What is a poem ? *The Philosophical Review, Vol. 66*(3), 337.

STOBÉE. (1884). Joannis stobaei anthologium. In O. H. WACHSMUTH (Éd.). Édition en 5 volumes. Berlin : Weidmann.

STOICHITA, V. I. (2015). *L'effet Sherlock Holmes*. Paris : Hazan.

STRAUSS, M. & CURTIS, L. (1981). *Infant perception of numerosity*.

STRAWSON, P. F. (1959a). *Individuals*. London : Methuen.

STRAWSON, P. F. (1959b). *Les individus*. L'ordre philosophique. Paris : Éditions du Seuil.

SWINBURNE, R. (2015). L'identité personnelle : la théorie dualiste. In *Identité et survie*. Traduit de l'anglais par Rémi Clot-Goudard, Vanessa Di Marino, Stéphane Dunand et Mathieu Mulcey. Ithaque.

TERASHIMA, N. (2001). The Definition of Hyperreality. In J. TIFFIN & N. TERASHIMA (Éds.), *Hyperreality Paradigm for the Third Millenium*. New York : Routledge.

THOMASSON, A. (1998). *Fiction and Metaphysics*. Cambridge Studies in Philosophy. Cambridge : Cambridge University Press. doi :10.1017/CBO9780511527463

THOMASSON, A. (2005). Speaking of Fictional Characters. *Dialectica*, *57*, 207–226.

THOMASSON, A. (2010). Fiction, existence et référence. mis en ligne le 09 avril 2010. Récupérée à partir de http://journals.openedition.org/methodos/2446%20;%20DOI%20:%20https://doi.org/10.4000/methodos.2446

THOMASSON, A. (2011). *Fiction et métaphysique*. Cahiers de logique et d'épistémologie. trad. française de C. Majolino et J. Ruelle de *Fiction and Metaphysics*, Cambridge University Press, Cambridge Studies in Philosophy, 2008. College Publications.

THOMSON, J. J. (1983). Parthood and Identity across Time. *Journal of Philosophy*, *80*, 201–220.

TODOROV, T. (1970). *Introduction à la littérature fantastique*. Collection Points. Paris : Éditions du Seuil.

TORZA, A. (2013). How to Lewis a Kripke-Hintikka. *Synthese*, *190*(4), 743–779.

TRAN, D. S. (2016). Le renoir du grenier de villeurbanne est un faux (27/05/2016). *Le Dauphiné*. Récupérée à partir de http://www.ledauphine.com/insolite/2016/05/26/le-renoir-du-grenier-de-villeurbanne-est-un-faux

TRIBUNAL CORRECTIONNEL DE PARIS. (p.d.). Procès intenté à M. Gustave Flaubert devant le tribunal correctionnel de paris (6e chambre) sous la présidence de M. Dubarle. Audiences des 31 janvier et 7 février 1857 : réquisitoire et jugement. Réquisitoire de M. l'Avocat Impérial Ernest Pinard. Consulté en ligne le 30/11/2016. Récupérée à partir de http://flaubert.univ-rouen.fr/oeuvres/mb_pinard.php

TROUBETZKOY, W. (1995). L'ombre perdue. In J. BESSIÈRE, A. FONYI & W. TROUBETZKOY (Éds.), *Le Double. Chamisso, Dostoïevski, Maupasant, Nabokov* (50). Collection Unichamp. Champion.

TROUBETZKOY, W. (1996). *L'ombre et la différence : le double en europe*. Collection Littératures Européennes. Paris : Presses Universitaires de France.
TULENHEIMO, T. (2009). Remarks on individuals in modal contexts. *Revue internationale de philosophie*, *4*(250), 383–394.
TULENHEIMO, T. (2017). *Objects and Modalities*. Logic, Epistemology, and the Unity of Science. Cham, Switzerland : Springer.
TURING, A. (1995). *La machine de Turing* (J.-Y. GIRARD, Éd.). Paris : Seuil.
TURNER, V. W. & INSTITUTE, I. A. (1968). *The Drums of Affliction : A Study of Religious Processes among the Ndembu of Zambia*. The international African institute. Clarendon Press.
TWARDOWSKI, K. (1979). On the Content and Object of Representations. In J. PELC (Éd.), *Semiotics in Poland 1894-1969 (1971)* (p. 7–12). Synthese Library. Studies in epistemology. Dordrecht, Holland : D. Reidel Publishing Company.
TWARDOWSKI, K. (1993). Sur la théorie du contenu et de l'objet des représentations (1894). In *Husserl-Twardowski. Sur les objets intentionnels (1893-1901)* (p. 85–200). Paris : Librairie Philosophique Vrin.
TYLOR, E. B. (1920). *Primitiv culture : researches into the development of mythology, philosophy, religion, language, art, and custom (2 volumes) (1871)*. Londres : John Murray.
TYMMS, R. (1949). *Doubles in Literary Psychology*. Cambridge : Bowes et Bowes.
TZOUVARAS, A. (1993, juin). Significant Parts and Identity of Artifacts. *Notre Dame Journal of Formal Logic*, *34*. doi :10.1305/ndjfl/1093634732
UEXKÜLL, J. v. (1965). *Mondes animaux et monde humain suivi de la théorie de la signification (1934)*. Agora. Paris : Presses Pocket.
UNGER, P. (1979a). I Dot not Exist. In G. MACDONALD (Éd.), *Perception and Identity*. London : Palgrave.
UNGER, P. (1979b). *There are no Ordinary Things*.
UNGER, P. (1980). The Problem of the Many. *Midwest Studies in Philosophy*, *5*, 411–467.
VALÉRY, P. (1957). Au sujet d'Adonis. In *Œuvres* (T. 1). Bibliothèque de la Pléiade. Paris : Gallimard.
VANDEN EECKHOUDT, J. (1970). *Défense et camouflage des animaux*. Visages de la nature. Paris : L'école des loisirs.
VAX, L. (1963). *L'art et la littérature fantastiques*. Que sais-je ? Paris : Presses Universitaires de France.
VAX, L. (1965). *La séduction de l'étrange*. Paris : Presses Universitaires de France.
VENDLER, Z. (1979). Vicarious Experience. *Revue de Métaphysique et de Morale*, *84*(2), 161–173.
VENDLER, Z. (1984). *The Matter of the Mind*. Oxford : Clarendon Press.
VERNANT, J.-P. (1989). L'individu dans la société (1987). In *L'individu, la mort, l'amour. Soi-même et l'autre en Grèce ancienne*. Folio essais. Gallimard.
VEYNE, P. (1983). *Les Grecs ont-ils cru à leurs mythes ?* Paris : Gallimard.
VOLTOLINI, A. (2006). *How Ficta Follow Fiction. A Syncretistic Account of Fictional Entities*. Philosophical Studies Series. Springer Netherlands.
WALTON, K. (1984). On the Nature of Photographic Realism. *Critical Inquiry*, *11*(2).
WALTON, K. (1990). *Mimesis as Make-Believe*. Harvard University Press.

WANSING, H. (2017). Remarks on the logic of imagination. a step towards understanding doxastic control through imagination. *Synthese*, *194*(8).
WEHMEIER, K. F. (2008, novembre). Wittgensteinian Tableaux, Identity, and Co-Denotation. *Erkenntnis*, *69*(3), 363–376. doi :10.1007/s10670-008-9118-x
WEHMEIER, K. F. (2012). How to Live Without Identity - and Why. *Australasian Journal of Philosophy*, *90*(4), 761–77. doi :10.1080/00048402.2011.627927
WEYL, H. (1931). *The Theory of Groups and Quantum Mechanics*. London : Methuen & Co.
WIGGINS, D. (1967). *Identity and Spatio-Temporal Continuity*. Oxford : Blackwell.
WIGGINS, D. (1968). On Being in the Same Place at the Same Time. *Philosophical Review*, *77*(1), 90–95.
WIGGINS, D. (1980). *Sameness and Substance*. Cambridge University Press.
WIGGINS, D. (2012). Identity, Individuation and Substance. *European Journal of Philosophy*, *20*(1), 1–25.
WILLIAMS, B. (1973). Imagination and the Self. In *Problems of the Self : Philosophical Papers, 1956-1972*. Cambridge : Cambridge University Press.
WILLIAMSON, T. (1988). First-Order Logics for Comparative Similarity. *Notre Dame Journal of Formal Logic*, *29*(4).
WIND, E. (1986). *Arte e anarchia*. Adelphi.
WITTGENSTEIN, L. (1961). *Investigations philosophiques*. Collection tel. Paris : Gallimard.
WITTGENSTEIN, L. (1965). *Le cahier bleu et le cahier brun [1958]*. Traduction française de M. Goldberg et J. Sackur. Paris : Gallimard.
WITTGENSTEIN, L. (1989). *Remarque sur la philosophie de la psychologie t.1*. T.E.R. Paris.
WITTGENSTEIN, L. (1993). *Tractatus logico-philosophicus (1921)*. Collection tel. Trad. fr. de Gilles-Gaston Granger. Paris : Gallimard.
WÖLFFLIN, H. (2016). *Principes fondamentaux d'histoire de l'art (1915)*. Agora. Paris : Presses Pocket.
WOLLHEIM, R. (1984). *The Thread of Life*. New Haven : Yale University Press.
WOLLHEIM, R. (2005). La représentation iconique. In J.-P. COMETTI, J. MORIZOT & R. POUIVET (Éds.), *Esthétique contemporaine*. Paris : Librairie philosophique Vrin.
WOODS, J. (1974). *The logic of Fiction : A Philosophical Sounding of Deviant Logic*. London : College Publications.
WOODS, J. (2002). *Paradox and paraconsistency*. Cambridge University Press.
WREEN, M. (2002). Forgery. *Canadian Journal of Philosophy*, *32*(2), 143–166.
WYNN, C. (1992). Addition and subtraction by human infants. *Nature*, *358*, 749–50. doi :10.1038/358749a0
XU, F. (1993). *The Concept of Object Identity*. New Orleans, LA.
XU, F. (1997). From Lot's Wife to a Pillar of Salt : Evidence that Physical Object is a Sortal Concept. *Mind and Langage*, *12*, 365–392.
ZALTA, E. [Edward]. (1983). *Abstract Objects*. Synthese library ; v. 160. D. Reidel Publishing Company.
ZALTA, E. [Edwards]. (2003). Referring to Fictional Characters. *Dialectica*, *57*(2), 243–254.
ZAZZO, R. (1993). *Reflets de miroir et autres doubles*. Paris : Presses Universitaires de France.
ZEH, D. H. (1970). On the Interpretation of Measurement in Quantum Theory. *Foundations of Physics*, *1*(1), 69–77.

ZEMACH, E. M. (2005). *La beauté réelle*. Traduction de Sébastien Réhault. Rennes : Presses Universitaires de Rennes.