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

NANCY UNIVERSITÉ
LPHS – Archives Henri Poincaré, CNRS UMR 7117
Ecole Doctorale "Langages, Temps, Sociétés"
Maison des Sciences de l'Homme Lorraine

Philosophie des Modalités Epistémiques

La logique assertorique revisitée

Fabien Schang

Thèse présentée en vue de l'obtention du Doctorat en Philosophie de Nancy Université
soutenue publiquement
Le 02 mars 2007

DIRECTEUR DE THÈSE
Gerhard Heinzmann

REPRODUCTION INTERDITE

JURY

Jean-Yves Béziau (Université Neuchâtel)
Gerhard Heinzmann (Nancy Université)
John T. Kearns (University Buffalo)
Bertram Kienzle (Universität Rostock)
Manuel Rebuschi (Nancy Université)
Anne-Françoise Schmid (Nancy Université)
Denis Vernant (Université Grenoble)

Remerciements

L'amitié, ça ne se dit pas.
(Lino Ventura)

Cette thèse est le résultat d'une longue et tumultueuse réflexion, conduite pendant plus de cinq années au sein des Archives Poincaré. Je ne peux que remercier vivement tous ceux qui, durant cette période, ont su m'entourer de leur bienveillance, me stimuler par leurs remarques, me renforcer par leurs compétences. Igor, Amirouche, Philippe, Manu, Pidou, Scott, Prosper, Eric … à tous, ma partie reconnaissante.

Je remercie tout spécialement notre directeur Gerhard Heinzmann, qui me soutient depuis mon année de maîtrise. Je n'oublie pas les secrétaires des Archives, dont la gentillesse et la disponibilité sont un atout évident du laboratoire, ainsi que Manuel Rebuschi, qui a su s'armer de patience pour me conduire dans la dernière ligne droite de ma thèse.

Un grand merci également aux membres du jury de thèse, parmi lesquels certains étrangers (et non-francophones, qui plus est) m'honorent de leur présence. J'ai une dette toute particulière envers John T. Kearns, avec qui j'ai eu le plaisir de correspondre à plusieurs reprises et dont les remarques ont encouragé le sens de mes recherches. Sans oublier Bertram Kienzle, qui m'accueillit en qualité de conférencier dans son université de Rostock, ainsi que Jean-Yves Béziau, dont la compagnie chaleureuse et les réflexions stimulantes ont représenté un atout décisif dans la poursuite de mon travail. Je leur suis tous extrêmement reconnaissant.

Mes derniers remerciements iront en direction de mes amis plus intimes et de ma famille : un salut complice à mes anciens camarades de la section philosophique 1995 (Céline, Hupt & Cathie, sans oublier Mike, victime de nos spéculations nocturnes), à qui je dois ce qu'ils sont ; un salut fraternel à mon *Groβbruder* Laurent et mes parents, à qui je dois ce que je suis ; un salut éternel, enfin, en direction de mes grands-parents : ma chère 'mamie' maternelle et mon grand-père paternel, dont l'esprit m'a guidé jusqu'ici et à qui je dédie ce travail.

Table des Matières

Introduction : faire du Moderne avec de l'Ancien i
1. Rendons à Aristote ce qui appartient à Aristote i
2. Ne rendons pas à Aristote ce qui ne lui appartient pas v
3. Comment strawsoniser un Aristote ? ix
4. Quantifier, c'est agir xii
5. Les 'véri-problèmes' entre parenthèses xv

Chapitre 1. Une logique d'attitudes 1
1.1 Une histoire problématique 2
1.1.1 D'Aristote à Kripke 3
1.1.2 Des axiomatiques modales aux mondes possibles 9
1.1.2.1 Les systèmes de Lewis 9
1.1.2.2 Deux types de sémantique 12
1.1.2.2.1 Sémantique algébrique 12
1.1.2.2.2 Sémantique relationnelle 18
1.2 Présentation de la logique modale épistémique 21
1.3 Les enjeux d'une sémantique relationnelle 29
1.3.1 Extensionalisme plural 29
1.3.2 Complétude et compréhension 34
1.3.3 Déflation ontologique 37
1.3.4 Enoncés et Propositions 47

Chapitre 2. Les attitudes en logique 53
2.1 Logique et pensée 54
2.1.1 Croyances et psychologie 54
2.1.2 Quelques adversaires d'une logique modale épistémique 62
2.1.2.1 Frege 64
2.1.2.2 Russell 75
2.1.2.3 Wittgenstein 88
2.1.2.4 Quine 99
2.1.3 Contre l'accusation du psychologisme 107
2.1.3.1 Une confusion sur la notion d'explication 107
2.1.3.2 Une confusion sur la notion de vérité 114
2.1.3.3 La neutralité du concept sémantique de vérité 119
2.1.4 Deux dogmes de la logique 122
2.1.4.1 Contre le dogme du rationalisme logique : Tolérance et relativisme logique 123
2.1.4.2 Logique comme langage universel, logique comme calcul 129
2.1.5 Vérité logique et validité dans une structure 138
2.1.5.1 Tolérance formelle, adéquation matérielle 139
2.1.5.2 Vérité et structure 143
2.2 La logique des concepts et leur usage 147
2.2.1 Une logique 'irréaliste' 147
2.2.2 Logique épistémique = une éthique de la logique classique ? 152
2.2.3 La théorie du modèle explicatif de Hintikka 157
2.2.4 Problèmes en logique épistémique propositionnelle 161
2.2.4.1 L'omniscience logique 161
2.2.4.2 L'introspection positive (et négative) 179

2.2.4.2.1 Les approches cognitives de la connaissance 180
2.2.4.2.2 Une approche pragmatique : le paradoxe de Moore 185
2.2.4.3 Un seul théorème 'naturel' : la clause de vérité ? 201
2.3 Logique et contexte 217
2.3.1 Le problème de l'opacité référentielle 217
2.3.1.1 Une 'solution' : la transquantification 223
2.3.1.2 Les obstacle selon Quine 228
2.3.2 Individuation et identification 239
2.3.3 Une autre solution : attitude et assertion 246

Chapitre 3. Quelle logique des attitudes ? 253
3.1 Vérité et connaissance 254
3.1.1 Réalisme et antiréalisme 255
3.1.1.1 Un mérite : le problème de Gettier 257
3.1.1.2 Une limite : le paradoxe de la connaissabilité de Fitch 261
3.2 Logique et rationalité 267
3.2.1 La charité logique, entre consistance et paraconsistance 270
3.2.2 Charité bien ordonnée … 272
3.2.3 … commence par soi-même 281
3.2.4 Des limites à l'analyse logique 296
3.3 Pour une logique modale illocutoire 301
3.3.1 Croyance justifiée et assertion 302
3.3.2 Négation et dénégation 323

Conclusion 339

Annexes 1 : Logique modale épistémique 344
Bertrand Russell (*Belief, disbelief, and doubt*) 345
Alan M. Hart (*Toward a logic of doubt*) 353

Annexe 2 : Logique illocutoire 375
John T. Kearns (*Deduction and cognition*) 376

Bibliographie 405

Glossaire 414

Symbolisme 426

Index des auteurs 429

Tout le monde a un maître.
(Georges Schang)
$(\forall x)(\exists y)\ M(y,x)$ …
… et $(\exists y)(\forall x)\ M(y,x)$?

Il est d'un homme cultivé de ne chercher la rigueur pour chaque genre de choses que dans la mesure où la nature du sujet l'admet.
(Aristote, *Ethique à Nicomaque*)

Introduction

Faire du Moderne avec de l'Ancien

Plutôt que de partir de définitions hasardeuses, le thème des modalités épistémiques peut être illustré par un problème : la quantification sur les termes vides. Ce problème hérité de la syllogistique catégorique remonte au temps de la logique traditionnelle, et son exposition permettra d'introduire les termes (en italiques) ainsi que l'enjeu du débat qui nous occuperont tout au long de ce travail. La solution proposée à ce problème permettra de mettre en lumière l'argument principal de la thèse, concernant la signification des modalités épistémiques.

1. Rendons à Aristote ce qui appartient à Aristote

Dans la partie logique de son œuvre, l'*Organon*, Aristote a présenté une liste de règles de *raisonnement* susceptibles de fonder le discours scientifique. Ces règles ont été décrites par le biais de la syllogistique (dans les *Premiers* et *Seconds Analytiques*) et de la théorie des oppositions (dans *De l'Interprétation*). L'étude de la connaissance prend son point de départ dans l'expression de *propositions* de type sujet-prédicat, que l'on définit selon leur qualité (affirmative, négative) et leur quantité (universelle, particulière). La syllogistique a pour but de montrer que la vérité de certaines propositions découle *nécessairement* de la vérité de deux prémisses (une majeure et une mineure) et, par 'nécessairement', on entend que ces propositions sont déduites uniquement en raison de la *forme* du raisonnement ; on trouve déjà dans cette théorie la caractéristique essentielle de la logique : une étude du *raisonnement valide*, c'est-à-dire des relations de conséquence toujours vraies entre un ensemble de formules (les prémisses) et une autre (la conclusion). L'exemple syllogistique le plus connu est le raisonnement en B*a*rb*a*r*a* : 'Tout homme est un être vivant ; or, tout Grec est un homme ; donc, tout Grec est un être vivant'. Ce raisonnement se compose de trois universelles affirmatives et sa conclusion est *toujours vraie* lorsque les prémisses le sont, quelle que soit l'*interprétation* donnée aux termes sujets et prédicats ; si l'on remplace dans l'ordre le grand terme 'être vivant', le petit terme 'Grec' et le moyen terme 'homme' par les trois variables G, P et M, la conclusion n'en est pas moins concluante sous la forme qui suit : Tout M est un G (M*a*G) ; or, tout P est un M (P*a*M) ; donc, tout P est un G (P*a*G). Ce syllogisme de base était considéré comme 'évident' par Aristote et n'était pas démontré : c'est un mode *parfait*, comme chacun des quatre modes de la première figure (B*a*rb*a*r*a*, C*e*l*a*r*e*nt, D*a*r*i*i, F*e*r*i*o) ; il en allait autrement pour les modes dits 'imparfaits', c'est-à-dire toutes les autres figures des modes concluants qui étaient démontrés à partir des modes parfaits.

Outre son intérêt historique, il faut surtout voir que la syllogistique contient en germe la discussion à venir sur la logique épistémique. Même si la logique n'est pas sortie entièrement achevée de l'œuvre d'Aristote, il convient de mettre à son crédit que la majeure partie du vocabulaire logique moderne se dégage de la syllogistique. Ainsi la preuve des syllogismes est *formelle* parce qu'elle ne s'occupe pas de la matière du discours, c'est-à-dire des *constantes* qui constituent les expressions du langage ordinaire. Leur démonstration repose uniquement sur l'*extension* des classes et peut être synthétisée en termes d'inclusion et d'exclusion : M*a*G signifie l'inclusion totale de la classe M dans la classe G, M*i*G son

inclusion partielle (certains individus de la classe M appartiennent également à la classe G, mais pas tous), M*o*G son exclusion partielle (certains individus de la classe M *ne* sont *pas* des G) et M*e*G son exclusion totale (aucun individu de la classe M n'appartient à la classe G). Extension et négation sont donc indispensables pour construire la syllogistique et décrire les règles de la pensée.

Au milieu du 20e siècle, le logicien et philosophe polonais Jan Łukasiewicz proposera de rationaliser la syllogistique d'Aristote afin de montrer ce qu'elle a d'intimement lié aux méthodes de démonstration des logiques modernes ; citons quelques-unes des comparaisons établies entre la logique des prédicats traditionnelle et la logique des propositions moderne. On retrouvera dans la description de la syllogistique ci-dessous les *règles syntaxiques de formation* des formules à l'intérieur d'un système formel, la *théorie de la preuve* des théorèmes à partir d'une *axiomatique* et, pour finir, la *sémantique* ou *théorie des modèles* qui correspond à l'interprétation des formules du système.

1. Règles de formation

La distinction entre une classe et les individus de cette classe permet de montrer que chaque terme de la démonstration syllogistique (petit, moyen ou grand) constitue déjà une proposition à part entière : P, M et G sont donnés comme des termes mais expriment des propositions reliées entre elles par des relations de quantification et d'affirmation ou de négation. Si l'on représente toute proposition élémentaire par un schéma fondamental de type 'S est P', où S symbolise le terme sujet et P le terme prédicat, la combinaison de la qualité et de la quantité donne quatre propositions possibles dans un syllogisme et peut être formalisée au sein de la théorie de la quantification moderne. On obtient ainsi l'universelle affirmative A, 'Tout S est P' : $\forall x(Sx \rightarrow Px)$; la particulière affirmative I, 'Quelque S est P' : $\exists x(Sx \wedge Px)$; la particulière négative O, 'Quelque S n'est pas P' : $\exists x(Sx \wedge \neg Px)$, et l'universelle négative E, 'Aucun S n'est P' : $\forall x(Sx \rightarrow \neg Px)$. Deux choses sont à relever dans cette logique des classes. D'une part, chacune de ces propositions peut être définie dans les termes d'une autre et permet d'établir des relations d'*équivalence*, comme nous le verrons plus bas avec la théorie des oppositions ; d'autre part, et contrairement aux apparences trompeuses du discours ordinaire, les termes sujets S ne désignent pas des individus mais des *classes*. 'S est P' symbolise donc de manière incomplète une relation entre une classe S et une classe de classes P, et la nature de cette relation doit être complétée selon la quantité et la qualité associée à la proposition : une inclusion totale (A) ou partielle (I) des éléments de S dans P, une exclusion partielle (O) ou totale (E) des éléments de S par rapport à P. Dans l'étude de Łukasiewicz, les classes d'Aristote deviennent des propositions et chaque prédication simple de type 'S est P' devient une proposition *complexe*, composée de deux propositions élémentaires ou *atomes propositionnels* dont la relation dépend de la qualité de la proposition. Les universelles deviennent des relations logiques de type *conditionnel* : $(\forall x)(Sx \rightarrow Px)$ pour l'affirmative universelle et $(\forall x)(Sx \rightarrow \neg Px)$ pour l'affirmative universelle ; les particulières deviennent des relations logiques de type *conjonctif* : $(\exists x)(Sx \wedge Px)$ pour la particulière affirmative et $(\exists x)(Sx \wedge \neg Px)$ pour la particulière négative. La distinction entre un élément *x* et sa classe S ou P donnera corps à la théorie des *descriptions définies* de Russell et permettra de montrer qu'un terme singulier, même s'il occupe la position *référentielle* du sujet grammatical dans une

proposition, n'a pas forcément une fonction référentielle. Quine mènera à l'extrême cette distinction entre sujet grammatical et sujet logique, quitte à faire violence aux langues naturelles en verbalisant les noms communs non référentiels (que l'on songe à sa fameuse transformation de 'Pégase' en 'pégaser', où Pégase n'est plus une valeur de variable x mais une classe à part entière).

2. Axiomatique

La relation entre les modes parfaits et imparfaits correspond à la relation déductive entre les *axiomes* et les *théorèmes* d'un système axiomatique, et les règles de démonstration indiquées par les consonnes des modes imparfaits (*c*onversion simple, conversion *p*artielle, *p*ermutation) servent d'*opérations* pour transformer *salva veritate* une vérité évidente ou immédiate (intuitive) en vérité démontrée ou médiate (formelle). Dans le cas de modes concluants imparfaits tels que B*a*ro*c*o et Bo*c*a*r*do, l'extension limitée de leurs classes oblige à démontrer leur validité de manière indirecte en *démontrant par l'absurde* que la négation de leur conclusion aboutit à contredire une de leurs prémisses de départ. On retrouvera couramment ce type de démonstration indirecte en théorie des modèles : pour montrer la validité de l'affirmation $[(\phi \wedge \beta) \rightarrow \psi]$, c'est-à-dire $[(\phi \wedge \beta) \Rightarrow \psi]$, on suppose la vérité de sa négation, soit $\neg[(\phi \wedge \beta) \rightarrow \psi]$ ou $[(\phi \wedge \beta) \wedge \neg\psi]$; puis l'on montre que la formule obtenue aboutit à une contradiction et doit être niée à son tour : $[(\phi \wedge \beta) \wedge \neg\psi] \rightarrow \perp$, donc $\neg[(\phi \wedge \beta) \wedge \neg\psi]$ ou $[(\phi \wedge \beta) \Rightarrow \psi]$. La démonstration par l'absurde signifie ainsi que le fait de nier une négation revient à réaffirmer la formule de départ : la loi de *double négation* ($\neg\neg\phi \Rightarrow \phi$) transparaît déjà dans la démonstration des modes B*a*ro*c*o et Bo*c*a*r*do.

3. Théorie des modèles

On distingue couramment entre ce qui est *parfois vrai*, *toujours vrai* et *toujours faux* dans un *modèle* pour expliquer la nature d'une proposition : elle est logiquement vraie ou *tautologique* lorsqu'elle est vraie pour toute *interprétation* de ses variables ; elle est empiriquement vraie ou *contingente* lorsqu'elle est vraie pour certaines interprétations et fausse pour d'autres ; elle est logiquement fausse, impossible, inconsistante, contradictoire ou *antilogique* lorsqu'elle est fausse pour toute interprétation de ses variables. Quant au *modèle*, il correspond au cadre général dans lequel les expressions sont interprétées : contrairement au modèle de l'arithmétique, par exemple, qui ne contient que des nombres dans son *domaine*, la vérité logique se caractérise comme la vérité dans *tout* modèle et vaut pour n'importe quel objet de discours. Si l'on parle de modèle*s* au pluriel, ici, c'est au sens où l'interprétation de nos affirmations quotidiennes peut être modifiée et passer du vrai au faux pour les variables de propositions, d'un objet à un autre pour les variables d'individus, etc. Lorsque modification de ce genre il y a, le modèle n'en est pas moins unique et *universel* ; l'introduction de relations entre des modèles aux structures et aux domaines variables viendra bien plus tard, avec l'apparition des sémantiques modales. Quoi qu'il en soit, la logique ne se préoccupe pas de ce dont elle parle : ses règles transcendent les *théories* et s'appliquent à n'importe quel objet de discours. Par voie de conséquence, chaque syllogisme peut être considéré comme une *implication* de type $(A \wedge B) \Rightarrow C$ entre trois propositions quelconques A, B et C, où la majeure A et la mineure B sont reliées par une *conjonction* et donnent

toujours la conclusion vraie C quelle que soit l'interprétation de ces variables. Cette formalisation complète a l'avantage de montrer que le mode B*arbara* des *Premiers Analytiques* n'est autre que l'expression semi-formelle de la loi de *transitivité* : $[(\phi \rightarrow \beta) \wedge (\beta \rightarrow \psi)] \Rightarrow (\phi \rightarrow \psi)]$; il suffisait pour le voir de permuter A et B, puis de remplacer dans l'ordre les variables M, P et G par les variables ϕ, β et ψ. De telles manipulations n'affectent en rien la validité du raisonnement et sont permises par la logique moderne : la conjonction est *commutative* (pour tout ϕ et β, on a : $(\phi \wedge \beta) \equiv (\beta \wedge \phi)$) ; la *règle d'inférence* de la *substitution uniforme* préserve la *valeur de vérité* d'une formule.

Pour finir le comparatif, les *opérations* syllogistiques sont des règles de transformation qui permettent de passer des modes parfaits aux modes imparfaits. On retrouve dans la conversion, la contraposition et l'obversion des règles d'inférence qui ont toujours cours en logique moderne. L'obversion, moins reconnue de nos jours que la contraposition, se retrouve toutefois dans la définition du conditionnel en termes de conjonction : 'Tout homme est mortel' signifie la même chose que 'Aucun homme est non-mortel', ce qui donne en logique des prédicats à l'équivalence $(\forall x)(Hx \rightarrow Mx) \Leftrightarrow \neg(\exists x)(Hx \wedge \neg Mx)$. Les démonstrations permise par ces opérations reposent sur la significations des constantes logiques de la négation, de la conjonction et du conditionnel ; leurs *conditions de vérité* générales de ces opérateurs permettent d'expliquer la validité des raisonnements logiques.

Si l'on a procédé jusqu'ici à une description quelque peu fastidieuse de la syllogistique catégorique d'Aristote, c'est dans le but de montrer que les notions qu'elle emploie se retrouvent tous dans la logique moderne quoique sous d'autres formulations. La démonstration des modes imparfaits est tout aussi fastidieuse, mais l'ouvrage de Łukasiewicz a permis de constater que l'inventaire empirique qu'Aristote avait dressé des règles de conséquence logique peut être obtenu par une théorie de la preuve *axiomatique*, sans aucun recours à l'intuition pour la démonstration des théorèmes. D'autres méthodes de preuve avaient été imaginées entre-temps pour faciliter la procédure, notamment la méthode visuelle des *diagrammes* : les relations d'inclusion et d'exclusion entre les propositions A, I, O et E peuvent être figurées par des relations spatiales entre les éléments des classes ; il suffit par ailleurs de marquer d'une croix la présence d'au moins un individu au sein de ces classes, et le lien est établi entre la syllogistique traditionnelle et la quantification moderne.

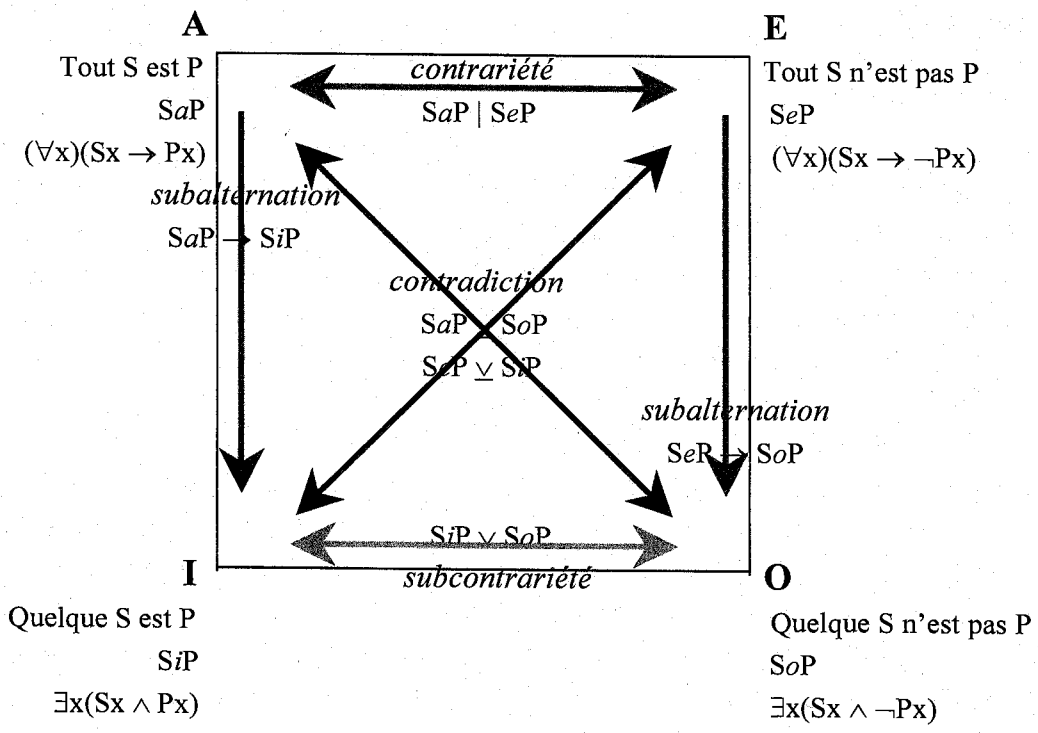
La syllogistique d'Aristote apporte ainsi une définition rigoureuse à plusieurs sortes de conséquence logique et annonce quelques concepts essentiels pour la philosophie de la logique contemporaine : la vérité, l'existence, l'implication, l'équivalence, la négation, l'affirmation, le modèle. En ce qui concerne la notion de vérité et son homologue de la fausseté, elles permettent de définir les relations fondamentales d'*opposition* entre les quatre types de proposition et de présenter l'envers du décor logique : à l'inverse de la conséquence logique, qui relie des propositions *toujours* vraies ensemble (l'une ne peut pas être vraie sans l'autre), la relation de *contradiction* oppose notamment des propositions *jamais* vraies ensemble (l'une ne peut pas être vraie avec l'autre). Ainsi peut-on dire des propositions logiquement liées qu'elles produisent des *implications*, c'est-à-dire des propositions dont la relation conditionnelle est 'toujours vraie' ou 'jamais fausse' ; inversement, on peut dire des propositions logiquement opposées qu'elles produisent des contradictions, c'est-à-dire des

propositions dont la relation conditionnelle est 'jamais vraie' ou 'toujours fausse'. Entre ces deux cas extrêmes figure l'immense catégorie des propositions *extra-logiques*, parfois vraies et parfois fausses selon l'interprétation donnée à leurs termes sujets et prédicats.

Hormis l'absence de propositions *relationnelles* (dont les prédicats s'appliquent à plus d'un argument), et bien que sa traduction en termes de logique moderne soit contestable par endroits (peut-on assimiler sans risque l'expression 'être non-P' à la forme 'ne pas être P' ?), force est donc de constater que la majorité des opérations logiques d'aujourd'hui se trouve déjà consignée dans la logique traditionnelle d'Aristote. Tout semble fonctionner à merveille dans cette machine déductive, donc ; un grain de sable a été signalé cependant par des auteurs de la logique médiévale ; or ce défaut n'est peut-être pas imputable à la théorie logique d'Aristote mais à l'interprétation qui en a été faite, et c'est cette erreur qui va conduire tout droit à l'objet de la thèse. Dans la théorie des oppositions, une difficulté apparaît en effet concernant la validité d'une relation logique particulière. Dans quelle mesure ne constitue-t-elle pas un défaut inhérent à la logique traditionnelle d'Aristote, et comment peut-on y remédier ?

2. Ne rendons pas à Aristote ce qui ne lui appartient pas

La théorie des oppositions d'Aristote apparaît aux chapitres 6 et 7 de *De l'Interprétation*, mais c'est à l'auteur latin Apulée (IIe siècle après J. C.) que l'on doit sa première exposition sous la forme d'un carré logique. Le rôle de ce carré est de représenter par une figure les quatre relations d'opposition des propositions A, I, O et E. Tout comme dans la théorie des modèles contemporaine, ces relations interpropositionnelles sont décrites en termes primitifs de valeurs de vérité. En supposant que tout ce qui *n'*est *pas* vrai est faux, et vice versa, on peut représenter les termes de la syllogistique catégorique de la manière suivante, où les universelles servent de formules primitives et peuvent être traduites en existentielles :



Deux propositions sont *contraires* lorsqu'elles ne peuvent pas être vraies ensemble mais peuvent être fausses ensemble. Deux propositions sont *contradictoires* lorsqu'elles ne peuvent être ni vraies ensemble ni fausses ensemble : la vérité de l'une implique la fausseté de l'autre, et inversement. Deux propositions sont *subcontraires* lorsqu'elles ne peuvent pas être fausses ensemble mais peuvent être vraies ensemble. Enfin, deux propositions sont *subalternes* lorsque la vérité de la première implique la vérité de la dernière.

Bien que l'opposition évoque en général une idée d'incompatibilité ou impossibilité d'être *vrai* conjointement, les relations de subalternation et de subcontrariété sont maintenues dans la théorie des oppositions et peuvent être comprises comme une certaine manière de connecter des propositions. A chacune de ces connexions est associée un connecteur de la logique propositionnelle. La relation de contrariété entre deux propositions est *satisfaite* lorsque ni l'une ni l'autre n'est vraie, ce qui correspond au connecteur logique de l'*incompatibilité* : (S*a*P | S*e*P), et (S*e*P | S*a*P) ; la relation de contradiction est satisfaite lorsque soit l'une soit l'autre des propositions est vraie, mais pas les deux à la fois, ce qui correspond au connecteur logique de la disjonction *exclusive* : (S*a*P $\veebar$ S*o*P), et (S*e*P $\veebar$ S*i*P) ; la relation de subcontrariété est satisfaite lorsqu'au moins une des deux propositions est vraie, ce qui correspond au connecteur de la disjonction *inclusive* : (S*i*P $\vee$ S*o*P), et (S*o*P $\vee$ S*i*P) ; enfin, la relation de subalternation est satisfaite seulement lorsque la vérité de la première proposition entraîne la vérité de la seconde, ce qui correspond au connecteur du *conditionnel* : (S*a*P $\rightarrow$ S*i*P), et (S*e*P $\rightarrow$ S*o*P).

Or cette dernière connexion ne semble pas s'appliquer toujours avec succès entre les affirmatives et les négatives : il existe des cas dans lesquels la vérité de la *super*alterne S*a*P (ou S*e*P) n'entraîne pas la vérité de la *sub*alterne S*i*P (ou S*o*P). Trois solutions apparaissent pour remédier à ce défaut de fabrication du carré : soit l'on impose une restriction d'usage *ad hoc* sur le modèle d'interprétation, soit l'on ne change rien au modèle et l'on redéfinit en revanche les règles d'interprétation dans ce modèle, c'est-à-dire les opérateurs logiques. Notre solution consistera à n'opter ni pour l'une ni pour l'autre et à regarder de plus près les côtés 'existentiels' du carré logique.

Dans la première approche, on décide que les relations d'opposition logique sont valides pour toute interprétation dans un modèle, mais pas n'importe quel modèle. Ce modèle doit être *non-vide* : le contre-exemple relevé par les logiciens médiévaux concerne en effet le cas où le sujet S désigne une classe vide, c'est-à-dire une classe à laquelle aucun individu *existant* n'appartient dans le modèle. Sans cette précaution d'usage, la conséquence de la classe vide serait un rejet des modes *mineurs* ou *subalternes*, ces modes imparfaits obtenus à partir des modes parfaits de la première figure. Parmi les modes subalternes, la preuve de la validité de B*a*m*a*l*i*p à partir de B*a*rb*a*r*a*, par exemple, est obtenue à partir de deux règles de transformation par *conversion*, l'une indiscutable et l'autre discutée. En vertu de la règle de conversion *simple* (*complète*), la vérité de 'Quelque S est P' implique la vérité de 'Quelque P est S' ; et en vertu de la règle de conversion *partielle* (par accident), la vérité de 'Tout S est P' implique la vérité de 'Quelque S est P'. La combinaison des deux règles permet de dériver tout d'abord le mode Br*a*m*a*nt*i*p de la quatrième figure (le mode 'galien'), par permutation et conversion partielle de la conclusion dans B*a*rb*a*r*a*, puis d'en déduire le mode subalterne

B*a*m*a*l*i*p par permutation et conversion complète de la conclusion dans Br*a*m*a*nt*i*p. Or ce mode, tout valide soit-il par la forme, donne un résultat absurde lorsqu'il affirme quelque chose d'un terme vide : la conclusion I ne peut pas être vraie, puisqu'il n'existe pas de sujet dont le prédicat puisse être affirmé. Pour régler ce problème de l'*import existentiel* de I, la seule solution pour sauvegarder les modes subalternes consiste à poser que tout syllogisme portant sur un terme sujet vide sera toujours faux 'par défaut', les prémisses devant être tout aussi fausses que la conclusion puisqu'elles affirment quelque chose d'impossible de manière universelle. Après tout, ce qui est faux particulièrement ne doit-il pas l'être d'autant plus lorsqu'il est affirmé universellement ?

Certes, mais une telle manœuvre enfreint les relations d'opposition logique décrites plus haut en termes de conditions de vérité et de fausseté. En effet, la fausseté de I implique la vérité de sa contradictoire E, ce qui n'a rien de dérangeant puisque conforme à l'idée selon laquelle aucun S ne saurait être P faute d'existence d'un S. Mais si E est vraie, A ne peut pas être vraie dans le même temps puisque deux A et E sont des contraires. Qu'à cela ne tienne : A sera fausse pour que E soit vraie lorsque I est fausse, et cela ne dérangera pas davantage puisque ce qui est faux particulièrement le sera d'autant plus universellement. Mais le problème se pose du côté de O : sa relation contradictoire avec A l'oblige à être vraie lorsque A est fausse, et la vérité de E oblige également à poser la vérité de O. On obtient de cette façon un carré des oppositions avec A et I fausses, E et O vraies. Une telle distribution des valeurs de vérité est certes conforme à la définition du conditionnel, auquel on associe la relation de subalternation : la vérité du superalterne (de l'antécédent) implique la vérité du subalterne (du conséquent), et la fausseté du superlaterne implique la fausseté du subalterne. Mais cela oblige à postuler la vérité de 'Quelque S n'est pas P', c'est-à-dire à supposer l'existence d'un S. Or il n'y a pas de S, depuis le départ de cette discussion portant sur la prédication de termes vides. Pour que l'ensemble des oppositions se conforme à notre intuition sur les déclarations d'existence sans objet, les subcontraires I et O semblent devoir être fausses toutes les deux et impliquer ainsi la vérité de leurs contradictoires A et E. De quelque côté que le problème soit pris, au moins une relation logique sera donc mise à mal tant que les règles de valuation des côtés restent ce qu'elles sont.

Dans la seconde approche du problème, il ne s'agit pas de décréter que tout modèle de discours doit être non-vide mais de modifier les règles de valuation. Pourquoi ne pas admettre que deux contradictoires puissent être vraies ensemble, ou que deux contraires puissent être vraies ensemble ? Si l'on admet le premier cas de figure, la fausseté de I n'impliquera plus toujours la vérité de sa contradictoire E et A pourra être interprétée alors comme vraie, ne serait-ce que pour garantir la fausseté escomptée du côté de O. Si l'on admet le second cas de figure, la fausseté de I implique toujours la vérité de E mais celle-ci n'implique plus la fausseté de sa contraire A ; cela permet d'accorder la vérité à A afin que sa contradictoire O puisse devenir fausse.

Ces deux modes de révision des relations d'opposition ne sont pas une pure fiction, puisqu'ils ont été imaginés et appliqués dans la théorie moderne des valuations ; ils ont même donné lieu à la création de deux *logiques philosophiques* réputées, le premier cas correspondant à la logique *paraconsistante* forte de Graham Priest et le second cas à la logique *intuitionniste*, ou paracomplète. Pour ne pas calomnier le discours du paraconsistant,

précisons que la contradictoire d'une proposition vraie ou fausse ne cesse pas tout à fait d'être fausse ou vraie mais peut se doter désormais des deux valeurs de vérité à la fois. Une proposition n'a donc plus une et une seule valeur de vérité, dans ce nouveau cadre de valuation, mais cela n'améliore pas pour autant les affaires d'Aristote ; au contraire, puisque celui-ci avait combattu dans sa *Métaphysique* (livre Γ3) toute tentative d'admettre la vérité simultanément de deux contradictoires. Les paraconsistants se réclamant de Héraclite ou Hegel, le remède apporté au problème de l'import existentiel sera pire que le mal pour un logicien classique ; et de là à prétendre que de telles révisions des propriétés de la vérité et de la fausseté constituent la panacée du logicien-philosophe moderne, il y a une marge que nous refuserons d'admettre durant la thèse. Cela vaut aussi bien pour la résolution du problème des modes subalternes que pour la résolution d'un paradoxe tel que celui du menteur : la logique paraconsistante de Priest a été imaginée en vue de résoudre un tel paradoxe, mais la procédure paraît coûteuse si elle exige une manipulation contre-intuitive de nos valeurs de vérité. Sans aller jusqu'à prétendre que les règles de valuation revêtent un caractère transcendantal (pourquoi la bivalence, plutôt qu'autre chose ?), nous prétendrons cela dit qu'il existe une autre solution capable de remédier au problème de l'import existentiel et sans modifier l'appareil logique standard hérité d'Aristote. Ou à moindre frais, ce qui revient au même.

Avant de dévoiler cette solution suggérée, signalons que jamais le 'problème du côté O' ne s'est posé en ces termes dans les textes logiques d'Aristote. Ceci pour une raison qui concerne l'exégèse, non la logique pure : si l'on reprend la traduction originale des textes d'Aristote sur le carré des oppositions, jamais celui-ci n'aurait exprimé la contradictoire de A en termes de *particulière* négative telle que 'Quelque S n'est pas P' ; si tel est le cas, le problème est résolu avant même d'être soulevé puisqu'il proviendrait d'une traduction erronée du côté O. Il faudrait lire ce côté comme la simple négation du côté A, c'est-à-dire 'Pas tous S est P' ou 'Tout S n'est pas P'. La grammaire naturelle a de quoi induire en erreur dans ce cas de figure, puisque la dernière paraphrase proposée s'apparente à l'universelle négative du côté E : dire que tout S n'est pas P serait une autre façon de dire que tout S est non-P, une fois admis que l'expression artificielle de la négation de terme 'non-P' se traduise par l'expression naturelle de la négation propositionnelle 'n'est pas P'. C'est là un problème vernaculaire de la langue française qui ne concerne pas la logique traditionnelle, toutefois : 'Not all S is P' est une expression parfaitement correcte dans la langue anglaise, contrairement à 'Pas tout S est P', et l'on peut en dire autant de l'expression allemande 'Nicht alle S sind P'. Le côté O signifiait uniquement la négation de A avant qu'il ne fut malheureusement traduit en particulière négative par son commentateur de langue latine Boèce (Ve siècle après J. C.), d'où la confusion malheureuse entre O et non-A lorsque le terme sujet est vide. Un tel détail historique n'est pas gratuit, parce que la relation de *dualité* exprimée entre les deux quantifications ci-dessus constituera plus loin une des pierres angulaires [*cornerstones*] de la logique épistémique modale.

Le problème de l'import existentiel est-il réglé pour autant, une fois la distinction maintenue entre 'Pas tout S est P' et 'Quelque S n'est pas P' ? Ce qui n'est pas consigné par écrit peut être en revanche deviné ou suggéré, compte tenu des règles rigoureuses de la logique traditionnelle. On imagine mal que son auteur ait refusé l'équivalence entre la négation de l'universelle affirmative et l'affirmation de la particulière négative. Aristote était-

il intuitionniste ? La question se pose dans la mesure où la non-équivalence des deux formes duales précédentes constitue une marque de fabrique de la logique de Heyting : $\neg(\forall x)Fx$ n'implique pas $(\exists x)\neg Fx$, parce que l'existence telle que les intuitionnistes l'entendent exige la construction d'une contre-preuve effective de Fx dans la pensée. Le problème des *futurs contingents* a de quoi étayer cette allégation puisque, comme le prétendra le même Łukasiewicz dans ses articles sur la logique trivalente, Aristote ne voulait pas défendre l'universalité de la bivalence mais celle du tiers exclu. Or puisque le principe du tiers exclu a constitué la cible principale de l'inspirateur (malgré lui) de la logique formelle intuitionniste, L. E. J. Brouwer, traiter Aristote d'intuitionniste ne tient pas ni ne saurait tenir sans commettre un certain anachronisme dans l'interprétation de ses textes. Bien qu'une certaine tension semble apparaître entre le problème de la bataille navale et le carré des oppositions, il sera donc difficile de contester la possibilité de traduire O par une particulière négative, malgré tout, et de ne pas voir dans la distinction entre non-A et O un simple détail linguistique sans incidence sur la portée des règles logiques.

Venons-en à la solution que nous proposerons et qui va inspirer l'objet de la thèse. Il ne s'agit pas de supposer la non-vacuité des modèles d'interprétation, et encore moins de modifier les règles de valuation qui caractérisent les opérateurs logiques : il s'agit plutôt de considérer les opérations de la logique comme des *actes* liés à un contexte de production. La conclusion qui en résulte servira de fil conducteur pour réorienter par la suite les analyses menées en logique épistémique.

3. Comment strawsoniser un Aristote ?

Supposons qu'il faille imputer finalement la faute à Aristote : celle de laisser sans réponse la difficulté du côté O. Rien ne dit qu'il ait songé au problème des termes sujets vides ; mais rien ne permet de dire non plus qu'il ait le droit de s'en dispenser dans le cadre d'un traité de logique formelle. A moins de trouver une raison suffisante du côté de sa métaphysique (était-il impossible ou interdit de parler seulement du non-être ?), le problème de l'import existentiel doit être réglé si l'on veut maintenir les égards dus à sa théorie des oppositions.

La chose paraît réalisable, mais à condition de revisiter (sans les réviser) les règles d'évaluation des propositions. Ou plutôt des *énoncés*, car ce ne sont pas sur les propositions que l'attention sera portée au final. Dans son article "On Referring" de 1950, les remarques conclusives du philosophe du langage Peter Strawson synthétisent l'essentiel de ce qui sera défendu dans la thèse et présenté en termes de logiques des attitudes : sa *théorie des présuppositions* revient à dire que toute affirmation ou négation au sujet d'un terme vide ne sera ni vraie ni fausse et, par conséquent, que la contradictoire de A ne sera pas susceptible d'être vraie tant que A n'est pas fausse. Cette théorie du langage et de l'usage de la référence avait pour cible la *théorie des descriptions définies* de Russell, en vertu de laquelle les termes vides ne sont pas de véritables sujets logiques et n'ont donc pas vocation à faire référence à quoi que ce soit. La distinction de Russell entre un sujet grammatical et un sujet logique peut être utile, notamment lorsqu'il s'agit d'expliquer la capacité d'une proposition à faire sens quand bien même son sujet grammatical ne désignerait rien. Plutôt que d'introduire des distinctions oiseuses entre une existence et une subsistance (ou inexistence), l'analyse logique de Russell consistait à présenter toute description définie telle que 'L'actuel roi de France'

comme une conjonction composée de trois propositions : 'Il existe un roi de France', 'Il existe un et un seul roi de France, 'Ce roi de France est chauve'. Cette analyse est commode, puisque capable de démontrer en vertu des propriétés logiques de la conjonction qu'une telle proposition sera toujours fausse. Mais elle jette le bébé avec l'eau du bain, pense Strawson, parce qu'une telle analyse considérera comme fausse toute proposition dont le sujet grammatical n'est pas un sujet logique ou ne fait pas référence à quoi que ce soit. Qui plus est, la fausseté de I impliquera de nouveau la fausseté de A et, donc, la vérité indésirable de O ; la théorie des descriptions définies ne règle pas plus le problème du côté O et son import existentiel.

Il existe certes la solution formelle de la *logique libre*, dont le rôle est d'accorder la vérité aux propositions portant sur des entités fictives en admettant un modèle d'objets fictifs ; mais cela n'apporte qu'une solution formelle et n'explique pas les conditions dans lesquelles une proposition peut recevoir un sens. Le problème semble plus profond et concerne les conditions d'attribution générale d'un sens à une expression.

Pour définir les conditions préalables à la signification d'une expression, nous ne parlerons pas en termes de conditions *de vérité* mais d'*attribution de vérité*. Pour attribuer une valeur de vérité à un énoncé, des conditions minimales doivent être garanties que les propositions portant sur des termes vides ne satisfont pas. De même que l'on ne peut pas rendre ce qui n'a pas été pris, et de même que l'on ne peut s'excuser d'une faute qui n'a pas commise, on ne peut *déclarer* la vérité d'une proposition qui ne porte sur rien. La déclaration 'Tout S est P' est donc dépourvue de sens, pour qui fait dépendre la signification des conditions de vérité ; elle sera en revanche pourvue d'un sens (de conditions de vérité) de façon générale mais insatisfaite dans ce cas particulier, lorsque le sujet grammatical S ne se rapporte à rien. Aussi triviale que soit cette remarque, son effet a une portée générale sur la logique traditionnelle et moderne : si la condition préalable à l'attribution d'une valeur de vérité est l'*assertabilité* d'une proposition, notre approche s'accordera avec Strawson sur l'idée qu'une déclaration *présuppose* l'existence de ce sur quoi elle porte. Mais qui parle de présupposition ne parle pas d'ontologie : une déclaration de vérité d'une expression de type A peut être effectuée si le locuteur pense ou croit que S n'est pas vide et fait référence à un objet existant, mais cela n'engagera personne d'autre que l'*auteur* de son expression.

En résumé, la solution au problème du côté O passe selon nous par une *logique des actes d'assertion et de supposition de vérité*, c'est-à-dire une logique dans laquelle ce qui tient lieu de 'proposition' n'est pas vraie ou fausse en soi mais rendue vraie ou rendue fausse par un acte d'affirmation et de négation. Aucun 'vérifacteur' [*truthmaker*], c'est-à-dire aucun fait correspondant à la proposition et la rendant vraie n'est exigée ici ; la métaphysique des 'truth-makers' est donc nulle et non avenue dans le cadre de ce traitement *illocutoire* des actes de discours. Le cas O ne porte donc pas un coup définitif à la théorie logique d'Aristote, sauf à supposer que toute proposition doive être vraie ou fausse et exprimer un état objectif du monde. C'est le cas chez Aristote, qui considère la proposition comme un énoncé vrai ou faux 'en soi'. Or comme le fait remarquer Strawson, il paraît curieux d'associer une valeur de vérité à une expression sans qu'il n'y ait personne pour l'interpréter au préalable. Une seule et même *phrase* peut recevoir des interprétations différentes selon son contexte d'*énonciation* ; c'est le cas des descriptions définies ou des indexicaux, puisque la phrase 'je ferme la porte'

ne recevra pas la même valeur de vérité selon le locuteur et le moment d'énonciation. Mais un énoncé peut n'être ni vrai ni faux tout en ayant un sens : chaque locuteur est capable de *comprendre* ce que veut dire la phrase ci-dessus, bien qu'il ne puisse la considérer vraie ou fausse sans avoir au moins une raison de le faire. La distinction entre une phrase, un énoncé et une proposition sera revue en détail plus bas. Elle paraît essentielle pour comprendre l'acte d'interprétation et régler un certain nombre de paradoxes en philosophie du langage ; et ce sans rien modifier dans les opérations abstraites de la logique traditionnelle, qui plus est : le problème ne vient pas des règles de valuation des opérateurs reliant des énoncés (révision des expressions logiques), il vient des règles d'interprétation de ce qui est relié par ces opérateurs (révisions des expressions extra-logiques).

La théorie des oppositions peut être sauvée, pensons-nous, si on la libère du carcan de la *logique propositionnelle* et qu'on la restreint à une analyse de ce que les anglophones qualifieraient de 'sentential logic', c'est-à-dire les phrases que les locuteurs décident d'affirmer ou de nier. Une fois que les conditions de cette décision ont été précisées selon l'usage qui est fait des phrases (pas forcément d'ordre *déclaratif*), rien n'a à être changé dans les règles de transformation de la logique traditionnelle.

Et pourtant, la solution de Strawson semble entraîner la *paracomplétude* du système logique, c'est-à-dire la possibilité que de deux contradictoires aucune des deux ne soit vraie et, par conséquent, la possibilité d'invalider le tiers exclu d'Aristote. Dans ce cas, n'est-ce pas faire vœu d'intuitionnisme que d'admettre des énoncés ni vrais ni faux, contrairement à l'objectif avoué dans cette thèse de ne rien modifier parmi les opérations logiques ? Si tel est le cas, alors toutes les règles d'inférence que Heyting avait exclues de la logique intuitionniste devront être proscrites également en logique traditionnelle, parmi lesquelles la preuve indirecte par l'absurde, la loi de double négation et le cadre général de la bivalence.

Mais ce n'est pas le cas : bien qu'un énoncé déclaratif soit nécessairement vrai ou faux, la phrase qui sert à l'exprimer n'est ni vraie ni fausse 'en soi' et ne reçoit aucune valeur tant que le locuteur n'y voit aucun *usage*. Asserter 'il n'y a pas de S qui soit P', c'est déjà supposer qu'il y a un S dans le modèle, donc c'est présupposer dans le cas d'un terme vide ce dont on suppose le contraire (que S n'est vrai de rien). Refuser d'asserter quoi que ce soit au sujet de S, c'est refuser d'affirmer ou de nier quoi que ce soit à ce sujet. C'est dans ce refus catégorique que nous situerons l'issue du problème du côté O ; ce refus exigera une distinction entre deux acceptions principales du concept de négation : la *négation* et la *dénégation*. Dire que 'quelque S n'est pas P', c'est nier d'un certain S qu'il soit P et c'est déjà trop en dire dans le cas du terme vide ; dire que 'Tout S n'est pas P', en revanche, c'est dénier que tout S soit P mais ce n'est pas forcément nier quoi que ce soit à ce sujet. La différence entre les deux formes de négation correspond à deux niveaux de discours : la négation correspond à la question ontologique de l'*être* et la dénégation à la question ontique du *dire*, ou à l'*être-dit* ; *ne pas* dire que quelque S n'est pas P ne laisse rien présupposer au sujet de S, contrairement au fait de dire que quelque S *n'est pas* P. La distinction entre dire et ne pas dire sera préférée à la distinction traditionnelle entre affirmer et nier.

Mais avant d'en arriver à cette distinction, formalisée dans la toute dernière section de la thèse, l'ensemble des problèmes posés en logique épistémique sera exposé et mis en évidence pour montrer tout l'intérêt d'une analyse logique illocutoire des verbes d'attitudes : penser,

croire, souhaiter, craindre ou percevoir sont de nombreux exemples susceptibles d'être traités dans ce cadre, mais la tâche sera amplement suffisante pour que l'on se contente des seuls verbes d'attitudes épistémiques (croire et savoir, notamment ; mais aussi être certain, douter de, avoir la preuve que, si l'on veut construire une logique exhaustive pour ces attitudes).

Là où la 'strawsonisation' de la logique traditionnelle d'Aristote aura ses limites, toutefois, c'est en ce qui concerne le rôle que Strawson accordait à l'analyse logique du langage ordinaire. Alors que celui-ci n'accordait qu'une confiance et une place limitées aux démonstrations formelles, nous tenterons d'accorder le discours de deux écoles couramment opposées en philosophie du langage : l'école du langage ordinaire d'Oxford, représentée dans les années 1950 par Austin, Ryle et Strawson ; l'école formelle, stimulée par von Wright à la même époque et confirmée aujourd'hui par des travaux de pragmatique formelle (Searle et Vanderveken).

Pour donner une orientation générale à nos réflexions, revisitons une fois encore les notions fondamentales de la théorie logique d'Aristote et examinons les ajustements considérés comme nécessaires pour une analyse logique du discours. L'accent mis sur l'acte de discours a une conséquence non pas sur le contenu des règles logiques, mais sur les conditions de leur application : quantifier, par exemple, c'est *agir* afin d'interpréter une phrase d'en faire un énoncé pourvu de sens.

4. Quantifier, c'est agir

L'idée n'est pas neuve et se retrouve aujourd'hui en sémantique des jeux et notamment en *dialogique*, où les objets d'un modèle ne sont pas donnés par avance mais sélectionnés par un locuteur. Dans la tradition héritée de Peirce puis éclipsée par l'approche aprioriste de Frege, quantifier n'est pas décréter l'existence d'un objet de discours mais sélectionner ou trouver un objet qui satisfait un jugement. Dans le même ordre d'idées, mais à l'intérieur d'un formalisme différent, la logique modale illocutoire traitera la quantification de la même manière mais dans la portée d'actes de discours symbolisés par des opérateurs modaux. Paradoxalement, le système logique auquel aboutira notre réflexion sera inspiré de la théorie du jugement de Frege et de son idéographie, la différence avec l'Allemand résidant dans l'interprétation logique des actes de jugement. Le fondement de la logique illocutoire visée reposera sur une distinction entre *fonction d'interprétation* et *fonction de vérité*, ainsi que sur un traitement de ces deux types de fonctions en termes d'opérateurs modaux. Expliquons-nous.

L'erreur attribuée à la théorie logique d'Aristote, s'il y en a une, ne concerne pas la définition des opérateurs logiques mais les circonstances dans lesquelles celles-ci sont applicables. Lorsqu'une proposition telle que 'Tout S est P' ou 'Quelque S est P' est affirmative, c'est qu'elle est affirm*ée* par quelqu'un en vue de fournir une information sur l'état du monde ; lorsqu'une proposition telle que 'Aucun S n'est P' ou 'Quelque S n'est pas P' est négative, c'est qu'elle est ni*ée* par ce même locuteur et s'adresse à un interlocuteur. On retrouve cette interaction dialogique entre un Proposant et un Opposant, entre ∃loïse et ∀bélard ou entre Moi et la Nature en sémantique des jeux, inspirée de la 'théorie' des jeux de langage du second Wittgenstein. Dans tous les cas de figure (que ce soit la sémantique dialogique constructive ou la sémantique des jeux basée sur un modèle pré-construit),

l'affirmation et la négation procèdent d'actes de discours au sein d'une conversation. La confusion est courante entre l'affirmation et la vérité, d'une part, la négation et la fausseté, d'autre part ; la raison de cette confusion est une affaire d'économie symbolique, probablement : la plupart des affirmations consistent à dire ce qui est plutôt que ce qui n'est pas, de sorte que la majeure partie des contenus de discours sont de nature affirmative lorsqu'ils cherchent à dire le vrai et de nature négative lorsqu'ils cherchent à se dédire du faux. Mais ces affirmations ne sont pas moins des interprétations de phrases qui, en tant que telles, n'ont aucun sens sans un interprétant préalable ; la différence entre la phrase et l'énoncé qui en découle est reconnue dans l'histoire de la logique, puisqu'elle oppose chez les stoïciens la *lexis* (le signe matériel) au *noêton* (le signe pensée), chez Wittgenstein le signe propositionnel à la proposition, chez Frege le signe au contenu jugeable, etc. Plus généralement, on distingue par là le signe matériel et le sens qui lui est associé. Pour mettre en évidence le rôle actif de l'agent dans la détermination du sens d'un signe linguistique, des opérateurs logiques existent en logique propositionnelle ; mais l'un de ces signes primitifs n'apparaît jamais dans la logique moderne : l'opérateur d'affirmation ⊢, dont l'absence est comparable à l'absence du signe + en arithmétique et a le tort de laisser croire qu'un énoncé aurait *un* sens indépendamment de l'interprétant. Une conséquence ressort de cette explication des actes d'assignation de vérité et de fausseté dans le discours quotidien.

Si l'affirmation consiste à assigner la vérité à une phrase, l'énoncé produit est considéré comme vrai par le locuteur ; si la négation consiste à assigner la fausseté à une phrase, l'énoncé produit est considéré comme faux par le locuteur. Dans les deux cas, l'opération se présente comme une *fonction d'interprétation* d'une phrase et produit un énoncé par le biais d'un acte d'énonciation. On utilise parfois l'expression 'contenu propositionnel' pour évoquer l''énoncé', mais l'allusion qui y est faite aux propositions serait trompeuse pour notre distinction à venir entre un énoncé et la proposition qu'il est censé exprimer. Un autre moyen de symboliser plus clairement l'acte d'interprétation consisterait à remplacer le symbolisme idéographique de Frege par un symbolisme plus 'parlant' : l'affirmation par Vp, où l'action de considérer comme vrai est associée à une phrase ou proposition grammaticale symbolisée 'p' ; la négation par Fp, où l'action de considérer comme faux est associée à cette même phrase. Il s'ensuit que l'opérateur de négation n'est pas seulement une *fonction de vérité*, mais d'abord et surtout une fonction d'*interprétation* : elle ne sert pas simplement à transformer le vrai en faux et le faux en vrai, comme c'est le cas en logique propositionnelle classique (bivalente) ; elle sert avant tout à produire un acte d'interprétation préalable sans lequel la transformation d'une valeur en une autre ne serait même pas concevable (faute de valeur de vérité assignée au départ à la phrase ou signe en tant que tel).

Les remarques ci-dessus peuvent sembler triviales, mais c'est tant mieux pour telles si elles se veulent à la fois convaincantes, c'est-à-dire capables de décrire fidèlement le processus de signification, et pertinentes, c'est-à-dire capables de résoudre un certain nombre de *paradoxes* en philosophie de la logique. Car des paradoxes il y en a, et la logique épistémique servira de champ d'étude pour les illustrer et, si possible, les régler. Si le paradoxe est regardé comme une maladie du langage ou un défaut de compréhension de nos méthodes de signification, l'intérêt d'une logique philosophique sera d'y mettre fin, mais à condition d'user de méthodes qui ne mettent pas 'la charrue avant les bœufs'. La révision des règles de valuation et la

redéfinition consécutive des opérateurs logiques sera considérée comme un stratagème de ce genre : incapable de saisir un problème conceptuel à sa racine et modifiant les règles de calcul au cas par cas, selon les besoins. Une telle critique reste à justifier et quelques passages de la thèse y pourvoiront tant bien que mal, parmi lesquelles la section consacrée au Principe de Charité : contre une révision imprudente des constantes logiques, il sera déclaré que le processus de traduction et d'interprétation des langues présuppose une procédure d'assignation de valeurs de vérité aux énoncés d'autrui ; cette procédure suppose à son tour un manuel de traduction minimal assimilé à la théorie des fonctions de vérité, donc aux règles de valuation de la logique classique (bivalente). Quoi qu'en pensera un paraconsistant, adepte des croyances inconsistantes et d'une valuation non-classique, le but de la manœuvre consistera à maintenir une moindre rigidité dans les règles logiques et à imputer toute traduction défectueuse (toute incompatibilité entre l'interprétation des phrases et le comportement de l'interprété) à un changement dans le *but illocutoire* du discours, c'est-à-dire à une expression d'un autre ordre que le discours *assertif* de référence (et vérifonctionnel : la conjonction adjoint des parties, la disjonction pose des alternatives, etc.).

A travers ce cas d'étude, c'est la thèse avancée par Grice de la *sous-détermination de la sémantique* qui sera défendue : la signification d'une expression n'est pas entièrement expliquée dans le cadre d'une analyse de ses conditions de vérité ; la stratégie de révision de ces conditions détourne du droit chemin et doit être remplacée par une analyse des intentions de discours (les buts illocutoires associés à l'expression). Un exemple : l'énoncé '*Ne* devrais-tu *pas* partir ?' n'est pas la preuve d'un comportement non-classique de la négation et d'une pluralité de significations de l'opérateur de négation : s'il équivaut à l'énoncé apparemment opposé 'Tu devrais partir', c'est parce que la première expression est un acte interrogatif où le locuteur suggère l'accomplissement par l'interlocuteur de l'affirmation 'Je vais partir'. Si les deux énoncés ci-dessus ne sont pas contradictoires, c'est parce qu'ils ne correspondent pas à deux énoncés déclaratifs opposés ; ils correspondent respectivement à une question (but illocutoire directif) et une promesse portant sur un *même* contenu propositionnel, c'est-à-dire deux actes directifs dont le but illocutoire est le même mais dont la force illocutoire n'est pas identique (suggérer le départ : 'tu devrais partir', promettre le départ : 'je vais partir'). Le logicien non-classique a selon nous le tort de ne pas tenir compte de cet aspect illocutoire du discours et de réformer la lettre des phrases avant de comprendre l'esprit de leur énoncé.

Pour résumer l'arsenal conceptuel étalé ici, la logique artificielle s'applique à des phrases pour produire des énoncés par le biais d'interprétation, et ce biais correspond à l'acte d'énonciation au sein d'un acte de discours. La diversité de ces actes et la relativité de leurs buts (leurs intentions), si elle rend justice à l'expression wittgensteinienne des 'jeux de langage', ne signifie pas pour autant le chant du *signe* pour l'application des méthodes formelles au langage ordinaire. Nous ne rentrerons pas dans les détails de l'analyse illocutoire des actes de discours, de portée bien trop large, et limiterons notre attention à quelques actes répertoriés par Searle et Vanderveken puis appliqués à l'analyse logique des notions épistémiques. Partagés entre les actes assertifs, directifs, déclaratifs et expressifs, la logique illocutoire justifie et surmonte à la fois l'argument massue lancé habituellement contre la logique épistémique : les concepts épistémiques auraient une trop grande variété de sens (d'intentions de signifier) pour être traités comme des constantes logiques ; le cadre

illocutoire permet au contraire d'étendre ce cadre rigide et de compléter les opérations formelles par une explication plus informelle des significations dans l'usage. Que l'on songe aux maximes de *quantité* et de *qualité* de Grice, en vertu desquelles le locuteur ne doit jamais livrer une quantité d'information excessive ou insuffisante par rapport à ce qu'il veut dire [*means*]. Qui prétend savoir ne dira pas qu'il 'croit' sans induire son auditoire en erreur, notamment ; cette règle a été relevée par Blanché, avec une conséquence négative pour l'analyse formelle du langage ordinaire : 'le savoir implique la croyance pour les uns et ne l'implique pas pour les autres, donc il n'y a aucune relation logique entre les deux concepts'. Le défi relevé par une logique illocutoire est de rendre justice au contraire à la diversité des usages sans tomber dans l'ambiguïté et sans multiplier les formalismes rivaux : une classification des règles de conséquence logique selon la force et le but illocutoire peut y remédier, et cela sera illustré dans le cadre des discussions sur les théorèmes de la logique épistémique. Ou plutôt des 'actes épistémiques', si nous admettons la fusion des logiques déclaratives et épistémiques au sein d'une même logique des actes de discours assertifs. Car c'est bien ce dont il s'agit en point de mire de la thèse : passer du carré d'Aristote à la logique de Hintikka, mais sans distinguer leur objet d'étude. Terminons par ce point afin de revenir aux moutons de la logique traditionnelle.

5. Les 'véri-problémes' entre parenthèses

Cette introduction a voulu montrer qu'un problème hérité de la logique des Anciens est tout aussi pertinent que les problèmes relevant de la logique des Modernes, sinon plus. Prenant à contre-pied l'approche courante de la logique épistémique, nous prétendons que la question de la *non-vérifonctionnalité* des opérateurs épistémiques n'est que le symptôme de surface d'un phénomène plus profond concernant la signification selon les usages.

La raison d'ajouter une logique modale épistémique à la logique classique est la suivante : les opérateurs de connaissance et de croyance n'obéissent pas à la théorie des fonctions de vérité, dans la mesure où la substitution uniforme de deux expressions de même *extension* ne préserve pas toujours la valeur de la vérité de la formule d'ensemble. Les exemples abondent dans la littérature logico-philosophique, depuis Frege jusqu'à Linsky en passant par Hintikka. Bien que les expressions 'Etoile du Matin' et 'Etoile du Soir' aient la même extension, l'analyse extensionnelle du langage et sa conséquence qu'est la théorie des fonctions de vérité incite à penser que la formule substituée conserve la même extension, c'est-à-dire la même valeur de sortie après que le terme sujet 'Etoile du Matin' ait été remplacé par 'Etoile du Soir'. Mais selon la connaissance du locuteur en astronomie, celui-ci croira ou ne croira pas que les deux termes singuliers visent le même objet (donc la même extension). Diverses approches de ce paradoxe épistémique seront examinées au sein de la thèse, mais elles ne feront qu'abonder dans notre idée principale selon laquelle la signification d'un énoncé dépend pour partie de l'intention du locuteur : une formule n'a pas de sens en soi, pas plus qu'une expression prononcée par deux locuteurs distincts n'est assurée de désigner le même objet ou d'exprimer la même idée (le même concept, c'est-à-dire la classe d'objets correspondants). La théorie néo-millienne des Nouveaux Théoriciens de la Référence donnera l'impression de s'opposer à notre posture néo-frégéenne ; mais l'opposition de ces néo-

statures n'aura d'intérêt que si elle s'inscrit dans une perspective d'ensemble, du carré des oppositions d'Aristote jusqu'à l'opacité référentielle de Quine.

Trois remarques concluront ce détour par Aristote, avant d'entrer dans le vif du sujet.

Notre mise en parenthèses des 'véri-problèmes' (tous les problèmes posés par la théorie de la signification comme analyse des conditions de *vérité*) a pour but principal de concevoir les conditions de vérité comme une étape seconde dans l'explication du *sens* des énoncés : l'interprétation des signes doit précéder ces conditions et, conformément aux vœux de Dummett, l'approche illocutoire devrait permettre de 'revenir suffisamment en arrière' dans l'explication du processus de signification. Mais cela n'empêchera pas pour autant un passage obligé aux conditions de vérité, dans l'analyse des actes de discours : les énoncés exprimant des questions, des ordres, des menaces, et surtout des assertions sont décomposables en termes d'interprétation d'une phrase combinée à un but précis. Une question doit susciter une affirmation ou négation de la part de l'interlocuteur, si elle veut être satisfaite (obtenir une réponse) ; une menace doit susciter l'obéissance et l'accomplissement d'un état de choses qui 'vérifiera' le contenu de la menace, si elle veut être mise à exécution ; une assertion doit s'accompagner de la vérité de l'état de chose affirmé (ou nié), si elle veut être réussie, etc. La condition de vérité est donc nécessaire sans être suffisante pour expliquer la signification des actes de discours. Dans la préface de son "Du discours à l'action", Vernant à mis en exergue l'expression d''illusion descriptive' reprochée par Austin : selon l'auteur de la tripartition locutoire/illocutoire/perlocutoire, le tort de l'analyse logique du langage est de réduire la fonction des énoncés à un rôle de représentation ou description passive du monde, et l'approche non locutoire a pour but de compléter cette approche largement insuffisante des actes intentionnels du langage ; si la condition de vérité est insuffisante pour analyser tous les actes de langage, elle n'en est donc pas moins nécessaire et n'a rien d'incompatible avec les énoncés autres que déclaratifs. Mais d'autres objections seront invoquées contre une telle analyse formelle du langage ordinaire.

Russell aurait un grief à poser contre notre distinction précédente, concernant les fonctions d'interprétation premières des phrases et les fonctions de vérité secondaires des énoncés. Contre la logique multivalente de Hugh MacColl, Russell accusa celui-ci de confondre fonctions *de vérité* et fonctions *propositionnelles* : dire qu'une proposition est *certaine* parce que vraie dans tous les cas, où dire qu'elle est *probable* parce que vraie dans au moins un cas, cela reviendrait au même que de dire d'une proposition qu'elle est vérifiée par *toutes* les constantes d'individus dans un modèle ou par l'*un* d'eux, tout au moins. En d'autres termes, MacColl fut accusé de traiter les variables individuelles comme des variables propositionnelles. Nous verrons au cours de la thèse que l'introduction de la sémantique relationnelle (des 'mondes possibles') permet de défendre le logicien modal en distinguant les niveaux de discours, loin de tout réduire à des fonctions de vérité binaires : la tradition du logique comme *langage universel* a imprégné à ce point le discours des logiciens mathématiques qu'il restreignit les procédures de calcul logique et limité le modèle d'interprétation à un seul : l'ensemble des faits du monde, symbolisé par un nombre binaire 1 pour l'affirmation et 0 pour la négation ; toute proposition est vérifiée ou falsifiée par un fait du monde, et rien d'autre ne put être exprimé dans la logique non-modale d'un Russell, un Frege ou un Wittgenstein. Dans le même sens que MacColl, partisan d'une logique comme

calcul dans plusieurs 'mondes', notre distinction entre affirmation et négation, d'un côté, assertion et dénégation de l'autre repose sur une valuation modale : affirmer, c'est supposer la vérité dans une circonstance ou état de choses, et le nier revient à contester cet état dans un 'monde' imaginé ; quant aux assertions et dénégations, elles constituent des actes de discours plus forts en ceci qu'elles évaluent des situations sur plusieurs mondes, c'est-à-dire dans plusieurs circonstances possibles. Le nombre des circonstances imaginables variera selon la complexité des énoncés interprétés, mais le procédé sémantique s'inspirera quoi qu'il en soit de la sémantique relationnelle et des modèles de Kripke.

Grâce à la typologie des actes de discours selon leur force illocutoire, enfin, la distinction ancienne entre *logique* et *rhétorique* perd son sens puisque la valeur de vérité n'est plus la composante ultime de la signification : tout acte de discours est lié à un état du monde, nous l'avons concédé ; mais la valeur de vérité n'est essentielle que lorsque l'acte de discours est d'ordre *déclaratif.* Nonobstant quelque difficulté sur le sens de 'déclaratif' (constater l'existence d'un état de choses indépendant, ou le stipuler par un acte d'autorité ?), les énoncés dont traite la logique déclarative (catégorique, ou apophantique) d'Aristote et la logique épistémique sont faits du même ordre puisqu'ils expriment un même acte de déclaration dont le but est assertif. La solution proposée au problème du côté O a reposé sur une distinction entre 'être vrai' et 'être considéré comme vrai' : c'est parce que l'on ne peut rien *affirmer* à propos d'un objet inexistant que O ne peut plus être traitée comme une particulière négative. La conséquence de cette fusion des actes déclaratifs et épistémiques est une autre analyse des formes logiques de l'énonciation ; basée sur une structure modale commune, les deux logiques précédentes sont d'ordre *assertorique* et traitent tout énoncé sur la base d'une fonction d'interprétation. Plutôt que de laisser cette fonction du côté du métalangage, elle peut être considérée comme un opérateur modal à part entière mais de portée plus limitée que les opérateurs illocutoires d'assertion et de dénégation : affirmer et nier sont des analogues de la possibilité, tandis que les assertions et dénégation correspondront à des déclarations de force analogue à la nécessité. Les nombreux cas d'étude traités tout au long de la thèse devront illustrer cette procédure et sa motivation première.

Pour terminer par le commencement, le traitement illocutoire des modalités épistémiques espère apporter une plus grande unité à l'analyse logique du langage ordinaire et, s'il en est besoin, rappeler le mérite de l'œuvre logique laissée par Aristote tout en revisitant son analyse des énoncés déclaratifs. Dans le cas problématique de l'import existentiel, le traitement de ce problème tient du même tonneau que celui proposé pour les modalités épistémiques : au côté A correspond une assertion affirmative, c'est-à-dire une affirmation forte, et au côté E correspond une assertion négative ; puisque deux suppositions affirmatives et négatives sont liées à I et O, le problème de la quantification sur des termes vides laisse à penser que le carré des oppositions est un polygone insuffisant pour régler la question de l'import existentiel. La thèse cherchera à montrer que l'acte de dénégation est plus complexe que l'assertion négative et la supposition négative ; dénier ne revient ni à affirmer fortement la fausseté d'une phrase, ni à supposer simplement sa possibilité de fausseté. Une formalisation originale de la dénégation sera proposée au final, clôturant l'examen de la logique épistémique sur un problème de logique traditionnelle. Faire du Moderne avec de l'Ancien est en tout cas le but espéré de ce travail, composé de trois parties et qui tentera de donner un aperçu des difficultés

liées au projet de formalisation d'une *logique des attitudes* : l'origine des modalités et l'évolution de leur sémantique ; les 'obstacles épistémologiques' rencontrés par la logique modale, parmi lesquels l'universalisme logique et l'approche référentielle (vériconditionnelle) du langage ; les paradoxes épistémiques et la pertinence de leur traitement en termes locutoires ou illocutoires (Moore, Fitch, opacité référentielle).

D'autres paradoxes seront évoqués sans recevoir un traitement détaillé (paradoxe du Menteur, bataille navale), mais dans le but de mettre en évidence l'intérêt d'une approche illocutoire des paradoxes épistémiques, au-delà du seul cas des modalités assertoriques. Maintenant que le projet global de la thèse a été exposé, il reste à le justifier dans les détails.

Chapitre 1

Une logique d'attitudes

Résumé : une histoire problématique de la logique modale épistémique est proposée. Comme toutes les logiques de la famille modale, son rôle est d'interpréter un type d'expressions irréductibles à des fonctions de vérité : les expressions modales. Les propriétés de cette logique modale épistémique seront ensuite exposées.

L'exercice classificatoire [des formes fondamentales de la prédication] n'est pas un exercice philosophique. C'est un prélude à la philosophie.
(Jules Vuillemin)

La langue ordinaire n'a pas de logique exacte (...) ni les règles d'Aristote ni celles de Russell ne fournissent la logique exacte de quelque expression que ce soit dans la langue ordinaire.
(Peter F. Strawson)

C'est une chose d'utiliser une sémantique comme stratagème formel pour modéliser une logique. C'en est une autre d'utiliser une sémantique comme stratagème interprétatif *pour* appliquer *une logique.*
(Greg Restall)

(...) la logique modale promet beaucoup par le biais de la clarification d'un très grand nombre de concepts et problèmes philosophiques des plus centraux, quoique moins en apparence par le biais de l'éclairage des concepts de nécessité et de possibilité logique que par l'éclairage de la nature des différentes attitudes propositionnelles, par exemple.
(Jaakko Hintikka)

1.1. Une histoire problématique

Revenir aux origines de la logique épistémique doit servir à justifier la forme modale qu'elle a prise dès le début de sa jeune histoire. Pourquoi traiter les verbes de croyance et de connaissance comme des opérateurs modaux, et de quelle nature sont les arguments susceptibles de justifier le choix d'une forme logique modale pour les attitudes propositionnelles ? On peut distinguer deux types d'explication : une explication *historique*, selon laquelle l'apparition d'un fait s'explique par le contexte historique dans lequel il se produit, les auteurs qui l'ont provoqué et leurs propres affinités théoriques, le contexte relatif dans lequel une théorie est élaborée ; une explication *théorique*, dans le cadre duquel un fait n'est pas un pur produit de l'histoire mais se justifie par des raisons plus générales, indépendantes des circonstances ou des auteurs. Le premier type d'explication tenant plus de l'habitude que de la véritable raison, on peut supposer qu'un fait est arbitraire ou n'est pas pleinement justifié lorsqu'il tient à une explication simplement historique. Le problème posé ici est d'abord celui-ci : la logique modale épistémique est-elle le produit arbitraire d'une histoire, ou le produit légitime d'un développement théorique ? La section qui suit propose une histoire problématique de la logique modale épistémique, dans le but de comprendre ce qui l'a constituée.

L'histoire de la logique épistémique est intimement liée à celle de la *logique modale*, dans laquelle la syntaxe de la logique classique est enrichie d'un opérateur primitif : la modalité *forte* ou boîte $\Box$, liée à la modalité *faible* ou diamant $\Diamond$ par la relation duale $\Box p \equiv \neg\Diamond\neg p$. C'est en raison de l'insuffisance des méthodes de démonstration de la logique classique que la logique modale a resurgi au début du 20e siècle. En effet, la logique modale étend le champ d'application de la logique formelle au sein des phrases des langues naturelles, et la *logique philosophique* va se développer sur la base de cette extension à travers ce que l'on appelle les 'sémantiques formelles des langues naturelles' (développées entre autres par Kripke, Kanger, Hintikka, puis Montague).[1]

Une modalité agit sur un énoncé comme un modificateur de valeur sémantique (de vérité), elle désigne son *mode* d'être : outre la valeur de vérité simple considérée par la logique classique, c'est-à-dire le vrai ou le faux, on peut dire de l'énoncé 'Metz est au Nord de Nancy' qu'il est vrai de manière contingente, tandis que '2 + 2 = 4' est considéré comme vrai de manière nécessaire. La modalité est souvent exprimée en termes de qualité (être vrai 'nécessairement', 'possiblement') ou de quantité (absolument vrai, plus ou moins vrai), et cette formulation explique en partie pourquoi la logique modale est assimilée parfois à une de ses expressions particulières : la logique modale *aléthique* (du grec 'aléthéia' : vérité). Bien que cette dernière soit assimilée parfois à l'analyse formelle des énoncés de nécessité (il est nécessaire que ϕ) et de possibilité (il est possible que ϕ), la notion de nécessité peut être utilisée pour définir d'autres types de modalités : la *certitude* exprime une connaissance nécessaire, l'obligation exprime une action nécessaire, etc. Le lien apparaît clairement ici

[1] Voir notamment l'essai de formalisation de la grammaire anglaise par Montague in *Formal Philosophy*, New Haven (1974), inspirée des résultats en sémantique des mondes possibles.

entre le concept général de nécessité et la logique aléthique, la logique épistémique et la logique déontique.[2]

En ce qui concerne la logique épistémique, on considère [Hintikka 1962] comme son ouvrage fondateur, mais le choix de cette date semble quelque peu arbitraire et ne rend pas justice à d'autres travaux antérieurs. Des analyses logiques du concept de croyance apparaissent en effet avant 1962 : citons les systèmes axiomatiques de [Pap 1957] et [Rescher 1960], dans lesquels apparaissent déjà les problèmes d'interprétation des axiomes doxastiques ; citons également [von Wright 1951], auquel Hintikka doit l'idée d'appliquer la sémantique modale à plusieurs concepts 'philosophiques' des langues naturelles, ainsi que des travaux moins connus telle que la logique d'assertion de [Łos 1948].

La raison pour laquelle [Hintikka 1962] est considérée comme l'ouvrage pionnier de la logique épistémique est sans doute due à sa double analyse syntaxique *et sémantique* des concepts épistémiques : son avantage est d'ordre technique, puisqu'il permet de démontrer la *complétude* du système (tout théorème est valide, et inversement)[3] et de satisfaire ainsi une des propriétés métalogiques principales d'un langage logique. Nous demanderons plus loin si ce type de propriété est une condition suffisante dans le cadre de la logique philosophique (voir section 2.2). Quoi qu'il en soit, ce critère de l'adéquation apparaît comme une possible raison pour laquelle les débuts de la logique épistémique sont associés à [Hintikka 1962].

Etant donné que cette logique épistémique appartient à la famille des logiques modales, commençons par retracer les grandes étapes de la logique modale et observons le moment où [Hintikka 1962] apparaît à l'intérieur de celle-ci.

1.1.1. D'Aristote à Kripke

L'historiographie de la logique modale peut varier selon les critères retenus pour en marquer les étapes. [Bull & Segerberg 1984] et [Meyer 2000] proposent deux ou trois vagues successives de logique modale.[4] Nous proposerons pour notre part une division de l'histoire de la logique modale en quatre périodes distinctes :

- période préhistorique (Moyen Age - fin du 19e siècle)
- période proto-historique (la fin du 19e siècle - années 1950)
- période moderne (années 1950 - années 1970)
- période contemporaine (années 1970 - aujourd'hui)

---

[2] Tandis que les notions de 'nécessité' et de 'possibilité' concernent uniquement la logique modale aléthique au sens *étroit*, nous les utiliserons ici au sens *large* : elles représentent respectivement la modalité forte $\Box$ et la modalité faible $\Diamond$, et ce quelle que soit leur interprétation (nécessité aléthique, nécessité épistémique, nécessité temporelle, nécessité déontique, nécessité physique, etc.). La nécessité aléthique désigne le mode de prédication du sujet et peut renvoyer à d'autres notions liées : la nécessité *logique*, la nécessité *ontique* et la nécessité *analytique*. Le choix d'interprétation de la nécessité aléthique est une question d'affinité philosophique, selon que le logicien admet ou non la nécessité ontique et sa doctrine conjointe de l'essentialisme.

[3] La complétude dont il s'agit ici est le métathéorème de complétude. Voir le glossaire.

[4] Pour une étude historique de la logique modale, voir également R. Goldblatt : "Mathematical modal logic : a view of its evolution", *Journal of Applied Logic* 1(2003), pp. 309-392. L'histoire de la logique modale est divisée généralement en trois temps : la période antique ou pré-formelle, la période formelle syntaxique et la période formelle sémantique. Nous ajoutons à ces trois moments un quatrième, dans le but de tenir compte des dernières avancées en logique modale depuis les années 1980. Cette quatrième période, contemporaine, correspond à l'application des systèmes logiques dans les domaines de l'informatique, de l'économie théorique et de l'intelligence artificielle.

La logique épistémique se situe dans les deux dernières.

Appelons *préhistorique* la première période de la logique modale. Celle-ci débute avec la syllogistique des modalités aléthiques, proposée par Aristote dans son *Organon* (*Premiers et Seconds Analytiques*, *De Interpretatione*), puis se prolonge avec les logiques stoïciennes et mégariques et conduit à la célèbre implication diodorienne ainsi qu'à la définition temporelle des modalités. Elle s'étend ensuite à travers le Moyen Age et les travaux de scolastiques tels que Guillaume d'Occam, Abélard ou le Pseudo-Scot, entre autres. On trouve à cette époque plusieurs travaux qui anticipent la logique modale dans sa version moderne. Une première formulation apparaît de ce qui constituera plus tard deux propriétés valables pour chaque structure modale réflexive : *A necesse esse ad esse valet consequentia* (de forme logique $\Box\phi \rightarrow \phi$), et *Ab esse ad posse valet consequentia* (de forme logique $\phi \rightarrow \Diamond\phi$), ainsi qu'une représentation des relations modales au sein d'un carré des oppositions logiques et des discussions sur le problème des *oratio obliquae* (contextes obliques), que l'on retrouvera bien plus tard sous l'appellation d'opacité référentielle au sein des logiques 'intensionnelles'.[5] Bien que les travaux de logique épistémique se situent essentiellement au sein des deux dernières périodes de la logique modale, on trouve donc déjà chez Aristote ou le Pseudo-Scot des réflexions liées à la logique des attitudes propositionnelles. Le problème de l'opacité référentielle, qui vient d'être évoqué, fut anticipé également par le Stagirite dans les *Seconds Analytiques* et les *Réfutations sophistiques* :

> Dans le cas d'un homme qui s'approche ou qui porte un masque, 'être en train de s'approcher' n'est pas non plus la même chose que 'être Coriscus', de sorte que si je suppose connaître Coriscus mais que je ne sais qui est l'homme qui s'approche, ce n'est cependant pas le cas que je connais et ne connais pas à la fois le même homme.[6]

quant à l'auteur médiéval, il présenta en termes syllogistiques des règles d'inférence que l'on associerait de nos jours à la logique modale 'non-normale'.[7] La différence avec la logique épistémique actuelle est que ces auteurs étaient alors incapables d'apporter une méthode d'interprétation formelle pour démontrer leurs affirmations. Malgré cette limitation technique

---

[5] Sur la logique médiévale et les difficultés posées par les attitudes propositionnelles, voir notamment Ivan Boh: "Epistemic and alethic iteration in later medieval logic", *Philosophy of Natur* **21**(1985), pp. 492-506; "Propositional attitudes in the logic of Walter Burley and William Ockam", *Franciscan Studies* **44**(1989), pp. 31-59; ainsi que *Epistemic Logic in the Later Middle Ages*, Routledge, New York (1984).

[6] Aristote : *Réfutations Sophistiques*, 179b2-179b6. L'exemple du masque de Coriscus était utilisé par les sophistes pour montrer le caractère contradictoire de nos connaissances, et Aristote y répondait par la distinction entre les propriétés accidentelles et essentielles d'un terme sujet : l'attribution de connaissance ci-dessus n'est pas contradictoire parce que seules les propriétés essentielles peuvent être attribuées à un sujet 'en propre' (*de re*), tandis que l'opacité référentielle n'est censée se produire qu'avec les propriétés accidentelles. Voir toutefois [Upton 1983], d'après lequel Aristote plaçait dans l'homme une sorte de faculté translucide (l'*oida*) capable de déjouer les pièges de l'opacité référentielle.

[7] Une logique modale est *non-normale* lorsqu'elle ne contient pas une des propriétés suivantes des logiques dites 'normales' : la distribution modale $\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi)$, et la nécessitation : $\vdash \phi \Rightarrow \vdash \Box\phi$. Nous reviendrons sur ces logiques non-normales, dans le cadre du traitement du paradoxe de l'omniscience logique. [Chisholm 1963] signale (p. 774) le fait que, dans la syllogistique épistémique du Pseudo-Scot, 'si les prémisses d'une inférence sont connues, la conclusion est connue' ne constituait pas une règle d'inférence.

de la logique antique et médiévale, de tels témoignages montrent clairement que certains problèmes de la logique philosophique contemporaine apparurent bien avant 1962.

Appelons *proto-historique* la seconde période de la logique modale. Après avoir joué un rôle majeur dans la syllogistique d'Aristote et les doctrines de l'être de la scolastique, la notion de modalité tomba en désuétude ou, pire, fut exclue du domaine de la 'bonne' logique sous l'impulsion des néo-positivistes du Cercle de Vienne.[8] Elle ne fut relancée qu'à la suite d'un problème posé par Lewis, relatif aux paradoxes de l'implication matérielle ou 'philonienne' : le faux implique n'importe quoi (*ex falso sequitur quodlibet*), et le vrai est impliqué par n'importe quoi (*verum sequitur ad quodlibet*).[9] Le concept d'implication stricte de C. I. Lewis, symbolisé par la flèche $-<$ et défini par la relation modale $\phi -< \psi =_{df} \Box(\phi \rightarrow \psi)$, c'est-à-dire $\phi -< \psi =_{df} \neg\Diamond(\phi \wedge \neg\psi)$, a réintroduit la notion de nécessité en 1918 afin d'éviter ces paradoxes et de tenir compte de la dépendance sémantique suggérée par le concept d'implication (si ϕ implique ψ, alors il est nécessaire que ψ soit vrai lorsque ϕ est vrai).

Trois remarques peuvent être faites sur cette seconde vague.

La logique modale est avant tout une théorie *abstraite* des modalités pures $\Box$ et $\Diamond$, sans interprétation informelle prédéfinie ; il n'y a en effet aucun lien apparent entre l'implication stricte et les attitudes épistémiques dont nous nous occuperons plus tard, hormis le fait qu'elles soient exprimables par un opérateur modal. L'implication stricte désigne un mode de vérité, la croyance un mode de pensée.[10]

La logique modale est intimement liée à la logique philosophique, dans la mesure où la première a été relancée pour satisfaire les besoins de la seconde : trouver une interprétation formelle adéquate pour des notions informelles des langues naturelles (l''implication', notamment) ; l'analyse logique de Lewis et Hintikka partent ainsi d'un même mode de questionnement philosophique que l'on peut résumer sous l'appellation de *philosophie analytique formelle*.

---

[8] Comme l'explique [Loux 1979] : "(...) en dépit de ses impressionnantes références historiques, la notion de logique modale fut regardée avec suspicion par les logiciens dans la première moitié [du 20e siècle]. On peut faire remonter leur suspicion en partie à l'orientation typiquement empiriste de la philosophie analytique du début du vingtième siècle ; et on peut les faire remonter en partie aussi à la croyance générale selon laquelle le système extensionnel formalisé dans les *Principia Mathematica* représenterait un paradigme de la logique" (p. 16). On retrouve ce type d'argumentation anti-modale à travers le naturalisme de [Quine 1993] : "Nous détectons l'emprise archaïque du mentalisme dans la préférence accordée à la cause finale sur la cause efficiente en tant que démarche explicative. (...) La modalité de la possibilité est peut-être une projection personnalisée du sentiment subjectif de capacité, rappelant la projection d'esprits dans les pierres et les arbres par les animistes. La nécessité serait alors une projection du sentiment subjectif de contrainte, ou de réduction de nos capacités". (pp. 110-1).

[9] Sur l'origine programmatique de la logique modale de l'implication stricte, voir C.I. Lewis : "Implication and the algebra of logic", *Mind* N. S. **21**(1912), pp. 522-531; *A Survey of Symbolic Logic*, University of California Press (1918). La liste des systèmes modaux S1-S5 sera établie par la suite dans C.I. Lewis & C.H. Langford : *Symbolic Logic*, Dover (1932).

[10] La logique classique est déjà une espèce de logique modale, si l'on conçoit l'opérateur d'affirmation $\llcorner$ comme un opérateur qui porte sur un énoncé ϕ mais ne modifie pas sa valeur de vérité. [Blanché 1968] (p. 84, note 1) et [von Wright 1951] considèrent ainsi la logique classique propositionnelle comme une logique modale de degré zéro (avec n=0 opérateur modal), dans laquelle les modes de vérité de l'énoncé ('est vrai', 'est faux') seraient comparables aux signes arithmétiques + ou – : de même que 2 désigne implicitement le nombre réel positif '+2', ϕ désignerait implicitement la proposition affirmative '$\llcorner\phi$'. Voir la littérature sur les 'logique de la vérité' (ou de 'l'actualité'), introduites par von Wright, ainsi que sur la 'logique de la fausseté'.

La formalisation des concepts informels est sujette à caution : d'une part, Quine reprochera notamment à l'implication stricte $-<$ de produire une confusion entre deux notions distinctes, à savoir l'implication $\Rightarrow$ du métalangage et le conditionnel $\rightarrow$ du langage-objet ; d'autre part, [Tarski 1944] a affirmé qu'un traitement formel des langages 'sémantiquement clos' (qui comportent des noms d'énoncés ainsi que des notions sémantiques telles que 'vérité') conduit inévitablement à des antinomies, d'où une limitation interne du formalisme dans son traitement des langues naturelles. Bien que le premier argument, celui de Quine parlant de la confusion entre langage-objet et métalangage comme d'un 'péché originel' commis par la logique modale (l'expression est de Quine), n'ait pas empêché le développement de cette logique par la suite,[11] la philosophie analytique formelle semble poser toutefois des difficultés concernant le champ d'application de la logique.

Puisque la thèse se concentrera uniquement sur la logique épistémique, elle s'intéressera aux troisième et quatrième périodes de la logique modale, qui commencent officiellement avec [Hintikka 1962] et coïncident avec l'introduction de la sémantique relationnelle en logique modale.

Cette troisième période, appelons-la *historique* ou *moderne*, voit l'émergence de la sémantique au sein des systèmes modaux. On distingue deux types distincts de sémantique des propositions : la sémantique *algébrique*, qui consiste dans l'application de matrices ou tables de vérité pour déterminer la signification des formules ; la sémantique des mondes possibles ou *relationnelle*, qui remplace les matrices par une évaluation similaire à la méthode des tableaux de Beth.[12] Puis les modalités seront étendues aux formules quantifiées à partir de 1946, sous la direction de Carnap et Barcan.[13]

La sémantique modale algébrique a été anticipée par les travaux de deux auteurs de la période précédente : Hugh MacColl et Charles Sanders Peirce.[14] Le premier a évoqué une logique tridimensionnelle et ajouté au vrai et au faux des valeurs sémantiques indiquant le variable et le certain ; le second imagina des mathématiques 'trichotomiques', basées sur une logique triadique dans laquelle une valeur de vérité indéterminée anticipait la future matrice trivalente de Łukasiewicz. On peut citer également le cas moins connu de la 'logique imaginaire' de Nicholas Vasil'év, qui préfigure la critique du principe de non-contradiction

---

[11] Dans [Quine 1997], "Reply to Professor Marcus", pp. 177-184, Quine précise bien que l'argument de la confusion entre usage et mention n'a pas l'effet d'un argument massue contre la logique modale : "Le Professeur Marcus a visé juste lorsqu'elle m'a décrit comme suggérant que la logique modale moderne avait été conçue dans le péché : le péché de confondre usage et mention. Mais elle a eu tort en revanche de me décrire comme soutenant que la logique modale *nécessite* une confusion de l'usage et de la mention. Mon argument était un argument historique, en rapport avec la confusion de Russell entre 'si-alors' et 'implique'." (p. 177)

[12] A la différence des matrices, les tableaux d'Evert W. Beth consistent à montrer par l'absurde qu'une formule est valide lorsque sa négation est contradictoire ; cette méthode de preuve *indirecte* se retrouvera dans [Hintikka 1962]. Pour une exposition générale de la méthode des tableaux, voir M. d'Agostino & D. M. Gabbay & R. Hähnle & J. Posegga (eds.) : *Handbook of Tableau Methods*, Kluwer Academic Publishers, Dordrecht (1999).

[13] Voir R. Carnap : "Modality and Quantification", *The Journal of Symbolic Logic* **11**(1946), pp. 33-64; Ruth Barcan : "A functional calculus of first order based on strict implication", *The Journal of Symbolic Logic* **11**(1946), pp. 1-16.

[14] Voir H. MacColl : "Symbolical reasoning", *Mind* **6**(1897), pp. 493-510, ainsi que *Symbolic Logic and its Applications*, 1906; sur la logique triadique de Peirce, dont les quelques notes datées d'avant 1910 ont été consignées dans ses *Collected Papers*, voir notamment Max Fisch & Atwell Turquette : "Peirce's triadic logic", *Transcriptions of the Charles S. Peirce Society* **11**(1966), pp. 71-85.

par Łukasiewicz et propose une logique non-classique analogue à la géométrie non-euclidienne de Lobatchevsky.[15]

Deux remarques peuvent être faites sur ces sémantiques algébriques : la première concerne la relation entre logique modale *pure* et logiques modales *interprétées* ; la seconde concerne le caractère adéquat de ces interprétations.

Dans les trois cas, on s'aperçoit que certaines interprétations particulières des modalités peuvent déboucher sur des théories plus abstraites : l'extension du nombre des valeurs de vérité a débouché en logique modale contemporaine sur une théorie des valuations, la n-chotomie.[16] Et inversement, des travaux d'algèbre pure peuvent trouver des applications ultérieures au sein de logiques non-classiques (non-bivalentes) interprétées : la logique imaginaire de Vasil'év a inspiré l'idée de paraconsistance,[17] et celle-ci prétend s'appliquer aussi bien en philosophie (dialectique de Hegel) qu'en ingénierie (systèmes experts à information incomplète) ou en théorie de la science (comment gérer des ensembles théoriques inconsistants) ; la valeur tierce ½ a donné lieu à plusieurs applications distinctes, que ce soit le problème des futurs contingents d'Aristote (Łukasiewicz), les énoncés mathématiques indécidables (Kleene) ou le paradoxe des énoncés auto-référentiels (Bochvar).[18] Cette relation à double sens entre interprétation particulière et abstraction générale est à retenir pour la suite, concernant la question des logiques 'philosophiquement pertinentes'.

Par ailleurs, l'interprétation particulière dont McColl est parti pour construire sa logique tridimensionnelle montre une certaine ambiguïté sur l'interprétation des modalités. Dans une matrice non-bivalente, les valeurs de vérité ne désignent plus seulement la réalisation (vérité) ou non (fausseté) d'un état de choses exprimé par l'énoncé mais expriment également notre connaissance de ces états ; lorsqu'un énoncé porte sur un événement futur ou indécidable, il peut être dit ni vrai ni faux si le fait ou état de choses qu'il exprime est considéré comme inconnu ou inconnaissable. Nous reviendrons plus loin sur cette confusion entre deux interprétations distinctes de la nécessité : un sens *métaphysique*, qui concerne la nature de la relation entre une propriété et son objet ; un sens *épistémique*, qui concerne la connaissance qu'a un sujet de cette relation. Faut-il voir dans cette confusion une raison de rejeter la logique philosophique et de la considérer comme le résultat d'une incompréhension profonde sur la signification des énoncés, c'est-à-dire une incompréhension sur leur valeur sémantique ? La réponse à cette question dépend de l'importance philosophique que les logiciens modaux accordent à leurs travaux, mais elle montre d'ores et déjà une certaine ambiguïté entre modalités aléthiques et modalités épistémiques.

---

[15] Voir N.A. Vasil'év : *Imaginary Logic. Selected Works* ; Nauka, Moscou (1989) ; sur le lien entre Vasil'év et Lobatchevsky, voir V.A. Bazhanov : "The imaginary geometry of N.I. Lobatchevsky and imaginary logic of N.A. Vasiliev', *Modern Logic* **4**(1994), pp. 148-156 ; sur l'héritage de Vasil'év en logique paraconsistante, voir G. Priest : "Vasil'év and Imaginary Logic", *History and Philosophy of Logic* **21**(2000), pp. 135-146.

[16] Sur la théorie de la valuation générale, voir N.C.A. da Costa : *Logiques classiques et non-classiques*, Paris, Masson (1997).

[17] La paraconsistance est une famille logique désignant l'ensemble des systèmes formels inconsistants mais non triviaux : il est possible d'admettre la vérité de ϕ et de sa négation sans admettre n'importe quoi.

[18] Voir J. Łukasiewicz : "On three-valued logic", *Ruch Filozoficzny* **5**(1920), pp. 170-1; D.A. Bochvar : "On a three-valued calculus and its application to the analysis of paradoxes of classical extend functional calculus", *Matématičeskij Sbornik* **4**, pp. 287-308 (en russe); S.C.Kleene: "On a notation for ordinal numbers", *The Journal of Symbolic Logic* **3**(1938), pp. 150-155.

La quatrième période ou période *contemporaine*, enfin, annonce la spécialisation des logiques philosophiques et leurs applications dans des domaines professionnels tels que l'économie théorique, l'intelligence artificielle, l'informatique ou les sciences cognitives (de 1980 à aujourd'hui).[19]

Cette dernière période, qui s'étend du milieu des années 1980 jusqu'à aujourd'hui, peut être résumée par deux de ses caractéristiques formelles principales : l'insistance sur les modèles formels *non-normaux*, destinés à apporter une formalisation plus 'réaliste' ou plus fine (donc plus complexe dans son expression) des raisonnements humains ; l'introduction de l'aspect *dynamique* en logique et l'apparition aux côtés de la théorie des modèles de la théorie des jeux, destinée à tenir compte de la constitution interactive (entre plusieurs agents distincts) des ensembles d'information. D'après Hintikka, le passage entre la troisième et la quatrième périodes de la logique modale correspond pour la logique épistémique à deux générations distinctes d'analyse : on passe d'une analyse *statique* du savoir à une analyse *dynamique* de savoirs communs ou d'ensembles de croyance révisables, et la notion platonicienne de connaissance est remplacée par la notion faillible et modulable d'information, dans la lignée des recherches effectuées en sciences cognitives.[20] La thèse portera toute son attention sur la troisième période, la période statique, et elle justifiera ce choix par le fait que c'est la logique philosophique en tant que telle qui est mise en cause plutôt que ses développements récents.

Au niveau du langage formel, cette période contemporaine incarne le passage à des modèles sémantiques plus complexes ainsi que l'affaiblissement de quelques propriétés générales de la relation de conséquence pour décrire les modèles d'information, parmi lesquelles la monotonie. Avec l'application des logiques modales épistémiques, la formalisation des attitudes propositionnelles donne une tournure de plus en plus *expérimentale* à la logique philosophique et fait appel à des structures formelles choisies 'au cas par cas', selon les besoins de modélisation des raisonnements effectifs. On assiste donc au passage d'un modèle de base (la logique modale normale) à des modèles à géométrie variable, en logique épistémique. A travers cette évolution historique des modalités, nous discuterons un problème posé par la logique philosophique à la philosophie de la logique : l'impact que peut avoir une réinterprétation des opérateurs sur leurs propriétés de *conséquence logique*, révisables notamment au sein d'autres branches de la logique pure (logiques substructurelles, logiques non-monotones). Nous nous demanderons plus tard si une redéfinition de ces propriétés apporte une solution proprement dite aux paradoxes de la logique épistémique (voir sections 3.1.1.2 et 3.2).

Après ce panorama historique de la logique modale, et après avoir situé la logique épistémique au sein de cette histoire, voyons à présent ce qui permet de justifier son évolution : l'axiomatique de Lewis et la sémantique relationnelle sont les deux caractéristiques principales dont [Hintikka 1962] s'est inspiré. Pourquoi avoir choisi une sémantique des 'mondes possibles', de préférence à une sémantique algébrique en termes de tables de vérité multivalentes ? N'y a-t-il qu'une différence de notation entre ces deux sortes de sémantique, ou faut-il voir une différence plus significative entre les deux ?

---

[19] Sur les applications de la logique épistémique, voir [Meyer & van der Hoek 2004] et [Hendricks 2005].
[20] Voir [Hintikka 2003].

1.1.2. Des axiomatiques modales aux mondes possibles

Comme la plupart des systèmes de logique formelle moderne, une logique modale se compose d'une partie syntaxique (règles de formation et de transformation des formules) et d'une partie sémantique (conditions de vérité des formules). Elle a été conçue tout d'abord comme une *extension* de la logique classique propositionnelle : tous les théorèmes de la logique classique sont également des théorèmes de la logique modale. Bien que la période proto-historique apporta une théorie de la preuve pour les formules modales, celles-ci demeureront sans sémantique jusqu'au milieu des années 1950.

Nous allons rappeler d'abord les propriétés principales de l'axiomatique modale, puis nous expliquerons les intuitions qui ont favorisé le choix d'une sémantique relationnelle pour la logique modale, de préférence à une sémantique algébrique.

1.1.2.1 Les systèmes de Lewis

Une propriété fondamentale d'une logique modale est la relation de dualité entre deux de ses opérateurs : $\Box\phi =_{df} \neg\Diamond\neg\phi$.[21] Quant aux axiomes choisis pour spécifier ces modalités, leur nombre varie selon le système modal choisi. A l'époque de Lewis, il s'agissait de voir si l'un de ces systèmes serait en mesure de saisir la notion d'implication. Mais l'idée est apparue peu à peu que les modalités pourraient s'appliquer tout aussi bien à d'autres concepts que celui de l'implication. Rappelons qu'une première interprétation des modalités en termes d'attitudes propositionnelles existait dès le Moyen-Age avec la syllogistique épistémique du Pseudo-Scot ; aucune axiomatique n'était concevable toutefois à cette époque, comme c'est le cas dans la description du système **S1** ci-dessous.

La logique épistémique modale apparut ainsi comme un prolongement direct de l'axiomatique modale de C. I. Lewis, caractérisée par la liste des systèmes **S** ci-dessous. Chacun des systèmes **S1**, **S2**, **S3**, **S4** et **S5** spécifie un ensemble d'axiomes de nombre croissant.[22] Les cinq systèmes de base sont les suivants :

---

[21] Nous parlons ici de la logique modale classique et non-intuitionniste, c'est-à-dire l'ensemble des systèmes modaux qui sont des extensions de la logique classique. Précisons-le parce que la relation de dualité ci-dessus n'est pas admise par les intuitionnistes : de même que l'intuitionniste refuse l'équivalence entre les quantifications $\forall$ et $\neg\exists\neg$ et admet uniquement l'implication ($\forall \Rightarrow \neg\exists\neg$) (la réciproque n'est pas valable), le logicien modal intuitionniste refuse l'équivalence entre les modalités $\Box$ et $\neg\Diamond\neg\phi$ et admet uniquement l'implication ($\Box \Rightarrow \neg\Diamond\neg$) (la réciproque n'est pas valable). Le parallèle établi entre les quantifications et les modalités peut être appliqué aussi aux valeurs de vérité, que [von Wright 1951] considérait par ailleurs comme des modalités : l'intuitionniste refuse l'équivalence entre vrai et non-faux et admet uniquement l'implication à sens unique (vrai $\Rightarrow$ non-faux) (la réciproque n'est pas valable).

[22] Il existe un grand nombre de systèmes modaux distincts, en plus des systèmes **S** de Lewis : **K**, **T**, **S4.2**, **Grz**, **Q**, etc. Le nombre des systèmes étudiés s'est accru sous l'effet de l'introduction de la sémantique modale, lorsque chaque axiomatique a été associée à une propriété relationnelle des mondes possibles. D'où la multiplication des systèmes obtenus en fonction du nombre des relations envisageables. La hiérarchie des modalités s'établit en termes de réduction du nombre des modalités : plus un système permet de démontrer l'équivalence de formules modales distinctes (de réduire les unes aux autres), plus il est fort. Ainsi, **S4** est plus fort que **T** parce qu'il permet la réduction d'une itération de *n* modalités de même nature (de $\Box^n$ à $\Box$), et **S1** est le plus faible de tous les systèmes modaux. Pour une exposition des modalités et leurs relations hiérarchiques, voir G. E. Hughes & M. J. Cresswell : *An Introduction to Modal Logic*, Methuen (1968).

S1

Axiomes :

$(\phi \wedge \psi) -\!\!< (\psi \wedge \phi)$

$(\phi \wedge \psi) -\!\!< \phi$

$\phi -\!\!< (\phi \wedge \psi)$

$((\phi \wedge \psi) \wedge \varphi) -\!\!< (\phi \wedge (\psi \wedge \varphi))$

$((\phi -\!\!< \psi) \wedge (\psi -\!\!< \varphi)) -\!\!< (\phi -\!\!< \varphi)$

$\phi -\!\!< \Diamond\phi$

Règles :

Substitution uniforme des formules bien formées (syntaxe du calcul propositionnel classique)

Substitution des équivalents stricts : de $(\phi >\!\!-\!\!< \psi)$ et φ, inférer toute formule inférée de φ en substituant ψ à ϕ.

Adjonction : de ϕ et ψ, inférer $\phi \wedge \psi$

Détachement strict : de ϕ et $\phi -\!\!< \psi$, inférer ψ.

S2

Axiomes :

S1 + $\Diamond(\phi \wedge \psi) -\!\!< \Diamond\phi$

S3

Axiomes :

S1 + $(\phi -\!\!< \psi) -\!\!< (\neg\Diamond\psi -\!\!< \neg\Diamond\phi)$

S4

Axiomes :

S1 + $\Diamond\Diamond\phi -\!\!< \Diamond\phi$

S5

Axiomes :

S1 + $\Diamond\phi -\!\!< \Box\Diamond\phi$

D'autres systèmes modaux plus riches (contenant d'autres théorèmes) seront ajoutés à la famille des systèmes **S**, et les plus connus d'entre eux se retrouveront dans la logique modale épistémique : les systèmes *normaux*. Ces systèmes correspondent à l'ensemble des théorèmes pour lesquels il existe un *modèle de Kripke* qui les valide : si un système modal est normal, alors tout théorème de ce système est également une formule valide, c'est-à-dire toujours vraie en vertu de son interprétation dans un modèle correspondant. La distinction entre les systèmes normaux et non-normaux est donc de nature sémantique : [Kripke 1963] a proposé des calculs propositionnels pour les systèmes normaux, par opposition aux systèmes de Lewis dont les axiomes n'ont pas de modèle correspondant (**S1**, **S2** et **S3**). Nous reviendrons en section 1.2 sur la correspondance entre les propriétés syntaxiques et sémantiques des axiomes modaux, ainsi que sur les relations de dépendance entre ces systèmes.

On distinguera ainsi deux groupes de systèmes modaux selon leur capacité ou non à être évalués dans un modèle sémantique.

Systèmes non-normaux :

S1, S2, S3

Systèmes normaux (la liste n'est pas exhaustive) :

K, T, D, B, S4, S5

Le plus faible des systèmes modaux normaux, le système **K** (du nom de Kripke), satisfait les deux propriétés minimales suivantes :

K
Axiome : de distribution
(second axiome de Gödel)[23] :
$\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi)$
Nécessitation :
De ϕ, inférer $\Box\phi$ (où ϕ est une thèse)

D (système **K** + **D** = **KD**)
Axiome : de consistance
$\Box\phi \rightarrow \Diamond\phi$

T (système **K** + **T** = **KT**)[24]
Axiome : de réduction (premier axiome de Gödel)
$\Box\phi \rightarrow \phi$

B (système **K** + **T** + **B** = **KTB**)
Axiome : de Brouwer
$\phi \rightarrow \Box\Diamond\phi$

S4 (système **K** + **T** + **4** = **KT4**)[25]
Axiome : d'itération positive
$\Box\phi \rightarrow \Box\Box\phi$

S5 (système **K** + **T** + **5** = **KT5**)
Axiome : d'itération positive
$\Diamond\phi \rightarrow \Box\Diamond\phi$

Si l'on applique ces différents axiomes de logique modale pure au sein de logiques modales interprétées, on obtient les quelques cas suivants :
Logique aléthique : $\Box$ = L (il est nécessaire que)
$\Diamond$ = M (il est possible que)
Logique temporelle : $\Box$ = G (ce sera toujours le cas que)/ H (il a toujours été le cas que)
$\Diamond$ = F (ce sera parfois le cas que)/ P (il a parfois été le cas que)
Logique déontique : $\Box$ = O (il est obligatoire que)
$\Diamond$ = P (il est permis que)
Logique topologique : $\Box$ = E (partout)
$\Diamond$ = S (quelque part)
Logique épistémique : $\Box = K_a$ (*a* sait que)/B_a (*a* croit que)
$\Diamond = P_a$ (*a* peut savoir que)/C_a (*a* peut croire que)[26]

---

[23] La référence à Gödel renvoie à [Gödel 1933], dans lequel l'auteur proposa une interprétation modale du calcul intuitionniste de Heyting. Nous y ferons référence par la suite.
[24] Le système **T** a été conçu par Robert Feys en 1937, qui consistait à retirer l'axiome d'itération (troisième axiome de Gödel) et à ajouter le premier. Il a été repris sous l'appellation de système **M** dans [von Wright 1951].
[25] Les systèmes **S4** et **S5** usuels sont légèrement distincts des systèmes de Lewis originels : nous n'avons plus des relations d'implication stricte $\prec$ mais d'implication matérielle $\rightarrow$. Les systèmes originels représentent une version nécessitée des systèmes modernes : pour prendre l'exemple de **S4**, $(\Box\phi \prec \Box\Box\phi) =_{df} \Box(\Box\phi \rightarrow \Box\Box\phi)$. Or puisque les formules $(\Box\phi \rightarrow \Box\Box\phi)$ et $(\Box\Box\phi \rightarrow \Box\Box\Box\phi)$ sont équivalentes dans **S4**, la différence entre les deux versions n'importe pas. On admet la formulation générale $\Box_n\phi \rightarrow \Box\Box_n\phi$, pour tout $n \geq 1$.
[26] Les premières interprétations épistémiques de la modalité étaient *impersonnelles* : dans [von Wright 1951], par exemple, l'opérateur de connaissance était traduit par '*it* is known that' et ne renvoyait donc pas à un agent individuel. La référence à un indice de personne *a* dans K_a/B_a prend toute son importance dans les systèmes de savoir commun entre des agents différents, notamment.

1.1.2.2. Deux types de sémantique

En ce qui concerne la sémantique modale, deux solutions ont été envisagées.

Soit l'on attribue aux formules d'autres valeurs sémantiques en plus des valeurs classiques du vrai (symbolisé 1) et du faux (symbolisé 0), et l'on opte alors pour une sémantique algébrique non-classique (avec $n > 2$ valeurs de vérité). Cette solution semble s'imposer, dans la mesure où il est impossible de définir les opérateurs modaux au sein d'un cadre sémantique bivalent : les quatre seuls opérateurs unaires bivalents sont l'affirmation $\llcorner$ et la négation $\neg$, la tautologie T et la contradiction $\perp$.[27] Assimiler les opérateurs modaux de la nécessité ou de la connaissance à l'un de ces foncteurs de vérité serait absurde, d'où l'idée d'une autre valuation pour caractériser ces opérateurs intensionnels.

Soit l'on maintient le cadre bivalent, et les opérateurs modaux doivent être définis alors par d'autres moyens qu'une multiplication des valeurs de vérité. C'est cette seconde voie qui va prédominer au sein de la logique épistémique modale, et nous donnerons une définition formelle de ce langage dans la section suivante. Essayons de comprendre pour l'instant ce qui a justifié cette option et relégué au second plan le choix d'une sémantique algébrique.

1.1.2.2.1. Sémantique algébrique

La sémantique classique repose sur une procédure de *bivaluation* : chaque formule reçoit une et une seule valeur parmi deux possibles, selon la loi d'application générale $\mathcal{L}_f \rightarrow \{0,1\}$: la bivaluation est une fonction d'interprétation qui associe une et une seule valeur de vérité à chaque formule d'un langage formel $\mathcal{L}_f$. Une sémantique modale algébrique étend l'ensemble des valuations à plus de deux éléments : elle est *multivalente*. Le problème posé par une sémantique de ce genre concerne la cardinalité et l'interprétation de l'ensemble des valuations : *combien* de valeurs distinctes faut-il pour définir les opérateurs modaux épistémiques ? Ont-elles une signification aussi claire que le vrai et le faux ?

La raison d'ajouter d'autres valeurs en plus du vrai et du faux tient au caractère *indécidable* des formules modales : il est impossible de déterminer la signification de ces formules dans le cadre d'une matrice classique, c'est-à-dire bivalente. Si de telles formules sont indécidables, c'est parce que la théorie de la signification ne peut s'en tenir à une analyse vérifonctionnelle et requiert d'autres critères que les seules conditions de vérité. Le tableau ci-dessus illustre l'impossibilité de déterminer la valeur sémantique des formules modales :

ϕ	$L\phi$	$M\phi$	$K\phi$	$B\phi$
1	?	1	?	?
0	0	?	0	?

La nécessité, la connaissance et la croyance de ϕ ne dépendent pas de la valeur de vérité de ϕ ; à l'inverse, la valeur de vérité de ϕ est indépendante de la possibilité de ϕ. La sémantique algébrique multivalente a pour but de surmonter cette difficulté : maintenir le critère

[27] Les quatre foncteurs de vérité de la logique standard se définissent ainsi : $v(\llcorner\phi) = v(\phi)$; $v(\neg\phi) = 0$ si $v(\phi) = 1$ et $v(\neg\phi) = 1$ si $v(\phi) = 0$; $v(T\phi) = 1$; $v(\perp\phi) = 0$.

vérifonctionnel de la signification, mais à condition de dépasser le cadre bivalent et d'opter ainsi pour des logiques non-classiques.

On peut citer trois exemples de logique épistémique algébrique, dans lesquelles on trouve une interprétation multivalente des formules d'attitudes propositionnelles : la logique d'assertion de [Łos 1948] ; les systèmes épistémiques **E1-E5** de [Lemmon 1957] ; la logique des probabilités subjectives de [Lenzen 1980a]. Or aucun de ces systèmes ne semble apporter un calcul logique *adéquat* pour les attitudes épistémiques, et ceci pour plusieurs raisons.

La logique d'assertion de Łos propose une sémantique n-valente, dont le nombre des valeurs est variable et dépend du nombre n des locuteurs : dans une assemblée de $n = 5$ personnes, la valeur sémantique de l'assertion d'une de ces personnes est définie par la formule $1/(n-1) = 1(5-1) = 1/4$ et donne ainsi une matrice logique 'pentavalente' de type $\{0,1/4,2/4,3/4,1\}$. Si l'on représente par $Lx\phi$ la formule 'x asserte ϕ', alors $v(Lx\phi) = 1$ si *tous* les membres de l'assemblée asserte que ϕ et $v(Lx\phi) = 0$ si p n'est assertée par *aucune* d'elles. Le calcul prend une tournure probabiliste, ou statistique : la valeur de l'assertion de ϕ ne dépend pas de la valeur de ϕ mais de l'opinion des personnes à son sujet. De la même façon, chacun des connecteurs propositionnels de ce systèmes se définit par des règles caractéristiques des matrices multivalentes : la négation est complémentaire ($v(\neg\phi) = 1 - v(\phi)$), la disjonction prend la valeur maximale de ses atomes ($v(\phi \vee \psi) = max(v(\phi), v(\psi))$, la conjonction prend la valeur minimale de ses atomes ($v(\phi \vee \psi) = max(v(\phi), v(\psi))$,

L'axiomatique de cette logique d'assertion contient quelques règles déjà aperçues précédemment dans les systèmes **S** de Lewis : la règle de distribution du système **K** ci-dessus, $Lx(\phi \rightarrow \psi) \rightarrow (Lx\phi \rightarrow Lx\psi)$; une version équivalente de la règle d'itération du système **S4** : $LxLx\phi \equiv Lx\phi$; une adaptation du système **D** ($\Box\phi \rightarrow \Diamond\phi$) : $Lx\phi \rightarrow \neg Lx\neg\phi$. Mais la différence essentielle avec la plupart des logiques d'attitude propositionnelle réside dans la définition de la *vérité* : en vertu de la valuation ci-dessus, une formule est vraie si et seulement si elle est assertée par *tous* les locuteurs du système considéré, soit l'axiome $\forall x(Lx\phi) \rightarrow \phi$. Une telle interprétation de la vérité diffère de la théorie de la *vérité-correspondance* habituelle, en vertu de laquelle un énoncé peut être vrai ou faux indépendamment des croyances (même justifiées) qu'un agent peut avoir à son sujet. L'ordre de définition des notions générales de connaissance et de vérité est donc inversé, ici : ce n'est pas la connaissance qui se définit en termes de vérité mais, au contraire, c'est la vérité qui se définit en termes de connaissance, à condition d'interpréter la connaissance comme une opinion unanime. La logique d'assertion de Łos apparaît pour cette raison comme la formalisation d'une sorte de théorie de la *vérité-consensus* : l'assignation de vérité à une formule, $v(\phi) = 1$ est la conséquence d'un accord unanime préalable entre les membres d'une communauté, $v(\forall x(Lx\phi)) = 1$. On retrouvera plus tard cette inversion du *definiendum* et du *definiens* pour les notions de vérité et de connaissance, dans le cas de l'antiréalisme sémantique (voir section 3.1.1.2).

Un autre exemple de logique épistémique multivalente est offert par [Lemmon 1957] : pour satisfaire les axiomes des systèmes **S** tout en évitant la règle de nécessitation (en vertu de laquelle toute thèse est nécessairement vraie), l'auteur propose une sémantique octovalente qui satisfait chacune des formules modales concernées. Mais le problème concerne l'interprétation informelle de ces huit valeurs : comment comprendre une valeur telle que 7, 4

ou 8 ? Le rôle de ces valeurs est purement algébrique : à défaut d'avoir un sens aussi clair que le vrai et le faux, elles produisent une valuation dans laquelle la valeur *désignée* (la vérité logique) est assignée à chacune des théorèmes du système. Une preuve de complétude est ainsi obtenue, mais on perd en clarté ce que l'on gagne en efficacité. A la différence de la sémantique modale relationnelle, dans laquelle les deux seules valeurs de vérité que sont le vrai et le faux ont une signification informelle parfaitement claire, une sémantique algébrique utilise les valeurs de vérité comme des symboles algébriques vides de sens. Cela ne posera aucun 'problème de conscience' pour le logicien ingénieur, dont le seul souci est de construire un système capable de calculer sur des symboles. Mais cela posera problème dans le cadre d'une logique philosophique : comment éclairer un concept tel que la croyance si l'on utilise une sémantique réduite à de l'algèbre pure ? [Quine 1975] a tenu ce genre de propos pour montrer que les logiques 'déviantes' ou non-classiques (non-bivalentes) n'ont aucun sens :

> Au début, la motivation [des recherches en logique multivalente] a été abstraitement mathématique : la poursuite de l'analogie et de la généralisation. Etudiée dans cet esprit, la logique à plusieurs valeurs n'est de la logique que par analogie ; en réalité il s'agit d'une théorie non interprétée ou d'algèbre abstraite.[28]

Cela veut-il dire qu'une valeur sémantique n'a aucune interprétation envisageable dès qu'elle ne représente ni le vrai ni le faux ? L'interprétation de la valeur tierce de Łukasiewicz, la valeur indéterminée ½, montre au contraire qu'une valeur sémantique peut recevoir une interprétation claire dans un cadre non-bivalent. Le Polonais précisa toutefois que certaines généralisations algébriques produisent des valeurs qui n'ont aucun sens, l'intérêt philosophique des matrices multivalentes se limitant ainsi à deux seuls cas de figure :

> Parmi tous ces systèmes plurivalents, il n'y en a que deux qui puissent prétendre à une *portée philosophique* : celui qui a trois valeurs, et celui qui en a une infinité. Car si l'on interprète les valeurs autres que 0 et 1 comme 'le possible', on ne peut raisonnablement distinguer que deux cas : ou bien l'on admet que le possible ne comporte pas de degrés, et alors on obtient le système trivalent ; ou bien l'on suppose le contraire, et alors il est naturel d'admettre, comme en calcul des probabilités, qu'il y a une infinité de degrés du possible, ce qui conduit au système à une infinité de valeurs.[29]

La logique de probabilité subjective fournit une matrice infinitaire de ce genre. Dans une interprétation épistémique des probabilités, la distinction classique entre savoir et croyance n'apparaît plus car la connaissance est réduite à la notion de conviction forte : la valeur 1 désigne un savoir subjectif synonyme de la certitude, tandis que la croyance correspond à une conviction dont la valeur se situe entre les valeurs ½ et 1. Les valeurs graduelles du doute, de

---

[28] [Quine 1975], p. 124. Dans un langage formel, une matrice logique peut être construite sans aucune motivation philosophique à l'appui, bien sûr : la théorie des matrices multivalentes de Post (1920) en est l'exemple. Mais rappelons que cet essai porte sur la logique philosophique et les applications possibles d'une théorie formelle. C'est en ce sens que la matrice octovalente de Lemmon nous pose problème.

[29] J. Łukasiewicz : "Philosophische Bemerkungen zu mehrwertigen Systemen des Aussagenkalküls", *Comptes rendus des Séances de la Société des Sciences et Lettres de Varsovie* **23** (Classe III) (1930), pp. 52-77, en part. p. 72 (nous soulignons). On pourrait peut-être étendre le jugement de Łukasiewicz au cas de la quadrivalence, dont nous aurons un aperçu dans un paragraphe suivant. Notons que l'auteur proposera lui-même un système quadrivalent, afin d'éliminer les conséquences involontaires de son système trivalent de 1920 ($V(\phi \wedge \neg\phi) \neq 0$ lorsque $V(\phi) = ½$).

la certitude, de la propension à croire, etc., trouvent leur expression formelle dans cette matrice *quantitative*, alors qu'elles seront inexprimables dans une sémantique relationnelle ou bivalente qui, pour sa part, est de nature *qualitative*.

Comment choisir la valeur adéquate pour une attitude épistémique ? On pourrait rétorquer que ce n'est pas le problème du logicien, si l'on considère que la logique est une science *conditionnelle* : elle ne s'intéresse pas à la valeur particulière des formules ϕ et ψ, mais à la valeur de la relation de conséquence $\phi \Rightarrow \psi$ *entre* ces formules. De ce fait, le logicien n'a pas à justifier la valeur assignée notamment aux hypothèses d'une inférence : il se contente de définir les règles permettant de *préserver la vérité* attribuée arbitrairement aux prémisses, et la justification de la valeur réelle d'une prémisse concerne uniquement les autres sciences. Cependant, la difficulté posée par une matrice multivalente ne se résout pas par une simple distinction entre vérité matérielle et vérité formelle : il s'agit de donner ici une interprétation plausible à l'assignation d'une valeur probabiliste située entre 0 et 1, fût-elle assignée de façon arbitraire. Supposer par exemple qu'une formule est vraie pour les besoins d'une démonstration (lui assigner la valeur 1, autrement dit) consiste à penser que ce qu'elle décrit 'est le cas' ; en revanche, comment comprendre le sens d'une assignation de valeur 0,8 à la formule exprimant un jugement de croyance, par exemple ?

Si l'on prend pour modèle les probabilités de tirage d'une boule dans un jeu de loterie, une interprétation *quantitative* de la croyance en termes de fréquence n'est pas recevable : on ne saurait évaluer statistiquement la croyance de ϕ comme un rapport entre le nombre de personnes qui considèrent ϕ et ceux qui croient que ϕ est vrai, dans la mesure où elle exprime un état mental personnel et dont l'intensité ne se mesure pas toujours en fonction de l'opinion publique. On pourrait concevoir sinon la valeur d'un énoncé comme le rapport entre le nombre de raisons de croire que ϕ et le nombre de raisons de croire que ϕ n'est pas le cas. On pourrait imaginer par exemple que la croyance de ϕ par un agent *a* pourrait être de valeur 0,8 si *a* possède quatre raisons de croire que ϕ est le cas et une seule raison de croire que ϕ n'est pas le cas. Si c'est de cette façon que la valuation probabiliste doit être entendue, alors sa définition semble assez arbitraire ou trop scientifique et éloignée de notre conception courante des attitudes propositionnelles : peut-on dresser une estimation précise du nombre d'arguments dont on dispose en faveur d'une croyance ? Tous les arguments se valent-ils pour justifier une croyance, ou ne faut-il pas plutôt distinguer entre différents coefficients selon la nature de l''évidence' et le degré de confiance que l'on peut lui accorder ?

Une telle interprétation *quantitative* des modalités épistémiques trouve malgré tout ses partisans. Pour justifier la distinction entre les notions de contradiction et d'inconsistance, [Williams 1981] fait appel ainsi au calcul des probabilités :

> La même preuve qui justifie une croyance que s_1 et justifie une croyance que s_2 à la fois peut justifier une croyance que $\neg(s_1 \wedge s_2)$. Supposez que la probabilité que s_1 soit la même que la probabilité que s_2 et que celle-ci soit de 2/3, i.e. plus probable que sa négation. Cela justifie à la fois une croyance que s_1 et une croyance que s_2. Mais elle est une preuve également que la probabilité que s_1 *et* s_2 est de $2/3 \times 2/3$, i.e. 4/9, i.e. moins probable que sa négation.[30]

[30] [Williams 1981], p. 602. Si l'auteur dit vrai, alors une logique de la croyance doit être *non-normale* et ne pas satisfaire la règle de distribution du système minimal **K** : un agent peut croire ϕ et ψ séparément sans croire la

Cet argument reste dépendant toutefois d'une condition préalable : que la théorie des probabilités (objective ou subjective) soit suffisante pour établir une théorie des croyances rationnelles, ce dont on est en droit en douter. En effet, un calcul probabiliste implique que l'agent n'acceptera pas un ensemble de croyances $\{\phi_1,\phi_2,\ldots,\phi_n\}$ qu'il accepte individuellement ($v(B\phi_1)$, $v(B\phi_2)$, ... et $v(B\phi_n) > ½$) dès que la valeur sémantique de leur conjonction entraîne le rejet de cet ensemble : $v(B\phi_1 \wedge B\phi_2 \wedge \ldots \wedge B\phi_n) < ½$; or la pratique semble infirmer cette interprétation probabiliste : un produit de valeurs inférieures à 1 (exprimant des attitudes épistémiques plus faibles que la conviction ou la certitude) est strictement décroissant, mais il est peu plausible que l'agent décidera de rejeter un ensemble de croyances dès que la valeur de celui-ci est inférieure à ½. Dans la mesure où l'attitude de croyance sert à conduire l'action du locuteur, celui-ci se comporte *comme si* le contenu de sa croyance est vrai, c'est-à-dire comme si sa valeur sémantique est égale à 1, et l'exemple ci-dessus de Williamson emploie un calcul de probabilités objectives qui ne semble pas approprié pour caractériser un ensemble de croyances.

Nous n'exclurons pas pour autant la possibilité de sémantiques algébriques en vue de définir les concepts épistémiques : la distinction platonicienne (qualitative) entre connaissance et croyance vraie est elle-même discutable et une approche quantifiable des croyances peut être appuyée par les résultats obtenus en théorie de la probabilité, même si ceux-ci semblent mieux s'appliquer aux croyances de groupes qu'aux croyances individuelles.

Par ailleurs, la critique selon laquelle une interprétation algébrique des modalités donne une sémantique *indécidable* n'est valable qu'à moitié : les formules non-bivalentes (autres que 0 et 1) ne peuvent certes pas être complètement déterminées en fonction de la valeur des formules bivalentes, mais rien n'indique *a priori* pas que la valuation bivalente soit la seule possible. Une théorie des fonctions de vérité non-bivalente serait-elle en mesure de garantir la décidabilité pour des sémantiques algébriques modales ? Pour procéder à un calcul à plus de deux valeurs de vérité, l'idée fut de construire une sémantique quadrivalente destinée à rendre compte de ces modes de vérité aléthiques : bien que la valeur de $\Box\phi$ soit indéterminable sur la seule base de ϕ, on peut construire en revanche une sémantique modale dans le sens inverse à celui de la vérifonctionnalité, c'est-à-dire sur la base des formules modales $\Box\phi$: soit 1^+ la valeur assignée à la vérité nécessaire de ϕ et 1 la valeur assignée à la simple vérité de ϕ, soit 0 la valeur assignée à la simple fausseté de ϕ et 0^- celle assignée à l'impossibilité (fausseté nécessaire) de ϕ.[31] Cette sémantique quadrivalente est adaptée aux modalités *aléthiques* et dispose d'une interprétation parfaitement claire de ses valeurs.

Or deux difficultés limitent l'intérêt de ce procédé, malgré tout : même s'il convient pour interpréter les vérités nécessaires et impossibles, on ne voit pas comment adapter ce genre de valuation à d'autres modalités, parmi lesquelles les modalités épistémiques. De plus, une

---

conjonction ($\phi \wedge \psi$). Cette position est liée au problème ultérieur de l'omniscience logique (section 2.2.4.1) et paraît séduisante, puisqu'elle permet apparemment d'éviter cet effet indésirable du système **K**.

[31] Pour des exemples de sémantique quadrivalente de ce genre, voir entre autres J. Łukasiewicz : "A system of modal logic", *Journal of Computing Systems* **1**(1953), pp. 111-149 ; R. Purtill : "Four-valued tables and modal logic", *Notre Dame Journal of Formal Logic* **11**(4) (1970), pp. 505-511 ; J.T. Kearns : "Modal semantics without possible worlds", *Journal of Symbolic Logic* **46**(1) (1981); J.-Y. Béziau : "Paraconsistent logic from a modal viewpoint", *Journal of Applied Logic* **3**(2005), pp. 7-14.

limitation interne au processus de valuation apparaît lorsque l'on veut calculer les valeurs des axiomes de Lewis : d'après un résultat de Dugundji, aucun des systèmes **S1-S5** ne peut recevoir de matrice vérifonctionnelle *finie*. Cette preuve de limitation constitue sans doute un des arguments théoriques les plus à même de justifier le déclin de la sémantique algébrique.[32]

En somme, le problème posé par la sémantique algébrique est aussi bien philosophique que technique, si l'on considère un problème d'interprétation des opérations logiques comme un problème philosophique. Mais cela ne signifie pas que les sémantiques algébrique et relationnelle n'ont aucun rapport entre elles ou portent sur des objets d'étude différents : une traduction entre les deux approches est possible. Comme l'a montré [Gödel 1933], la logique intuitionniste de Heyting peut être traduite dans le système modal **S4** de Lewis, avec $\Box$ interprété par 'est prouvable' (symbolisé B : 'beweisbar'). La valeur ½ de l'indétermination de Kleene trouve donc sa contrepartie relationnelle ; les deux sémantiques peuvent contribuer, chacune à leur manière, à invalider le principe du tiers exclu,[33] à l'origine des motivations intuitionnistes de Brouwer puis de Heyting.

Il reste à savoir si la préférence marquée pour la sémantique relationnelle est uniquement due à une raison d'ordre technique : la préférence pour la sémantique des tableaux de Beth, sur laquelle s'appuiera largement la théorie des mondes possibles de Hintikka ou Kripke ; ou si cette préférence est due à l'orientation philosophique des logiques modales de l'époque, dans laquelle une explication des opérations sémantiques était souhaitée. On peut résumer l'évolution historique de la sémantique modale en ces termes :

> L'utilisation de sémantiques multivalentes en vue du développement de la logique modale a été complètement laissée de côté. On peut l'expliquer pour deux raisons : d'une part les résultats négatifs prouvés par Dugundji, montrant que S5 et d'autres logiques modales standards ne

---

[32] Voir J. Dungundji : "Note on a property of matrices for Lewis and Langford's calculi of propositions", *The Journal of Symbolic Logic* **5**(1940), pp. 150-151. Les matrices de **S1-S5** seront soit fonctionnellement incomplètes (certaines formules ne reçoivent aucune valeur de vérité) et finies, soit fonctionnellement complètes (chaque formule reçoit une valeur de vérité) et indéfinies. A. N. Prior a utilisé des matrices indéfinies pour sa logique temporelle (*Time and Modality*, 1957) en associant à chaque instant du temps une valeur 1 ou 0 dans une suite indéfinie. Dans le sens inverse, la sémantique relationnelle propose une méthode de décision de type apagogique (preuve par l'absurde).

[33] La critique des intuitionnistes porte-t-elle réellement sur la loi classique du tiers exclu, de forme $\vDash (\phi \vee \neg\phi)$? Si l'on applique la définition de la vérité logique de Quine, une tautologie demeure vraie pour toute substitution uniforme de ses termes non-logiques. Or la substitution $[B\phi/\phi]$ devrait donner un tiers exclu modal de forme $B\phi \vee \neg B\phi$, et non la symbolisation qu'en proposent les intuitionnistes : $B\phi \vee B\neg\phi$ (toute formule ou sa négation est prouvable) ; [Gödel 1933] propose sinon de traduire la négation $\sim\phi$ du système de Heyting par $B\neg B\phi$, ce qui donnerait l'autre schéma : $B\phi \vee B\neg B\phi$. Dans les deux cas de figure, la forme logique initiale du tiers exclu est altérée. La traduction de l'intuitionnisme en termes d'opérateurs modaux explicites pose ainsi un problème en ce qui concerne la définition des *vérités logiques* : soit l'on rejette le critère syntaxique de substitution uniforme, soit l'on affirme que logiciens classiques et intuitionnistes ne parlent pas de la même 'chose' lorsqu'ils évoquent le tiers exclu. La même objection peut être adressée aux logiciens paraconsistants, lorsqu'ils prétendent rejeter le principe de non-contradiction : $\vDash \neg(\phi \wedge \neg\phi)$; dans l'article de Jean-Yves Béziau : "New light on the square of oppositions and its nameless corner", *Logical Investigations* **105**(2003), pp. 218-232, une traduction modale de 'leur' non-contradiction semble donner $\neg(\Diamond\phi \wedge \Diamond\neg\phi)$, et non $\neg(\Diamond\phi \wedge \neg\Diamond\phi)$. Au final, le *sens* d'une formule peut-il être déterminé par des critères structurels (préservation de sa forme logique) ? Sur cette question des critères d'identité des formules logiques, voir notamment Walter Sinnott-Armstrong & Amit Malhotra : "How to avoid deviance (in logic)", *History and Philosophy of Logic* **23**(2002), pp. 215-236, qui optent pour le critère sémantique.

peuvent pas être caractérisées par des matrices finies ; d'autre part, le gain de popularité des sémantiques de Kripke.[34]

Nous allons tenter de montrer que cette prétendue popularité de la seconde sémantique, celle qui triomphera dans l'histoire moderne de la logique modale, est liée à son héritage philosophique aussi bien qu'à sa plus grande simplicité d'application : non seulement l'argument de Dugundji montre que toute matrice de **S1-S5** est non finie, mais la sémantique relationnelle présente également l'avantage d'apporter une procédure de décision finie.

1.1.2.2.2. Sémantique relationnelle

La tradition plus récente de la sémantique modale, la sémantique relationnelle, s'imposa dans les années 1950 sous l'impulsion de Kripke, Kanger et Hintikka (entre autres).[35] L'inspiration de cette sémantique remonte bien avant la logique moderne : on associe traditionnellement la notion de 'monde possible' à la personne de Leibniz, qui aurait inspiré les définitions des modalités en termes de vérité relativement à des mondes. L'idée de base consiste à dire qu'un énoncé est *nécessairement* vrai s'il est vrai dans *tous* les mondes possibles, qu'il est *possiblement* vrai s'il est vrai dans *au moins un* monde possible.

La définition ci-dessus témoigne du lien logique étroit entre modalités et quantifications, et cette façon de définir est aussi pratique qu'évidente : elle permet d'expliquer des formules *intensionnelles* (c'est-à-dire non-vérifonctionnelles) en termes d'*extensions* dans plusieurs modèles, si l'on entend par 'modèle' la totalité des faits exprimés par des propositions et qui caractérisent l'état du monde.[36] Cette explication transparaissait déjà dans les tables de vérité classiques : si l'on considère chaque distribution de valeur de vérité de la dernière colonne comme un états de choses possible, alors les notions de tautologie, de contradiction et de simple satisfiabilité expriment respectivement le fait pour une formule d'être nécessairement vraie, nécessairement fausse et possiblement vraie. [Von Wright 1951] confirme cette affinité entre logique modale et logique du premier ordre :

> La logique des termes 'possible', 'impossible' et 'nécessaire' est, en d'autres termes, pour beaucoup similaire à la logique des termes 'quelque', 'aucun' et 'tous'. Il n'est pas surprenant que ce soit le cas, en fait. Pour le dire en termes populaires, le possible est ce qui est vrai dans certaines circonstances, l'impossible ce qui n'est vrai dans aucune circonstance, et le nécessaire ce qui est vrai en toute circonstance.[37]

L'analogie entre les mondes possibles de la logique modale et les tables de vérité de la logique classique s'obtient en quantifiant sur des assignations de valeurs de vérité. Dans une

---

[34] J.-Y. Béziau, "Many-valued and Kripke semantics", in *The Age of Alternative Logics*, J. van Benthem & G. Heinzmann & M. Rebuschi & H. Visser (ed.), Springer Dordrecht (2006), p. 90.

[35] Voir notamment [Kripke 1963], [Kanger 1957b] et [Hintikka 1957].

[36] Dans une sémantique relationnelle, les 'modèles' au pluriel ne désignent pas des *parties* de l'univers de discours, c'est-à-dire catégories particulières d'objets : un modèle pour les nombres, un autre modèle pour les êtres vivants, etc. Chaque modèle comprend la totalité de l'univers de discours et exprime la totalité des faits du monde, mais sous des configurations variables. La différence entre modèles *restreints* (qui expriment une partie des faits du monde seulement) et modèles *variables* (qui expriment la totalité des faits mais sous différents aspects) sera essentielle pour comprendre par la suite l'idée de logique comme calcul (voir section 2.3.1.2).

[37] [Von Wright 1951], p. 19. Sur le statut des modalités chez le premier Wittgenstein, voir [von Wright 1986]. Sur l'origine de l'expression 'mondes possibles' dans l'histoire de la logique, voir [Copeland 2002].

table de vérité bivalente, chaque colonne indique une 'proposition' atomique reflétant un état de choses, et la dernière colonne indique une proposition exprimant un fait ; chaque ligne exprime un ensemble d'état de choses ou monde possible et détermine le statut 'modal' de la formule : la nécessité correspond à l'assignation de la vérité pour chaque ligne, l'impossibilité à l'assignation de la fausseté pour chaque ligne, la contingence à l'assignation de la vérité pour certaines lignes et de la fausseté pour d'autres.

fait : littéraux / complexe

mondes possibles
états de choses
descriptions d'état

p	¬p	p∨¬p
1	0	1
1	0	1
0	1	1
0	1	1

'proposition' : tautologique
nécessaire

p	¬p	p∧¬p
1	0	0
1	0	0
0	1	0
0	1	0

contradictoire
impossible

p	q	p∨q
1	1	1
1	0	1
0	1	1
0	0	0

satisfiable
contingente

On constate un traitement semblable chez Carnap et Wittgenstein, le promoteur de la méthode des tables de vérité, puisque la théorie de l'image tableau [*Bildtheorie*] de ce dernier a été retranscrite par le premier en termes de *descriptions d'états* : une proposition nécessaire (ou analytiquement vraie) est une proposition vraie pour toute description d'état, c'est-à-dire pour toute description du monde en termes d'atomes propositionnels qui représentent des faits ; une proposition possible est vraie pour au moins une description d'état.[38] Si l'on imagine un monde composé de seulement trois propositions atomiques p, q et ¬r, alors {p, q, ¬r} est une description d'état de ce monde qui se compose de trois faits : *p* est le cas, *q* est le cas, *r* n'est pas le cas. Notons également que Wittgenstein concevait la bipolarité des propositions comme leur *possibilité* de vérité, et cette possibilité correspond également à la bivaluation des sémanticiens formels : toute proposition peut appartenir soit à la classe du vrai, soit à la classe du faux. Le lien entre mondes possibles et logique classique peut être développé comme suit.

Soit par exemple la proposition complexe ϕ : 'p ou q', dont la formule exprime un fait ou combinaison d'états de choses. Il y a quatre façons de combiner les atomes p et q, à savoir : p et q, p et non-q, non-p et q, non-p et non-q. Une table de vérité 'montre' cette évaluation des états de choses possibles en dressant la liste exhaustive des faits possibles, chaque combinaison possible de *littéraux* (un atome p ou sa négation non-p) constituant un 'monde'. Le lien entre tables de vérité, descriptions d'état et mondes possibles montre ainsi que la distinction marquée ici entre logique classique et logique modale est superficielle, ou tout au moins que l'on peut se dispenser des modalités tant que l'on considère la nécessité comme un

[38] On dit aussi d'une description d'état qu'elle est 'maximalement consistante' : l'adjonction d'une autre formule dans cette description la rend inconsistante (voir le glossaire : 'Consistance syntaxique'). Une description d'état est donc un ensemble qui contient, pour tout énoncé atomique du langage, soit un énoncé, soit sa négation (mais pas les deux), *et rien d'autre*. Nous expliquerons de cette manière la différence entre la nécessité selon Carnap et Kripke : les mondes possibles de Kripke ne sont pas maximalement consistants.

prédicat métalinguistique synonyme de tautologie.[39] Mais loin d'être inutile, cette comparaison conduira au contraire à la théorie de l'information de Carnap, puis à la sémantique des attitudes propositionnelles ; mais l'innovation majeure de la logique modale sera apportée par les modèles de Kripke, comme nous le verrons dans la prochaine section.

Suite à Wittgenstein, Carnap qualifiait d'*informative* ou non-triviale toute proposition qui nous apprend quelque chose sur l'état du monde. Une proposition atomique est bipolaire par définition et exprime un état de choses dont la réalisation ou l'absence la rend vraie ou fausse, contrairement aux 'pseudo-propositions' qui sont toujours vraies (tautologies) ou toujours fausses (antilogies) et ne disent rien sur le monde.[40] Les propositions contingentes sont celles dont on ne peut pas connaître la valeur de vérité *a priori* : elles nécessitent une vérification ou connaissance du fait qu'elles expriment, et cette connaissance une fois acquise permet d'exclure leur négation.

La notion de *connaissance* peut être expliquée par cette approche en termes de valuations. Supposons que l'agent *a* ignore si ψ est le cas ; il ne peut donc exclure ni la possibilité sa vérité, ni celle de sa fausseté. Supposons en revanche que *a* apprenne que p est le cas ; l'information alors acquise exclut de son stock de connaissance les états de choses disant le contraire, c'est-à-dire les deux situations possibles où p ne serait pas le cas : $(\neg p \vee q)$ et $(\neg p \vee \neg q)$. En d'autres termes :

(a) Si l'espace *logique* détermine la totalité des combinaisons possibles de valeurs de vérité, et cet espace se réduit chaque fois que l'agent est informé de la réalité de certains faits ;
(b) Un espace *épistémique* correspond à un sous-ensemble de l'espace logique qui exclut tous les états de choses en contradiction avec les connaissances acquises par l'agent (*omnis determino est negatio*).

Un accroissement de la connaissance entraîne une diminution des états de choses possibles, comme le souligne Hintikka :

> Supposons que je sache d'un énoncé qu'il est vrai. Je peux dès lors laisser par exemple de côté dans mes projets et décisions toutes les alternatives exclues par l'énoncé. De toute évidence, plus je peux exclure de possibilités, plus je peux dire que je sais. Si je ne peux pas exclure de possibilités, alors je ne sais absolument rien sur le sujet en question. (...) Cette connexion entre l'exclusion des possibilités et le degré d'information que fournit un énoncé a été explicitée tout d'abord par Karl Popper et mise en valeur par lui-même par la suite ; voir par exemple *La logique de la découverte scientifique* (...).[41]

---

[39] Hintikka fait état de l'affinité entre l'espace logique de Wittgenstein et sa propre sémantique des mondes possibles ; voir [Hintikka 1973a] : "Quantification and the picture theory of language", pp. 26-51.
[40] "Je ne sais rien au sujet du temps (qu'il fait) lorsque je sais qu'il pleut ou qu'il ne pleut pas" (*Tractatus*, §4.461). Sur la notion d'information, Voir Carnap, "Semantic Information", *The British Journal of the Philosophy of Science* **4**(1954), pp. 147-157, dans lequel l'auteur définit la tautologie comme un énoncé à information nulle. Voir également [Hintikka 1975b], en part. Ch. 5 : "Carnap's heritage in logical semantics", pp. 76-101.
[41] [Hintikka 1965], p. 183, note 7.

espace logique

p	q	p∨q
1	1	1
1	0	1
0	1	1
0	0	0

espace épistémique (*a* sait que p et ne sait pas si q)

Une façon similaire d'expliquer les modalités épistémiques passe par la sémantique relationnelle et l'idée de modèles. Si l'on considère chaque ligne d'une matrice comme un possible état du monde dans lequel un état de choses particulier a lieu, alors l'ensemble des possibilités d'assignation de valeur distinctes décrit différents modèles possibles, soit 2^n possibilités pour *n* propositions.

La notion de *relation* est nécessaire pour distinguer les états de choses qu'un agent peut admettre ou ne pas admettre en fonction de son stock d'informations ; elle correspond ici à une relation entre des mondes possibles, c'est-à-dire l'ensemble des propositions dont l'agent admet ou peut encore admettre la vérité. L'exemple précédent posait comme condition que cet ensemble ne contienne pas la proposition non-ϕ si *a* sait que ϕ. C'est donc un critère de *consistance* entre plusieurs ensembles de propositions qui exprime la relation minimale entre ce qu'un agent sait et les hypothèses qu'il peut encore former sur des faits inconnus.[42]

Nous allons décrire à présent la sémantique relationnelle des modalités épistémiques ; dans la mesure où [Hintikka 1962] a proposé une terminologie de celle en vigueur aujourd'hui, nous proposerons dans le même temps une analyse comparative des différentes notions utilisées par Hintikka et Kripke pour décrire la logique modale relationnelle.

1.2 Présentation d'une logique modale épistémique propositionnelle

Commençons par une définition abstraite de la sémantique relationnelle, avant de passer à son application aux attitudes épistémiques.

On distinguera ici trois principaux types d'interprétation dans un système logique : la logique classique non-modale, la logique modale de Carnap, et la logique modale de Kripke. Tout comme en logique classique, la logique modale de Carnap interprète ses théorèmes relativement à un modèle, mais ce modèle n'interprète plus les formules dans un ensemble unique et parcourt plusieurs ensembles qu'il appelle 'descriptions d'état'. Par opposition à la logique du premier Wittgenstein, celle du *Tractatus*, où le seul monde dans lequel un langage peut être interprété est le monde actuel (écrivons-le ainsi : w_0), la logique modale de Carnap fait parcourir son modèle sur plusieurs mondes et traite le monde actuel comme un sous-ensemble de l'ensemble des mondes possibles. Tout comme dans la logique modale de Carnap, la logique modale de Kripke interprète ses théorèmes dans plusieurs mondes appelés 'mondes possibles', mais elle admet également plusieurs modèles. Il y a donc un seul monde

[42] La sémantique relationnelle repose donc sur une clause de consistance : si ϕ appartient à un espace épistémique W, alors $\neg\phi$ n'appartient à aucun des mondes possibles w de W. Ce critère n'est-il pas davantage syntaxique que sémantique, puisqu'il décrit une combinaison de formules et ne dit rien sur leur référent ? Des auteurs se sont posé la même question au sujet de la sémantique des tableaux de Beth, parce que cette dernière produit une certaine confusion sur la distinction courante entre syntaxe et sémantique. Voir notamment Philippe de Rouilhan : "Les tableaux sémantiques : syntaxiques ou sémantiques ?", *Philosophia Scientiae* 3(4) (1999), pp. 305-322.

et un seul modèle, en logique classique ; plusieurs mondes et un seul modèle, dans la logique modale de Carnap ; plusieurs mondes et plusieurs modèles (les modèles de Kripke), dans la logique modale de Kripke. Hormis ces distinctions parmi les modèles utilisés pour interpréter les formules d'un langage, la procédure d'évaluation reste l'interprétation référentielle léguée par la sémantique de Tarski. Revenons-y un instant avant de spécifier les modèles de Kripke.

Conformément à la théorie des modèles mathématique, un énoncé se définit de deux façons complémentaires selon que le langage choisi se compose d'atomes propositionnels (de type *p*) ou de prédications (de type F*x*) : les extensions assignées aux atomes propositionnels sont des valeurs de vérité qui sont le vrai et le faux, tandis que les extensions assignées à des prédications sont des *n*-uplets liés par une relation d'arité *n*. Par exemple, l'énoncé 'Metz se situe au nord de Nancy' est interprété comme vrai dans la langue française parce que le monde actuel contient une paire ordonnée d'éléments ⟨Metz, Nancy⟩ qui satisfait la relation binaire 'se situer au nord de'. Quant aux énoncés modaux de la langue française tels que 'Metz est nécessairement au nord de Metz', leur interprétation exige un modèle plus structuré, donc plus riche que le modèle des énoncés déclaratifs.

Pour interpréter un langage modal propositionnel (calcul des énoncés) ou prédicatif (calcul des prédicats), on définit la sémantique relationnelle en termes de structure et de modèle sur une structure.

Une *structure* [*frame*] est un couple $\mathcal{F} = \langle W,R \rangle$, où W est un ensemble non vide et où R est une relation binaire sur l'ensemble W. Une structure est constituée d'un ensemble W de mondes possibles liés par une relation d'accessibilité R. Si par exemple le monde w* est accessible à partir de w, on écrira wRw* la relation binaire d'accessibilité entre w et w*. Dans [Hintikka 1962], les mondes possibles $w \in W$ sont appelés des 'ensembles modèles' et l'ensemble des mondes possibles W est appelé un 'système modèle', c'est-à-dire un ensemble d'ensembles modèles. Pour éviter la confusion entre le modèle $\mathcal{M}$ (qui correspond à la notion de modèle utilisée en théorie des modèles) et les modèles w, il sera préférable d'employer la terminologie actuellement en vigueur et de distinguer les notions de modèle et de mondes possibles : un monde possible w représente un ensemble de formules interprétées dans un modèle $\mathcal{M}$.

Pour interpréter un langage propositionnel, un *modèle* (de Kripke) $\mathcal{M}$ *sur une structure* $\mathcal{F}$ est un triplet $\mathcal{M} = \langle W,v,R \rangle$, où *v* est une valuation qui associe à chaque atome propositionnel *p* du langage propositionnel un sous-ensemble *v*(p) de W. Soit w un élément de W et soit □p une formule modale de ce langage. Pour exprimer le fait que □p est vraie dans le monde w du modèle $\mathcal{M}$, on a le choix entre des symbolisations interchangeables telles que :

$\mathcal{M} \models_w \Box p$ (□p est rendue vraie dans le monde w par l'interprétation du modèle $\mathcal{M}$) ;

$\Box p \in w \in W$ (notation de [Hintikka 1962] : 'p' appartient à l'ensemble modèle w du système modèle W) ; ou

$v(w,p) = 1$ (p est vraie dans le monde w).

Si une formule ϕ est un théorème du système **K**, c'est-à-dire valide ou vraie dans tous les mondes du modèle de **K**, par exemple, on dira de ϕ qu'elle est valide dans **K** ou **K**-valide, et nous écrirons cette validité en indexant **K** à la relation $\models$, soit : $\models_{\mathbf{K}} \phi$.

Pour interpréter un langage prédicatif, le modèle (de Kripke) correspondant est un quadruplet $\langle W,R,\mathcal{D},f \rangle$, où $\mathcal{D}$ est une fonction qui associe un domaine $\mathcal{D}_w$ à chaque monde w et

f une fonction qui associe une extensions aux symboles de la signature dans chaque monde w. Le domaine $\mathcal{D}$ se compose d'expressions fonctionnelles (prédicats et propriétés) et de constantes d'individus, et l'on dira d'une formule ϕ qu'elle est vraie dans un monde w si $\mathcal{D}$ associe à w un domaine $\mathcal{D}_w$ qui satisfait la formule ϕ.

L'innovation majeure apportée par les modèles de Kripke est la relation d'accessibilité R (que [Hintikka 1962] appelle 'alternativité') : contrairement à la définition de la nécessité de Carnap en termes de descriptions d'états, une formule nécessaire dans un modèle de Kripke n'est pas toujours une vérité logique et l'on peut distinguer plusieurs types de structures selon la propriété imposée à R. Contrairement à la logique modale de Carnap, où la classe des théorèmes est unique se confond avec la classe des vérités logiques, la vérité modale est modifiée sous l'instigation de plusieurs logiciens isolés parmi lesquels Guillaume, Bayart, Hintikka, Kanger et Kripke. La vérité modale n'est plus absolue et devient relative à plusieurs niveaux de discours, caractérisés par une ramification des niveaux de vérité et de validité en logique modale.

On distingue dans les modèles de Kripke quatre niveaux de vérité distincts selon leur ordre de puissance croissant : la vérité dans un modèle, la validité dans un modèle, la validité dans une classe de modèles (structure) et la validité modale. On peut les formuler ainsi :

Vérité dans un modèle
Une formule ϕ est vraie dans un modèle $\mathcal{M}$ si elle vraie dans au moins un monde w du modèle $\mathcal{M}$. Formellement : $w \models_{\mathcal{M}} \phi$

Validité dans un modèle
Une formule ϕ est valide dans un modèle $\mathcal{M}$ si elle est vraie dans tous les mondes w du modèle $\mathcal{M}$, c'est-à-dire si $w \models_{\mathcal{M}} \phi$ pour tout $w \in W$. Formellement : $\models_{\mathcal{M}} \phi$

Validité dans une classe de modèles
Une formule ϕ est valide dans une classe de modèles $C(\mathcal{M})$ si elle est valide dans tous les modèles de la classe $C(\mathcal{M})$, c'est-à-dire vraie dans tous les mondes de tous les modèles $\mathcal{M}$ de $C(\mathcal{M})$. Formellement : $\models_{C(\mathcal{M})}\phi$.
Comme on le verra plus bas, la correspondance entre la classe syntaxique des axiomes modaux et la classe sémantique des modèles de Kripke signifie qu'à chaque classe de structures correspond une classe de modèles : tout ensemble d'axiomes modal est modélisable, et tout ensemble de modèles de Kripke est axiomatisable.

Validité modale
Une formule ϕ est valide modale si elle est vraie pour toute structure $\mathcal{F}$. Formellement : $\models \phi$.

On peut ranger l'ordre des formules vraies en logique modale selon l'ordre d'inclusion suivant $\mathcal{M},w \models \phi \subseteq \mathcal{M} \models \phi \subseteq C(\mathcal{M}) \models \phi \subseteq \models \phi$. Comme nous le verrons plus bas dans la liste des propriétés structurelles de R, la relation de dépendance entre ces propriétés détermine l'ensemble des théorèmes d'un système modal : toute structure équivalente est également réflexive, symétrique et transitive, par exemple, donc les théorèmes de **S5** appartient à la classe des modèles à la fois réflexifs, transitifs et euclidiens {**K**,**T**,**4**,**5**}. Comme le système **K**

n'impose aucune propriété sur $\mathcal{F}$, il appartient à la classe des modèles la plus faible, à savoir la classe d'unité {**K**}.

Il existe ainsi divers ensembles de théorèmes dans les modèles de Kripke. Les différents types de relation d'accessibilité imposés entre les mondes (réflexive, transitive, euclidienne, etc.) permettent de distinguer plusieurs sortes de modèles en sémantique relationnelle, contrairement au modèle de Carnap où chaque monde est accessible à partir de n'importe quel autre. La classe des modèles réflexifs, par exemple, est la classe des modèles $\mathcal{M}$ où le schéma **T** ($\Box\phi \rightarrow \phi$) est valide; le schéma **4** ($\Box\phi \rightarrow \Box\Box\phi$) est valide, c'est-à-dire vrai dans tous les modèles transitifs, etc. Le modèle de Carnap correspond à un cas particulier des modèles de Kripke : c'est la classe maximale des modèles équivalents {**K**,**T**,**5**} qui valide le schéma **5** ($\Diamond\phi \rightarrow \Box\Diamond\phi$) ; à l'opposé, le système **K** représente le modèle de Kripke minimal parce qu'il correspond à un singleton, le modèle {**K**} sur une structure quelconque (où R n'a aucune propriété spécifique) qui contient toutes les tautologies, l'axiome de distribution et la règle de nécessitation. Nous reviendrons dans le second chapitre sur l'importance qu'ont prises les innovations en théorie des modèles dans l'histoire de la logique modale, ainsi que sur leurs applications en logique philosophique. Le premier effet de ces innovations est un langage modal plus riche, capable de distinguer d'autres concepts que la tautologie et la contradiction. A partir d'une variation des propriétés de R, la logique philosophique va proposer une traduction de plusieurs concepts des langues naturelles en termes de modèles de Kripke. En voici trois exemples.

L'interprétation informelle de la relation R et des éléments w de W varie selon le type de logique modale interprétée. Dans une logique modale *temporelle*, w correspond à un certain instant du temps W, et w*Rw signifie que w* est ultérieur à w; dans une logique modale *topologique*, w correspond à un certain point de l'espace W, et wRw* signifie que w* se situe à un autre point ; dans une logique modale *épistémique*, w correspond à un état de l'ensemble des informations d'un agent W, et w*Rw signifie que w* contient une somme d'informations au moins égale à celle de w.[43] L'espace épistémique, c'est-à-dire l'ensemble des informations dont dispose un agent et qui s'exprime par des ensembles de propositions connues, s'organise en vertu de certaines conditions de consistance minimale entre des états d'information w, w* dans W :

> (...) je sais un énoncé si cet énoncé est valable dans tous les mondes que je ne peux pas *distinguer* de mon monde réel, c'est-à-dire s'il est valable à travers mon 'champ d'incertitude' dans son ensemble.[44]

La définition logique des opérateurs épistémiques est la suivante, avec K_a et B_a désignant respectivement la connaissance et la croyance de l'agent *a*. On retrouve dans cette définition l'analogie entre les opérateurs modaux et les quantificateurs du premier ordre. Soit ϕ un énoncé quelconque (atomique ou complexe) d'un modèle, et soit $\{w,w^*,\ldots,w^{*n}\}$ un ensemble

[43] "Les conditions dans lesquelles nous essayons de saisir ici la logique du savoir et de la croyance sont en termes de certaines alternatives à un état de choses. En gros, ces alternatives sont des états de choses dans lesquels une certaine personne en sait au moins autant –voire plus, d'habitude –qu'il n'en sait dans l'état donné. En bref, nous nous intéressons aux différentes possibilités en présence pour qu'une personne obtienne une information supplémentaire" ([Hintikka 1962], p. 56).

[44] [van Benthem 1991], p. 5.

de *n* mondes. Les modalités de connaissance et de croyance sont définies comme suit dans [Hintikka 1962], les paires K/P et B/C correspondant à l'interprétation épistémique des modalités pures $\Box/\Diamond$. Sous chaque caractérisation de Hintikka est associée sa traduction dans la terminologie des modèles de Kripke, où R_k désigne la relation d'accessibilité épistémique et R_b la relation d'accessibilité doxastique :

Connaissance

(C.K*) $K_a\phi \in w$ ssi $\phi \in w^*$ pour tout w* alterne à w (par rapport à *a*)[45]

= (**K**) $w \models_K K\phi$ ssi $\forall w^* R_k w$, on a $w^* \models_K \phi$

(C.P*) $P_a\phi \in w$ ssi $\phi \in w^*$ pour au moins un w* alterne à w

= (**K**) $w \models_K P\phi$ ssi $\exists w^* R_k w$, on a $w^* \models_K \phi$

Croyance

(C.B*) $B_a\phi \in w$ ssi $\phi \in w^*$ pour tout w* alterne à w

= (**K**) $w \models_K P\phi$ ssi $\forall w^* R_b w$, on a $w^* \models_K \phi$

(C.C*) $C_a\phi \in w$ ssi $\phi \in w^*$ pour au moins un w* alterne à w

= (**K**) $w \models_K P\phi$ ssi $\exists w^* R_b w$, on a $w^* \models_K \phi$

La logique modale épistémique étudiée ici est une logique *enrichie*, ce qui veut dire qu'elle représente une extension de la logique classique et que la signification d'une formule modale dépend également des connecteurs logiques standards. Dans la section précédente, la comparaison faite entre l'espace logique et l'espace épistémique a montré que ce dernier espace est inclus dans le premier ; il conserve ainsi les propriétés des connecteurs habituels : la manière dont l'agent organise son espace de connaissance est établie à partir des règles de la logique classique. Croire que la proposition $(p \wedge q)$ 'est le cas' (exprime un fait actuel) engage l'agent à croire que les deux états de choses exprimés par les atomes propositionnels p et q sont le cas, par exemple.

Le résultat est un ensemble de conditions sémantiques conformes à la logique classique : la logique modale épistémique *étend* ainsi les propriétés du calcul des propositions tout en les *conservant*. On trouve dans [Hintikka 1962] les conditions sémantiques de base (C.) qui suivent, auxquelles on ajoutera de nouveau la traduction dans les modèles de Kripke :[46]

(C.¬) $\phi \in w$ ssi $\neg\phi \notin w$

= $w \models \phi$ ssi $w \not\models \neg\phi$

(C.∧) $(\phi \wedge \psi) \in w$ ssi $\phi \in w$ et $\psi \in w$

= $w \models (\phi \wedge \psi)$ ssi $w \models \phi$ et $w \models \psi$

(C.∨) $(\phi \vee \psi) \in w$ ssi $\phi \in w$ ou $\psi \in w$

---

[45] Dans une logique d'attitudes telle que la logique épistémique, R se rapporte à un sujet et chaque ensemble modèle correspond à un ensemble d'états de choses imaginé par lui.

[46] Si la plupart de ces conditions sémantiques contiennent la négation, c'est parce que, rappelons-le, la sémantique relationnelle raisonne par l'absurde : pour montrer que ϕ est valide, on suppose d'abord la négation ϕ afin de montrer ensuite que $\neg\phi$ entraîne une contradiction. Le fait que la réfutation par l'absurde présuppose le tiers exclu et la double négation (C.¬¬) indiquent nettement que l'inspiration *classique* de la sémantique modale de [Hintikka 1962].

=	$w \vDash (\phi \vee \psi)$ ssi $w \vDash \phi$ ou $w \vDash \psi$ (ou les deux)
(C.$\neg\wedge$)	$\neg(\phi \wedge \psi) \in w$ ssi $\neg\phi \in w$ ou $\neg\psi \in w$ (ou les deux)
=	$w \vDash \neg(\phi \wedge \psi)$ ssi $w \vDash \neg\phi$ ou $w \vDash \phi$ (ou les deux)
(C.$\neg\vee$)	$\neg(\phi \vee \psi) \in w$ ssi $\neg\phi \in w$ ou $\neg\psi \in w$ (ou les deux)
=	$w \vDash \neg(\phi \vee \psi)$ ssi $w \vDash \neg\phi$ et $w \vDash \neg\psi$
(C.$\neg\neg$)	$\neg\neg\phi \in w$ ssi $\phi \in w$
=	$w \vDash \neg\neg\phi$ ssi $w \vDash \phi$

Puisqu'un espace épistémique conserve les opérations classiques, alors les théorèmes de la logique classique seront préservés dans ces espaces et un agent n'admettra pas d'états de choses qui violent les règles de la logique classique : pour tout ϕ et ψ, si $\phi \in w$ et si ψ est une conséquence de ϕ, alors $\psi \in w$. Un ensemble d'états de choses contradictoire n'est pas une alternative envisageable en logique épistémique, tant que celle-ci définit un état d'information minimal par les propriétés d'une logique modale *normale*.

Pour justifier l'expression de ces conditions sémantiques en termes de modèles, Hintikka fait appel à la notion de *partition*. Selon lui, toute acquisition d'une information implique une élimination de certaines alternatives, et l'opérateur qui exprime cette acquisition nécessite ainsi une distinction entre deux types d'ensembles modèles : ceux qui sont compatibles avec l'attitude en question, et ceux qui ne le sont pas :

> Mon hypothèse fondamentale (dont je simplifie la formulation à l'extrême) est que l'attribution d'une attitude propositionnelle quelconque à la personne en question implique une *partition* de tous les mondes possibles (...) en deux classes contenant, l'une, des mondes possibles qui *concordent* avec l'attitude en question, et l'autre, ceux qui sont *incompatibles* avec elle.[47]

Nous avons dit que les opérateurs de connaissance et de croyance se comportent comme l'opérateur modal fort $\Box$, ce qui signifie que la connaissance ou la croyance par *a* de ϕ exclut non-ϕ de son espace épistémique : il n'est pas possible que *a* admette la négation de ce qu'il sait ou croit, sous peine de se contredire et de violer la condition minimale (C.$\neg$). Par analogie avec la relation duale entre $\Box\phi$ et $\neg\Diamond\neg\phi$, la condition de consistance introduit ainsi une *clause de rationalité* minimale pour chaque agent épistémique, quel que soit le type d'attitude en question. On retrouve dans [Hintikka 1962] ces critères de consistance des attitudes propositionnelles :

(C.$\neg$K)	$\neg K_a\phi \in w$ ssi $P_a\neg\phi \in w$
= (**K**)	$w \vDash_{\mathbf{K}} \neg K_a\phi$ ssi $w \vDash_{\mathbf{K}} P_a\neg\phi$
(C.$\neg$P)	$\neg P_a\phi \in w$ ssi $K_a\neg\phi \in w$
= (**K**)	$w \vDash_{\mathbf{K}} \neg P_a\phi$ ssi $w \vDash_{\mathbf{K}} K_a\neg\phi$

[47] [Hintikka 1969b], p. 41.

En d'autres termes, l'ignorance de ϕ signifie la possibilité de sa négation $\neg\phi$, et l'impossibilité de ϕ signifie la connaissance de sa négation $\neg\phi$. Cette relation de dualité vaut aussi bien entre K et P qu'entre B et C, quel que soit *a*.

En revanche, les propriétés de R peuvent différer selon l'interprétation donnée à l'opérateur modal. La propriété logique la plus 'intuitivement évidente' de l'opérateur K est la *clause de vérité* (thèse **T**), que l'on peut appeler également la clause de simplification : pour n'importe quel modèle w de W, si *a* sait que ϕ alors ϕ est vraie,

(C.K) Si $K_a\phi \in w$, alors $\phi \in w$
= (**T**) Si $w \models_T K_a\phi$, alors $w \models_T \phi$

La croyance n'est pas caractérisée par cette clause de vérité, par contre, puisqu'il est possible d'avoir des croyances fausses : (C.B) n'est donc pas un théorème épistémique, donc il est possible d'avoir un ensemble modèle w dans lequel $B_a\phi \in w$ et $\neg\phi \in w$. Une autre propriété attribuée à l'opérateur K dans [Hintikka 1962], plus discutable, est la clause d'itération (thèse KK) ou *clause d'introspection positive* : si *a* sait que ϕ, alors *a* sait qu'il sait que ϕ,[48]

(C.KK*) Si $K_a\phi \in w$, alors $K_a\phi \in w^*$ pour tout w* alterne à w (par rapport à *a*)
= (**4**) Si $w \models_{S4} K_a\phi$, alors $w \models_{S4} K_aK_a\phi$, c'est-à-dire: $\forall w^*Rw$, on a $w^* \models_{S4} K_a\phi$

Dans [Hintikka 1962], le 'système modèle' W (l'ensemble des mondes w) qui caractérise l'opérateur K est interprété dans une classe de modèles particulière, celle où les clauses de consistance, de vérité et d'itération sont validées. Le concept de connaissance y est spécifié ainsi par les axiomes **K** + **T** + **4** = **KT4**, interprétés dans la classe des modèles transitifs {**K**,**T**,**4**}. A noter que, en vertu de la caractérisation de K par les schémas **T** : $K\phi \rightarrow \phi$ et **4** : $K\phi \rightarrow KK\phi$, on peut dériver la validité de $KK\phi \rightarrow K\phi$ par substitution de $K\phi$ à ϕ (donc $\models_T KK\phi \rightarrow K\phi$) et obtenir ainsi l'équivalence $K\phi \equiv KK\phi$ (donc $\models_{T4} K_a\phi \equiv K_aK_a\phi$). Cette équivalence permet la réduction d'un nombre *n* quelconque de modalités identiques à une seule : savoir que l'on sait que l'on sait que ... l'on sait que ϕ équivaut au simple fait de savoir que ϕ est le cas, pour une suite indéfinie de déclarations de savoir de type $K\phi$.

Le concept de croyance est spécifié dans [Hintikka 1962] par les axiomes **K** + **D** + **4** = **KD4**. L'opérateur de croyance satisfait donc la relation de transitivité, tout comme l'opérateur de connaissance : $\models_{S4} B_a\phi \rightarrow B_aB_a\phi$. Mais à la différence de ce dernier, la propriété réflexive de **T** ne fait pas partie des caractéristiques de l'opérateur de croyance (donc $\not\models_T B_a\phi \rightarrow \phi$), $B\phi \equiv BB\phi$ n'est pas valide. Conformément à la relation de dualité entre $\Diamond\phi$ et $\neg\Box\neg\phi$, **D** rappelle le fait qu'un agent rationnel ne doit pas croire une formule et sa négation à la fois (donc $\models_D B_a\phi \rightarrow \neg B\neg\phi$).[49]

---

[48] Le pronom personnel 'il' a un usage anaphorique et renvoie ici à l'agent *a*, bien entendu. Mais comme la référence d'un pronom personnel peut varier selon les contextes de discours, cette ambiguïté donnera lieu à des critiques sur la formulation de la thèse KK. Voir notamment Castañeda [1970]; [Hintikka 1970a] proposera en réponse une version quantifiée plus précise de la thèse KK.

[49] Si l'on laisse la dualité de côté, le théorème de **D** : $\Box\phi \rightarrow \Diamond\phi$ signifie aussi que le système modèle W est non vide ; s'il était vide, alors on serait exposé au même problème que celui posé par la relation de *subalternation*

Au total, on peut dresser quelques propriétés caractéristiques de K et B selon le tableau ci-dessous : chaque axiome modal a sa contrepartie sémantique exprimée par une propriété spécifique de R. L'interprétation épistémique du schéma **5** est appelée la *clause d'introspection négative*, et Hintikka ne l'associe pas à l'opérateur K parce que son interprétation informelle aboutit à une absurdité : si tout agent savait ce qu'il ne sait pas, il devrait être capable de déterminer l'étendue de son ignorance, à la manière du sage Socrate ($\models_{S5} \neg K_a\phi \rightarrow K_a\neg K_a\phi$). Cette clause est admissible dans un autre contexte d'usage du concept de connaissance, lorsqu'il s'agit de modéliser le comportement 'intelligent' des systèmes distribués ; mais une telle interprétation ne nous intéresse pas dans le cadre de ce travail : nous nous intéressons avant tout à l'impact de la logique épistémique sur la réflexion en théorie de la connaissance, et celle-ci porte en particulier sur la connaissance humaine.

Correspondance modale entre syntaxe (axiomes) et sémantique (modèles de Kripke)

Axiomes	**Propriétés de R**
K $(\phi \rightarrow \psi) \rightarrow \Box(\phi \rightarrow \psi)$ $\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi)$	Aucune
T $\Box\phi \rightarrow \phi$	Réflexivité $\forall w(wRw)$
B $\phi \rightarrow \Box\Diamond\phi$	Symétrie $\forall w\forall w^*(w^*Rw \rightarrow wRw^*)$
4 $\Box\phi \rightarrow \Box\Box\phi$	Transitivité $\forall w\forall w^*\forall w^{**}[(w^*Rw \wedge w^{**}Rw^*) \rightarrow (w^{**}Rw)]$
D $\Box\phi \rightarrow \Diamond\phi$	Sérialité $\forall w\exists w^*(w^*Rw)$
5 $\Diamond\phi \rightarrow \Box\Diamond\phi$	Euclidianité $\forall w\forall w^*\forall w^{**}[(w^*Rw \wedge w^{**}Rw) \rightarrow w^{**}Rw^*]$

Selon [Hintikka 2003], la première génération de la logique épistémique (la période moderne de la logique modale) s'est principalement consacrée à la recherche de la 'bonne' logique épistémique :

---

dans la scolastique médiévale : 'tout S est P' est vraie et 'quelque S est P' est fausse lorsque S est un terme vide, et l'analogie entre quantificateurs et modalités signifierait que le nécessaire n'implique pas toujours le possible dans un modèle de Kripke ($\not\models \Box\phi \rightarrow \Diamond\phi$).

> Ce que Hintikka a récemment surnommé 'logique épistémique première génération' est caractérisé par l'ambition qu'un catalogage des systèmes complets possibles de ces logiques permette de choisir la (les) plus 'appropriée(s)' ou 'intuitive(s)'.[50]

Ce programme de formalisation a-t-il un sens, dans la mesure où il suppose que certains concepts des langues naturelles sont pleinement caractérisables en termes de structures ? S'il en a un, alors la logique épistémique prouvera pour sa part que les critiques anti-formalistes de la philosophie du langage ordinaire reposaient sur une connaissance insuffisante des ressources de la logique.

Avant de discuter le bien-fondé de la logique épistémique première génération, nous reviendrons d'abord sur quelques enjeux de cette logique modale interprétée. Un outillage formel de ce genre permet-il de mieux comprendre des concepts philosophiques par le biais du formalisme logique (quitte à réviser ces formalismes lorsque le besoin s'en fait sentir, par exemple)? A l'inverse, n'est-ce pas un problème philosophique en soi que de vouloir traiter ces concepts par le biais du langage logique ? Nous n'aurons pas la prétention de vouloir apporter une réponse définitive à ce vaste problème, concernant les mérites et limites de la philosophie formelle ; nous chercherons plus modestement à l'illustrer et voir ce qui peut se faire effectivement dans le cadre d'une analyse formelle de ce genre.

1.3 Les enjeux d'une sémantique relationnelle

La sémantique algébrique a été progressivement délaissée en logique modale au profit d'un autre traitement des modalités : celui de la sémantique dite des 'mondes possibles'. Quels avantages cette approche relationnelle apporte-t-elle pour les modalités épistémiques, et quelles difficultés philosophiques a-t-elle créé en contrepartie ? L'avantage est d'ordre explicatif : la sémantique relationnelle restaure l'analyse extensionnelle des énoncés. L'inconvénient est d'ordre ontologique : l'existence des individus dans des 'mondes possibles' pose problème, tout comme l'idée de 'quantifier sur un monde'.

1.3.1 Extensionalisme plural

La logique modale *aléthique* reposait d'abord sur une distinction catégorique entre extension et intension : pour préserver la vérité dans un contexte modal, le critère de l'identité des extensions avait été remplacée par celui d'identité des intensions. Dans cette perspective, la co-extensionalité des expressions n'est plus une condition suffisante pour déterminer la signification des énoncés, et les notions de concept d'individu ou de postulats de signification furent introduites par plusieurs auteurs (Church, Carnap) pour garantir le principe de substitution uniforme. Si l'on suit cette démarche, la signification ne se réduit pas à une analyse des formules en termes de classes et nécessite le recours à des attributs. La vérité de $\Box\phi$ ne serait préservée par $\Box\psi$ que si ϕ et ψ sont non seulement co-extensionnels mais co-intensionnels (ont même attribut, c'est-à-dire partagent la même propriété).

Or si la valeur de vérité est une notion extensionnelle insuffisante, à quelle autre valeur sémantique faut-il recourir pour comprendre la signification d'une expression modale ? Si l'on prend les énoncés ϕ : 'les célibataires ne sont pas mariés' et ψ : 'l'homme est un rationnel

[50] [Hendricks 2003], p. 93.

bipède', les couples de propriétés 'célibataire'/'non marié' et 'homme'/'rationnel bipède' semblent être co-intensionnels dans la mesure où l'on suppose que ϕ et ψ sont vrais nécessairement. La substitution *salva veritate* de ϕ par ψ repose alors sur une condition renforcée de substitution *salva congruate* : deux formules modales ont la même signification si elles ont *toujours la* même valeur de vérité, mais le fait d'être toujours vrai n'a de sens que si l'on introduit la notion d'*analyticité* (un célibataire est non marié *par définition*) ou d'attribution essentielle (l'homme est un bipède rationnel, *par nature*) pour justifier l'existence de vérités logiques modales. La notion d'analyticité est-elle indispensable à la logique modale ?

Nous ne reviendrons pas en détail sur la critique de Quine contre la notion d'intensionalité : après avoir déconstruit le dogme empiriste de la distinction entre vérités synthétiques et vérités analytiques, donc après avoir sapé les bases de ce qui constituait pour Carnap et Church un moyen d'expliquer les modalités, [Quine 1961b] explique que le critère de l'identité de signification repose sur la notion arbitraire de vérité analytique ; en effet, un tel critère suppose que l'on ait un manuel de traduction pour déterminer l'ensemble des propriétés synonymes, mais ce manuel n'est concevable que dans les limites d'un langage relativement pauvre. Un langage artificiel peut satisfaire le critère du postulat de signification de Carnap, mais pas les langues naturelles. Si le projet d'une logique intensionnelle repose sur des identités d'intensions exprimées en termes de synonymie ou de postulats de signification, elle est compromise dès qu'elle veut s'étendre aux langues naturelles car

> cela reviendrait en outre, pour les besoins d'une quantification de ce genre, à n'admettre aucun objet concret et seulement ce que Frege appelait les sens des noms, et ce que Carnap et Church ont appelé des concepts d'individus. L'inconvénient d'une telle ontologie, c'est que le principe d'individuation de ses entités repose invariablement sur la notion putative de synonymie, ou d'analyticité.[51]

En plus d'être arbitraire ou insuffisant, le recours aux intensions semble menacer par ailleurs le caractère purement formel de la logique : s'il faut tenir compte du sens des concepts pour garantir leur substitution *salva congruate*, la logique modale est-elle encore formelle à proprement parler ? Contrairement aux partisans de l'extensionalisme, en vertu duquel une vérité logique reste vraie lorsque l'on y remplace toutes les expressions du lexique par d'autres expressions de même extension, la vérité logique ne peut plus se contenter des extensions si la logique modale exige pour sa part une identité de sens ou synonymie entre ces expressions. Or si le sens, c'est-à-dire le *contenu* de signification, doit être pris en compte en vue d'une analyse logique des constructions modales, cela risque d'amener à la conclusion que la logique n'est pas une science des formes pures et qu'elle doit finalement tenir compte de la matière du discours. Bien que Carnap ait proposé de garantir l'aspect formel de la substitution uniforme par un critère d'iso*morphie* intensionnelle (deux expressions sont co-intensionnelles si elles ont la même structure), l'identité de structure ne va pas de soi et un tel critère ne fonctionne pas par ailleurs dans n'importe quel système modal interprété. Le problème de l'*hyperintensionalité*, évoqué plus loin, montre les limites de cette procédure intensionnelle dans un contexte épistémique.

[51] [Quine 1961d], p. 152.

Face à l'approche intensionnelle précédente, la sémantique relationnelle propose de remplacer l'identité de deux intensions dans un modèle par l'identification d'une seule et même extension dans plusieurs modèles : être co-intensionnel dans un monde, c'est être une même extension dans tous les mondes possibles. [Carnap 1947] a lui-même emprunté cette démarche extensionnelle, lorsqu'il proposa de redéfinir la synonymie en termes de vérité dans *toute* description d'état.[52] [Gardiès 1979] nous rappelle que cette procédure ne date pas de la logique modale moderne mais remonte à l'Antiquité, en marge de la logique traditionnelle d'Aristote. Les logiciens de l'Ecole de Mégare avaient anticipé en quelque sorte les descriptions d'état de Carnap : dans la première prémisse de l'Argument Dominateur de Diodore Cronos, le possible est défini comme ce qui est ou sera, et la nécessité diodorienne est présentée comme la vérité à *tous* les instants du temps. Le parallèle avec la logique modale temporelle est évident, permettant ainsi une réduction des intensions à des extensions multiples :

> Ainsi l'originalité de l'analyse mégaro-stoïcienne des modalités ontiques tient-elle à ce qu'elle réussit à dépouiller les représentations ordinaires de ces modalités de ce qu'elles peuvent apparemment comporter de rigoureusement intensionnel. Dans une telle transcription, la modalité s'analyse, au même titre que la simple universalité, en strictes valeurs de vérité et de fausseté. Il n'y a plus de καθ'αυτο irréductible au κατα παντος Les mégariques montraient la voie dans laquelle s'engagent beaucoup de travaux récents.[53]

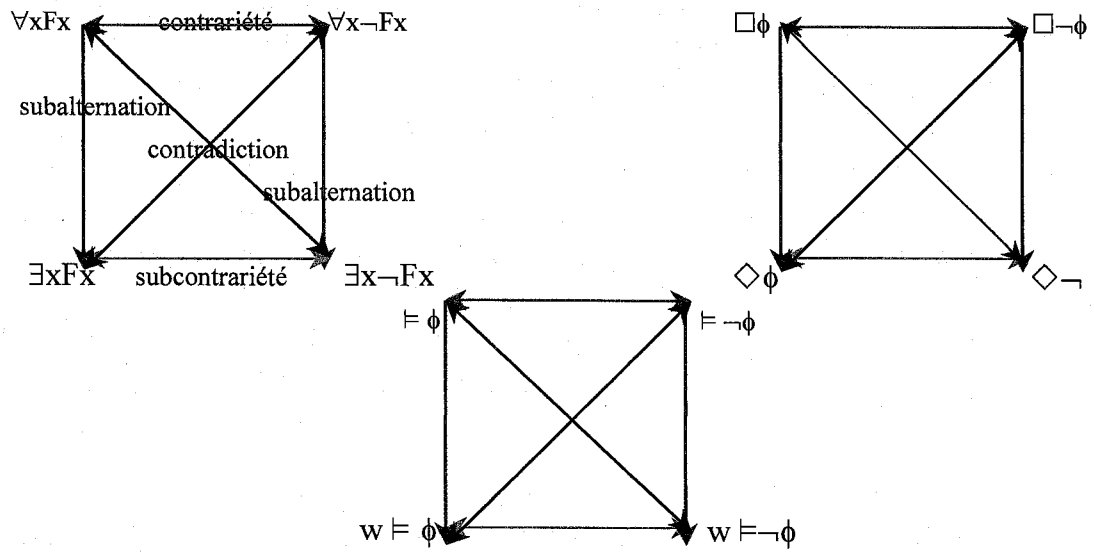
Toute la question reste de savoir si, au-delà de sa commodité, la réduction des intensions à des extensions n'enlève rien à l'information transmise par les premières. Le débat est connu, il concerne le fossé à combler entre les notions d'attribut et de classe : un attribut est-il réductible à une classe constante, c'est-à-dire un ensemble d'objets qui satisfont une propriété dans *tous* les cas (mondes possibles), et a-t-on vraiment compris le sens d'un attribut lorsque l'on affirme qu'il est *toujours* satisfait par les mêmes objets ? Quoi qu'il en soit, [Gochet 1972] parle d'*extensionalisme plural* pour qualifier ce processus de réduction d'une intension au singulier à des extensions au pluriel : la vérité modale dans un modèle ne serait rien de plus que la vérité dans plusieurs modèles. Pour prolonger l'exemple antique, on peut représenter les modalités au sein d'un carré des oppositions logiques ; la comparaison précédente entre modalités, quantification et vérité dans un monde donne ces structures isomorphes :

---

[52] [Quine 1961b] reproche toutefois à cette redéfinition de confondre deux notions distinctes : la synonymie (vérité analytique) et la tautologie (vérité logique). Chez Carnap, une vérité logique ou 'L-vérité' est vraie dans toute description d'état ; or comment garantir que deux propositions synonymes auront même extension dans toutes les descriptions d'état ? "Le critère d'analyticité en termes de descriptions d'état ne vaut ainsi que pour les langages dénués de paires de synonymes extralogiques, telles 'célibataire' et 'homme non marié' –des paires de synonymes du type de celles qui donnaient lieu à la 'seconde classe' d'énoncés analytiques. Le critère en termes de descriptions d'état est une réinterprétation de la vérité logique tout au plus, non de l'analyticité" (pp. 23-24). Quine réitèrera cette critique pour les contextes épistémiques : comment garantir un critère d'identité extensionnelle suffisant entre des mondes possibles ?

[53] [Gardiès 1979], p. 39. La réduction du καθ'αυτο au κατα παντος signifie la réduction de la nécessité ontique (vraie en *soi*) à la nécessité temporelle (vraie à *tout* instant). Aristote n'admettait pas la réduction des modalités ontiques à des modalités temporelles : la nécessité temporelle est une condition *nécessaire* à la nécessité ontique, mais elle n'est pas une condition *suffisante*. Si cette remarque concerne la nécessité absolue de Carnap, elle ne concerne pas en revanche la nécessité relative (à R) des modèles de Kripke : une vérité est nécessaire relativement à un instant du temps qui le précède (les vérités passées sont nécessaires, selon Diodore), mais elle n'est pas forcément éternelle ; les vérités éternelles correspondent uniquement aux formules **S5**-valides.

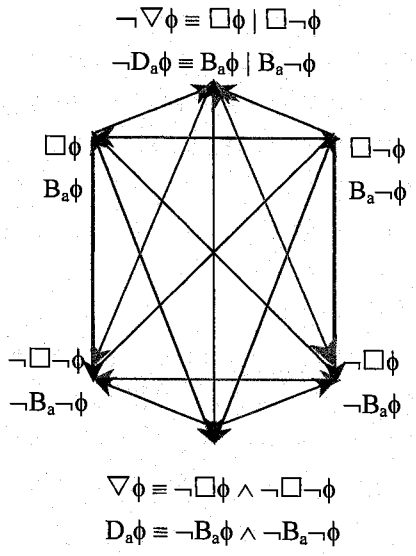
Analogie structurelle entre quantification, modalité, et vérité dans un monde



Les trois schémas montrent une même structure pour plusieurs notions distinctes ; modalité, quantification et vérité ont un comportement logique commun. L'avantage de cette schématisation est-il strictement technique et artificiel, ou est-il possible de tirer un enseignement de ces identités de structure pour notre compréhension de plusieurs concepts des langues naturelles ? Le cas des modalités aléthiques avait montré un relatif succès du processus de réduction : ce qui est nécessaire n'est pas impossible, le nécessaire implique le réel et le réel implique le possible. Ces différentes validités modales permettent de retranscrire correctement et par un simple jeu de structures formelles le sens attribué à plusieurs concepts ordinaires.

De ce fait, l'extension des modalités ontiques à toute une famille de concepts modaux devrait montrer qu'il est possible de *comprendre* (donner les règles de signification de) plusieurs concepts philosophiques en termes d'opérateurs formels. Si de telles schématisations fonctionnent pour les attitudes propositionnelles, la logique épistémique devrait montrer également que le projet de logique philosophique est réalisable. Pour conforter cette opinion, on peut poursuivre la méthode des oppositions et formaliser d'autres concepts philosophiques. [Hart 1980] a étendu ainsi la sémantique relationnelle de [Hintikka 1962] à la notion de *doute*. La relation entre croyance et doute correspond à la relation entre un carré d'Apulée et un hexagone de Blanché, dans lequel la notion de doute D correspondrait à une sorte de possibilité bilatérale ou *contingence épistémique* : de même que ϕ est contingent si ni ϕ ni non-ϕ ne sont nécessaires : $\nabla\phi =_{df} \neg\Box\phi \wedge \neg\Box\neg\phi$, *a* doute que ϕ soit le cas si *a* ne croit ni que ϕ ni que non-ϕ : $D_a\phi =_{df} \neg B_a\phi \wedge \neg B_a\neg\phi$.[54]

[54] Cet hexagone ne retranscrit pas exactement la notion de doute et exigerait quelques complexifications ; comme le note Hart, en effet, le doute ne signifie pas la simple absence de croyance mais plutôt la croyance de l'absence de croyance, ce qui donne $D\phi \equiv B(\neg B\phi \wedge \neg B\neg\phi)$. Le doute ne serait donc pas une contingence épistémique simple ∇, mais une contingence nécessaire épistémique $\Box\nabla$. Une meilleure représentation de cette modalité complexe sera proposée en section 3.3.3.2, en liaison avec la notion de dénégation.



La schématisation ci-dessus semble prometteuse, si elle s'applique avec succès à un nombre indéfini de concepts ordinaires. A l'aide de la conjonction et de la négation, des expressions telles que 'douter' ou 'ne pas savoir si' peuvent être construites sur une même structure modale de type exclusif (ni … ni) : $\neg\Box\phi \wedge \neg\Box\neg\phi$.[55] Mais le succès relatif de ces constructions suffit-il à produire une logique philosophique intéressante, notamment une logique de l'épistémologie pour ce qui nous concerne ? Même s'il est possible de construire d'autres formes polygonales, plus complexes, pour représenter les autres propriétés des opérateurs épistémiques (distribution, simplification, itération),[56] l'élégance du résultat ci-dessus ne supprime pas le problème que rencontrent l'axiomatique de Lewis ou la sémantique relationnelle de Kripke : dans les trois cas, la difficulté n'est pas tant de trouver une méthode de preuve effective que de garantir leur adéquation informelle avec les notions des langues naturelles. Nous verrons que, dans une logique philosophique, chaque méthode de preuve logique est confrontée aux mêmes problèmes philosophiques. En logique propositionnelle, des applications douteuses voire paradoxales fragilisent le succès de la formalisation des attitudes propositionnelles : comme nous le verrons par la suite (section 2.2.4), une vaste littérature s'est formée sur la critique de deux axiomes de [Hintikka 1962], à savoir l'introspection positive et, surtout, l'omniscience logique. En logique des prédicats, le problème concerne les critères d'identité des extensions entre plusieurs 'mondes' et alternatives épistémiques. Quine fera de ce problème d'individuation un argument majeur

[55] Sur les relations d'opposition observées dans plusieurs groupes de concepts des langues naturelles, voir R. Blanché : *Structures intellectuelles, essai sur l'organisation systématique des concepts*, Paris, Vrin (1966) ; ainsi que [Gardiès 1975], en part. pp. 17-21, et [Vernant 2003].

[56] Voir Alessio Moretti : "Geometries for Modalities? Yes: Through n-Opposition Theory"; in *Aspects of Universal Logic*, J.-Y. Béziau & A. Costa-Leite & A. Facchini (éds.), Neuchâtel (2004), pp. 102-145. L'identité des structures ci-dessus montre que trois notions distinctes (quantificateurs, vérité, modalité) se comportent toutes comme un même opérateur unaire. Une généralisation des propriétés de la relation de conséquence logique est apportée avec l'opérateur de conséquence Cn de Tarski : appliqué à un ensemble X de propositions, Cn est définie comme réflexive ($X \subseteq Cn(X)$), monotone (si $X \subseteq Y$, alors $Cn(X) \subseteq Cn(Y)$) et transitive ($CnCn(X) \subseteq Cn(X)$). On obtient ici une abstraction des systèmes modaux **T**, **K** et **S4**, avec $Cn = \Diamond$. Voir A. Tarski, "Remarques sur les notions fondamentales de la méthodologie des mathématiques", *Annales de la Société Polonaise de Mathématique* (1928).

contre l'applicabilité des logiques modales, auquel Hintikka répondra par l'argument de la référence multiple.

La sémantique relationnelle permet en tout cas une réduction des intensions à une pluralité d'extensions ; elle apporte également une preuve de complétude du système modal, qualité métalogique qui est censée garantir le bien-fondé du système. La complétude de $\mathcal{L}_f$ est-elle un gage d'adéquation informelle entre la signification des validités dans $\mathcal{L}_f$ et la compréhension des expressions dans $\mathcal{L}_n$? C'est cette question que nous allons aborder dans ce qui suit.

1.3.2 Complétude et compréhension

Déjà dans [Hintikka 1969b], l'auteur déclarait que la majeure partie des critiques dont sa logique épistémique est la cible repose sur une définition axiomatique des théorèmes : leur aspect contre-intuitif exige le recours à une interprétation des formules dans un modèle, une sémantique qui représente la façon dont les modalités s'interprètent dans la 'réalité' de ce modèle. Il s'explique ainsi :

> (...) comprendre un langage logique ce n'est pas, et ne peut pas être, simplement maîtriser son comportement syntaxique et déductif. Il nous faut savoir également comment ce langage et sa logique sont reliés avec la réalité. En d'autres termes, il nous faut connaître sa théorie des modèles.[57]

Autrement dit, l'intérêt d'une sémantique relationnelle n'est pas simplement d'ordre métalogique : quelle que soit la classe de structures susceptible de satisfaire le méta-théorème de complétude, ce système est supposé donner une description sémantique *réaliste* du lien entre langage et réalité ; mais une telle relation des mots aux choses ou des énoncés aux faits ne signifie pas simplement la coordination algébrique d'un domaine à un co-domaine. Nous verrons qu'une sémantique, si elle ne corrige pas l'aspect contre-intuitif de certains axiomes épistémiques, montre tout au moins quelles conditions l'agent épistémique devrait satisfaire pour que ces théorèmes deviennent intuitivement valides.

Et cependant, cet avantage de la sémantique suppose que la *signification* de la connaissance et de la croyance en termes de *conditions de vérité* dans des modèles nous les fasse mieux *comprendre*. Suffit-il d'établir des règles de conditions de vérité pour comprendre une phrase ? En termes formels, l'adéquation formelle de la logique modale épistémique indique qu'à chaque axiome des systèmes de Lewis utilisés correspond une propriété spécifique de R : tous les théorèmes **T**-valides sont valides dans les classes de modèles réflexifs, tous les théorèmes **S4**-valides dans les classes de modèles transitifs, etc.

Hormis la rigueur de la définition sémantique, dans quelle mesure cette correspondance syntaxe-sémantique nous éclaire-t-elle sur le lien dont parle Hintikka entre langage et réalité ? Si des conditions de vérité sont définies en termes d'application ou de fonctions intensionnelles, en quoi cette procédure apporte-t-elle une meilleure compréhension des concepts d'attitude propositionnelle ? Nous verrons que Hintikka procède selon une sorte d'*exégèse* pour montrer l'intérêt de la logique philosophique : définir l'opérateur de connaissance en termes de ses conséquences logiques, par exemple, cela revient à représenter

[57] [Hintikka 1997b], p. 464.

sa signification complète dans une classe de modèles et à indiquer les propriétés algébriques qu'il doit satisfaire pour correspondre à la notion de connaissance. Une fois que le logicien a décrit l'éventail des propriétés disponibles dans un modèle, le travail du philosophe consiste à sélectionner celles qui seront en mesure de caractériser le concept de connaissance.

Mais un esprit critique pourrait reprocher au formalisme de trop insister sur la *clarté* et d'oublier le critère de l'*évidence* : qu'est-ce qui garantit la justesse des conditions de vérité décrétées par le logicien, à la base de sa sémantique ? Peut-on montrer qu'une loi d'application entre une formule et des valeurs de vérité dans un modèle rend bien compte de ce que signifie un concept, à supposer que ce concept puisse être formalisé à juste titre comme un opérateur propositionnel ? En vertu de l'approche vériconditionnelle de la sémantique, toute phrase pourvue de sens peut être expliquée en termes de ses conditions de vérité : comprendre cette phrase, c'est connaître ses conditions de vérité, et la tâche revient à la sémantique formelle d'établir ces conditions. En réaction contre cette approche, [Dummett 1973] reproche à cette sémantique formelle d'être trop éloignée du contexte d'usage des concepts, qu'elle traite comme des opérateurs, et de traiter les notions centrales de vérité et de fausseté comme de simples marque-places algébriques. L'intérêt d'une logique proprement philosophique devrait en être affecté, si le lien prétendu entre langage et réalité se réduit à un jeu de valuations arbitraire :

> Quelle est alors, de ce point de vue, l'importance ou l'intérêt d'une preuve de correction ou de complétude ? Un grand nombre de philosophes esquiveraient la question en disant que ces preuves on un intérêt 'simplement technique' (...) Cela reviendrait à réduire la notion sémantique de conséquence logique à un outil purement algébrique.[58]

Mais nous attendons plus d'une sémantique formelle qu'un résultat technique ; nous attendons d'elle qu'elle apporte un éclairage sur des concepts ordinaires ambigus. De ce point de vue, l'exercice de la sémantique formelle doit se prévaloir d'une moindre intuitivité dans les règles de calcul qu'elle établit.

Le problème posé par Dummett est celui de savoir si une sémantique sans compréhension à l'appui permet d'exprimer un véritable lien entre le langage et la réalité. Dans une sémantique formelle, la signification complexe des énoncés du langage naturel est résumée en quelques règles simples de conditions de vérité. Par un procédé de reconstruction qui va des parties simples à la structure complexe d'ensemble, les énoncés complexes du langage ordinaire peuvent être reconstruits ensuite en termes sémantiques clairs et récursifs. Cette analyse est-elle suffisante ? Dummett prétend qu'une analyse en termes de conditions de vérité ne suffit pas pour donner la signification d'un énoncé. Il reproche notamment aux sémanticiens formels de confondre *garantie* et *justification* : les règles de calcul permettent de

[58] [Dummett 1973b], p. 204. L'auteur reproche à la logique formelle de rechercher des correspondances entre syntaxe et sémantique abstraites des expériences de langage. Or la pratique des logiciens philosophiques semble intimement liée à la pratique constructive décrite ici par Meixner : "La logique moderne procède d'une manière assez proche de la maxime : 'Donnez-moi le système axiomatique, *quel qu'il soit*, et je vous donne ses modèles (s'il y a). Ou vice versa : 'Donnez-moi les modèles, *quels qu'ils soient*, et je vous donne le système axiomatique qui leur convient (s'il y a)." Mais le problème soulevé par la logique philosophique reste ouvert : quelles sont les règles d'interprétation adéquates des attitudes épistémiques ?

garantir abstraitement la validité d'un théorème, par application d'une méthode de décision ; mais elles ne *justifient* pas par là même la conformité de ses résultats à l'usage du concept formalisé, et l'existence de 'paradoxes' en logique philosophique (dont ceux de l'implication matérielle) semble montrer en effet les limites d'une analyse vériconditionnelle des langues naturelles. Le traitement en termes de conditions de vérité doit-il être abandonné pour cette raison ? Dummett semble émettre des réserves quant à la valeur ajoutée d'une analyse logique du langage ordinaire. Pour justifier une sémantique des attitudes propositionnelles, il nous faudra préciser la façon dont sont liées les notions de signification et de compréhension, et nous nous référerons pour ce faire à l'auteur de prédilection de Dummett : Frege. Sans quelques précaution d'usage concernant le processus de signification, les notions sémantiques centrales de vérité et de fausseté seraient réduites à de simples marques symboliques comparables aux valeurs 1 et 0 des algébristes. Comment éviter cette réduction de la sémantique formelle à une manipulation de *signes* vides de sens, et que peut nous apprendre l'analyse logique que nous ne sachions déjà au sujet des concepts philosophiques ? Répondre à cette question, c'est montrer que l'adéquation formelle (correction et complétude) d'un système logique n'est pas seulement un résultat technique et peut aboutir à une adéquation informelle de son langage formel. Il s'agit de montrer que les notions formelles sont les *symboles* d'usages réglés de notre discours sur les choses.

Hintikka a déjà justifié en partie l'adéquation informelle de sa sémantique relationnelle : lorsqu'il compare la valuation à la représentation d'un état de choses possible, le modèle à un ensemble d'états de ce genre et la relation R à une description des alternatives possibles, le vrai et le faux sont conformes à notre usage de ces termes et n'agissent pas comme des signes non interprétés : la vérité d'une formule dans un modèle exprime bel et bien un rapport au monde, et les propriétés de R imposées par les opérateurs modaux déterminent bel et bien l'ensemble des éventualités que nous pouvons admettre dans ce monde selon la modalité en question (aléthique, épistémique, temporelle, etc.). En ce sens, une sémantique relationnelle est préférable à la sémantique algébrique : toute valeur autre que le vrai et le faux n'a pas d'interprétation précise, hormis peut-être la valeur indéterminée ½, et même cette dernière valeur peut être redéfinie en termes d'alternative ouverte dans un cadre bivalent. Les valeurs de vérité utilisées dans [Hintikka 1962] échappent à cet égard à la critique relativiste de Dummett : les *valeurs logiques* (le vrai et le faux) ne sont pas de pures *valeurs algébriques* et retranscrivent de façon justifiable notre rapport indéterminé vis-à-vis du monde ; lorsque l'attitude d'un sujet est indéterminée vis-à-vis d'une formule ϕ, l'assignation de la valeur ½ à cet énoncé correspond en sémantique relationnelle à l'application de l'opérateur modal faible $\Diamond$ à ϕ et exige la prise en compte d'autres mondes possibles. La pluralité des ensembles modèles restitue cette idée d'indétermination et fait de la logique épistémique un langage formel capable de représenter une théorie de la décision.[59]

L'intérêt de la complétude pour la logique philosophique repose donc sur une exigence d'adéquation informelle minimale. Ce réquisit minimal paraît satisfait dans la sémantique relationnelle de Hintikka, mais il reste à s'entendre sur le choix des propriétés relationnelles

[59] La distinction entre valeurs *algébriques* et valeurs *logiques* est empruntée à [Suszko 1975], selon lequel toute matrice multivalente peut être réduite à une matrice bivalente composée de deux groupes de valeurs : les valeurs *désignées* et *non-désignées*. Nous reviendrons sur ce point dans le troisième chapitre.

de R. Comment interpréter la vérité formelle des axiomes de **K**, **T** ou **S4** ? L'intérêt d'une sémantique relationnelle, on l'a vu, est de définir récursivement une fonction de valuation v en vertu de laquelle un ensemble de formules sera validé dans un ou plusieurs modèles. Si l'une de ces validités formelles s'avère être contre-intuitive, l'avantage de la sémantique formelle sera de mettre en évidence le lieu de l'opération au niveau duquel la formalisation est susceptible de s'écarter de l'intuition. Nous reviendrons sur les détails de cette méthode d'analyse philosophique formelle, dans la section consacrée à l'idée de logique comme *modèle explicatif* (section 2.2.3).

Mais d'autres ambiguïtés sont encore à dissiper avant cela, parmi lesquelles le processus de quantification sur des mondes et son impact sur la notion d'existence. Quantifier sur des mondes possibles, n'est-ce pas admettre des entités irréelles, les *possibilia* et, si tel est le cas, la logique épistémique nous engage-t-elle irrémédiablement dans un discours métaphysique ?

1.3.3 Déflation ontologique

La sémantique détermine les conditions de *vérité* des formules dans un langage formel. Mais quelles les conditions de *possibilité de vérité* d'une formule ? De prime abord, le logicien philosophique peut considérer comme tout bon logicien son langage formel comme un simple jeu de *structures* sans prétention ontologique, exprimant des relations entre des objets quelconques : si, comme on l'a dit plus tôt, la logique concerne la forme et ne s'intéresse pas au contenu, on peut en conclure que le tiers exclu ($\phi \vee \neg\phi$) ou la non-contradiction $\neg(\phi \wedge \neg\phi)$ sont validés uniquement en raison de la signification de leurs constantes logiques. Dans la tradition des logiques abstraites, une logique se caractérise ainsi comme une paire $\langle \mathcal{L}_f, \Vdash \rangle$, où $\mathcal{L}_f$ désigne un langage formel doté de règles de formation de formules bien formées et $\Vdash$ une relation de conséquence (syntaxique ou sémantique) entre ces formules. Rien jusqu'ici ne définit la logique autrement que comme une combinatoire de signes, un pur formalisme. Mais l'introduction d'un modèle crée un problème d'ontologie au sein de la logique : quelle est la portée ontologique d'un énoncé quantifié, par exemple ? Lorsque le philosophe interprète l'interprétation des formules logiques dans un modèle, il traite le modèle comme un reflet des états de choses du monde et des objets qui le constituent. Faut-il justifier l'existence des individus dans un modèle, ou l'assignation d'individus à des variables n'est-elle qu'un acte arbitraire sans connexion avec le 'réel', avec ce que Russell appelait l''ameublement du monde' ? Un usage purement abstrait de la logique ne susciterait pas ce genre d'interprétation ; mais un usage philosophique crée apparemment ce genre de problèmes : il faut justifier la nature des entités introduites dans le modèle. De quelle façon cet engagement ontologique affecte-t-il la logique épistémique ? Le problème concerne le statut ontologique des individus dans les mondes possibles, et nous allons observer la réponse de Hintikka à ce problème d'interprétation.

D'un côté, parler de vérité logique revient à ne rien dire *sur* le monde. C'est le sens que [Wittgenstein 1969] donnait au caractère vide de sens [*sinnlos*] des tautologies. C'est le sens qui semble se dégager également de la vérité logique telle que Quine la définit :

> Trivialement, les vérités de logique sont vraies en vertu de n'importe quelle circonstance qu'il

vous plaise de nommer : le langage, le monde, n'importe quoi.[60]

Mais d'un autre côté, ne rien dire ne signifie pas tout à fait que la vérité logique ne porte *sur* rien. Deux problèmes se posent, concernant les conditions de vérité d'une formule : dans un langage logique, la notion de conséquence décrit les conditions sous lesquelles la vérité d'une formule est préservée en vertu de règles de transformation ; la simple notion de vérité est donc présupposée dans cette procédure d'interprétation, laquelle suppose que les prémisses soient pourvues de sens et puissent recevoir une valeur sémantique : ψ est la conséquence logique de ϕ si la vérité de ϕ implique la vérité de ψ. Or pour qu'une formule reçoive une telle valeur, ne faut-il pas que son terme singulier *existe*, c'est-à-dire appartienne au modèle d'interprétation du langage à titre de condition préalable ? Le problème est connu, qui concerne les termes singuliers sans référent et dont l'explication exige que l'on passe d'une interprétation d'ordre zéro (le calcul des propositions) à une interprétation du premier ordre (calcul des prédicats).

Dans une interprétation référentielle de la signification, une formule est interprétée comme vraie si l'énoncé qu'elle représente contient une extension pour chacun de ses termes. Si l'énoncé F*a* est faux, l'interprétation signifie qu'il existe dans le modèle correspondant une interprétation de la constante d'individu *a* et que cette interprétation satisfait la propriété d'être F ; si en revanche F*a* est faux, cela veut toujours dire que *a* existe (que la constante d'individu *a* reçoit une interprétation dans le modèle) mais ne satisfait pas la propriété d'être F. Dans les deux interprétations envisagées, le terme singulier de l'énoncé est supposé exister. Mais que dire dans ce cas d'un énoncé tel que 'Pégase est identique à Pégase' : exprime-t-il une instance du principe d'identité à soi ou, au contraire, l'identité formelle entre les deux formules 'x = x' et 'Pégase = Pégase' est-elle insuffisante pour garantir la *validité* de l'énoncé 'Pégase = Pégase' ? Peut-on dire d'un objet qu'il est identique à lui-même lorsqu'il n'existe pas, et cela a-t-il un sens de parler d'objets identiques dans un modèle vide ? Si le critère d'isomorphie entre les deux formules précédentes ne suffit pas, cela semblerait montrer qu'une loi logique présuppose un engagement ontologique minimal. Dans la logique des prédicats du premier ordre, les lois logiques expriment des termes d'objets qui appartiennent à un modèle. Dans le cas de la généralisation existentielle, par exemple : $Fa \rightarrow (\exists x)Fx$, ce principe est logiquement vrai parce qu'il est vrai pour tout *a* appartenant à un modèle quelconque, que ce soit le modèle des nombres réels, des habitants de Paris, des noms de pays commençant par la lettre 'P', etc. Peu importe donc le modèle et la valeur de *x* choisis, mais cet individu doit exister, c'est-à-dire appartenir à l'un de ces modèles pris au hasard.

Le fait que la notion de vérité logique contienne le mot 'vrai' indique un engagement ontologique minimal. Est-ce nécessaire ? Toute valuation présuppose-t-elle une interprétation de l'opération (une relation langage-monde sous-tend-elle chaque assignation de valeurs) ? On trouve à ce sujet une opposition entre les théories de la vérité de Frege et Wittgenstein (voir chapitre 2), l'un exigeant l'existence du terme singulier pour prédiquer la vérité ou la fausseté de lui, l'autre affirmant au contraire qu'un énoncé logiquement vrai ne porte sur rien. La difficulté porte sur le sens du mot 'rien' : signifie-t-il que la vérité logique ne porte sur 'aucun' objet du modèle correspondant, ou qu'il porte sur 'n'importe quel' objet de ce

[60] [Quine 1975], p. 96.

modèle ?

Les considérations qui suivent n'ont d'intérêt que pour un logicien admettant un engagement ontologique minimal lors de l'interprétation d'un énoncé ; pour un logicien dont l'interprétation dans un modèle est toute formelle et n'exprime aucun lien véritable avec le monde réel, la question de la déflation ontologique posée ici ne sera que pure spéculation ou, pire, l'effet d'une confusion entre *supposition de vérité* et *engagement sur la vérité*. C'est notamment le cas des logiciens libres et des adeptes de la quantification substitutionnelle (opposée à sa version objectuelle) : pour ces derniers, une expression peut avoir un sens sans que ses termes portent un référent, parmi lesquels les énoncés tels que 'Pégase est Pégase'. Nous suivrons ici la voie engagée de Quine et des empiristes logiques, pour lesquels la logique ne parle pas de n'importe quoi mais a vocation à parler de la réalité. La thèse de Quine ('pas d'entité sans identité') restreint le droit de parler d'individus en fonction de notre connaissance scientifique de leurs propriétés distinctes et de leurs relations avec d'autres individus, sans quoi tout devient possible et le domaine de référence arbitraire. Le lien entre langage et monde justifie ce souci ontologique de départ, qui s'exprime en logique des prédicats du premier ordre standard par la règle de généralisation existentielle : $Fa \Rightarrow (\exists x)Fx$. Affirmer quelque chose de vrai d'un objet, c'est supposer son existence dans notre ontologie, que l'on traduit par un modèle ou domaine de discours ; dans le cas contraire, l'énoncé sera soit ni vrai ni faux et dépourvu de sens (théorie de la présupposition d'existence de Strawson), soit toujours faux (théorie de la description définie de Russell). Or si la vérité d'un énoncé signifie l'application d'une propriété à un objet *existant*, comment affirmer quelque chose de vrai en logique modale sans présupposer l'existence de possibles inactualisés, c'est-à-dire l'existence d'objets dans d'autres modèles qui expriment des mondes possibles ?

Le critère d'engagement ontologique de Quine ('être, c'être la valeur d'une variable liée')[61] s'inscrit également en faux contre une version formaliste de la pratique logique : la doctrine linguistique de la vérité logique de Carnap. Selon cette doctrine, la vérité logique repose sur les règles de significations accordées aux connecteurs dans des tables de vérité, mais ces significations sont choisies par convention au sein d'un langage formel arbitrairement constitué. Contre cette restriction de la vérité logique à une manipulation de *signes* plutôt que de *symboles* (dont le signe porterait un référent), Quine maintient l'idée que la notion de vérité logique présuppose celle de vérité et que celle-ci suppose à son tour un lien entre le langage et le monde. [Gochet 1978] explique ainsi que la définition tarskienne de la vérité comme satisfaction d'une formule par un modèle ne se réduit pas à un jeu d'assignations de valeurs *arbitraires* :

> Notre raisonnement peut se résumer ainsi : la vérité de logique 'p ou non p' est fabriquée à l'aide de deux matériaux qui sont respectivement le *sens* de 'non' et de 'ou', et de la *variable d'énoncé* 'p'. Or la définition de l'*énoncé* mentionne le mot 'vrai'. (...) Le concept de *vérité logique* présuppose donc le concept de *vérité* tout court. En d'autres termes, pour qu'il y ait de la vérité de logique n'importe où, il faut qu'il y ait de la vérité *tout court* quelque part. Or, la vérité se définit en termes de 'satisfaction par des séquences d'individus'. Dès lors, la vérité tout court, et du même coup la vérité de logique s'enracine finalement dans la référence, et plus

[61] Cette maxime apparaît pour la première fois dans un article de 1939 : "Designation and Existence", et sera reprise dans [Quine 1961a]. Elle signifie que le verbe 'être' n'a de sens que lorsqu'il s'applique à des termes désignatifs, c'est-à-dire désignant un objet dans un modèle.

particulièrement dans un domaine d'individus. Il y a des vérités de logique qui ont manifestement une portée ontologique, c'est le cas de l'énoncé $(\forall x)(x = x)$ qui est tout autant une vérité d'ontologie qu'une vérité de logique.[62]

Si l'on prend l'expression 'quelque part' comme signifiant 'quelque part dans le modèle', et si les éléments du modèle ne sont pas choisis arbitrairement en logique philosophique, alors l'énoncé 'Pégase est identique à Pégase' sera une vérité logique pour qui s'engage sur l'existence d'entités mythologiques, mais ce ne sera pas une instance de vérité logique pour un empiriste radical tel que Quine.

Par voie de conséquence, l'idée d'un engagement ontologique minimal en logique ajoute un critère de *signifiance* que n'exprime pas à lui seul le processus de valuation $\mathcal{L}_f \rightarrow \{0,1\}$ des logiciens abstraits. Certes, un énoncé logiquement vrai ne dépend pas d'un choix d'interprétation particulier de ses termes, par opposition aux vérités matérielles : le propre d'une vérité logique est d'être formelle ou de reposer essentiellement sur la forme de l'enchaînement des énoncés, comme le montre un syllogisme valide (donc formellement vrai) dont les prémisses peuvent être matériellement fausses.[63] De plus, certaines prémisses matériellement fausses sont susceptibles de devenir matériellement vraies tant que leurs termes singuliers et généraux ont un sens pour nous, c'est-à-dire tant qu'ils appartiennent à l'ontologie que nous associons au langage interprété. Et si tel est le cas, la logique ne se réduit pas uniquement à un jeu de structures défini par des règles de déduction naturelle, une axiomatique ou un calcul des séquents. Elle exige une interprétation, un modèle dans lequel les expressions du langage prennent leur sens.

Les relations de conséquence logique se produisent ainsi entre des formules susceptibles d'être vraies ou fausses, c'est-à-dire interprétables. Pour s'assurer des conditions de possibilité de vérité de ces formules, le logicien doit donc s'assurer de deux choses préalables. Premièrement, que les objets exprimés dans ces formules trouvent une place dans son modèle, quelle que soit les catégories d'éléments que l'on puisse y trouver ; parler de 'notre' ontologie comme d'une ontologie commune et admise par n'importe quel locuteur était peut-être présomptueux, mais l'idée de relativité des ontologies (admise par Quine) n'est pas incompatible avec un engagement ontologique minimal de la part du logicien philosophe ; au contraire, cette idée affirme seulement qu'une assomption préalable d'existence de l'objet dont on parle est nécessaire pour donner sens à son énoncé. Deuxièmement, que le logicien puisse affirmer que l'objet dont il parle existe en un certain sens : le nombre des objets existants varie entre des logiciens philosophes dont les affinités seraient tantôt nominalistes tantôt réalistes ; mais lorsque l'assomption d'existence s'étend au domaine des énoncés modaux, il s'agit de donner désormais un sens à l'idée d'une vérité nécessaire ou simplement

[62] [Gochet 1978], pp. 152-3. L'auteur use également d'une métaphore pour marquer la différence entre être vrai de n'importe quoi et de rien : "De ce que 'p ou non-p' est *vrai quelle que soit* la valeur de p, on ne peut pas conclure que 'p ou non-p' soit vrai *indépendamment* de la valeur de p, pas plus qu'on ne peut dire que si un moteur consomme *indifféremment* n'importe quel combustible, il fonctionne *indépendamment* de tout combustible" (p. 152).

[63] Le syllogisme qui suit, bien que contenant deux prémisses matériellement fausses, est formellement vrai. 'Les triangles sont des carrés, et les cylindres sont des triangles ; donc tout cylindre est un triangle', qui correspond à la 1ère figure du première mode (Barbara : tout B est A, et tout C est B; donc tout C est A). Le problème ne se pose pas lorsque les prémisses sont matériellement fausses, mais lorsque les termes des prémisses désignent des termes vides : 'les dragons', 'les chimères', etc.

possible. Comment justifier une ontologie en prétendant que certains de ses objets ont certaines de leurs propriétés nécessairement, par exemple ? C'est à cette difficulté que le logicien épistémique doit répondre, s'il veut satisfaire des besoins d'ordre philosophique avec son langage formel. L'avantage technique de la sémantique relationnelle ne suffit pas à éliminer par ailleurs cet inconvénient ontologique : justifier le *mode* d'existence de certains objets dans un modèle, et cela avant même de s'intéresser en logique aux relations déductives entre des formules modales dont le sens même est mis en cause.

Une fois admis le critère d'engagement ontologique, le problème que pose la logique épistémique est donc double : non seulement le logicien modal quantifie sur des mondes possibles, en vertu de l'analogie structurelle entre modalités et quantificateurs ; mais lorsqu'il s'agit d'analyser ces procédures au niveau de la logique du premier ordre, on constate que ces mondes possibles contiennent souvent des individus qui n'appartiennent pas en tant que tels à l'ontologie déclarée du monde actuel. En effet, le dernier changement apporté aux mondes possibles par [Kripke 1963] est la variation des domaines : soit un domaine est constant d'un monde à un autre, et seules les propriétés de la classe constante des individus changent parmi certains de ces mondes ; soit un domaine est inconstant, et certains individus peuvent apparaître ou disparaître d'un monde à un autre ; soit l'on combine les deux scénarios précédents, et le nombre des individus dans un monde peut changer aussi bien que leurs propriétés. Une formule modale n'aura pas la même signification selon ces trois cas de figure, notamment la Formule de Barcan : $\Box(\forall x)Fx \rightarrow (\forall x)\Box Fx$, validée ou non selon le type de modèle (à domaines constants ou non constants). Non seulement l'identité des individus possible pose un problème pour la théorie référentielle de la signification, apparemment, mais il s'y ajoute une relativité des domaines qu'il reste à expliquer.

Nous n'entrerons pas plus en détail sur la thèse de la relativité de l'ontologie de Quine : la logique ne statue pas ce qu'il doit y avoir, mais seulement ce qu'une théorie *dit* qu'il y a. Or malgré le droit théorique de postuler des entités fictives ou abstraites, Quine reprocherait au logicien modal de reproduire les difficultés de ses interlocuteurs de [Quine 1961a], Wyman et McX : grossir l'ontologie (l'univers de discours) d'un nouvel 'être' chaque fois qu'un énoncé pourvu de sens porte sur quelque chose qui n'existe pas 'réellement' mais 'aurait pu' exister. D'où le risque de retomber dans la jungle ontologique décriée par Quine, lorsque le logicien épistémique introduit des mondes possibles sous prétexte de traduire efficacement les intensions en termes d'extensions multiples.

Ne faut-il pas admettre en logique modale interprétée l'existence d'entités abstraites, et comment s'engager sur des individus possibles (des *possibilia*) dont les propriétés sont inconnaissables ? Le problème ontologique posé à la fois par McX (partisan d'un premier conceptualisme, celui des idées mentales) et Wyman (partisan d'un second conceptualisme, celui des individus possibles inactualisés) semble devoir se poser également pour Hintikka : imaginer des mondes possibles entraînerait la postulation d'individus que nos moyens de connaissance ne sont pas capables d'identifier ou de différencier les uns des autres. L'exemple donné par [Quine 1961a] du gros homme possible dans l'embrasure de la porte est connu (voir section 2.3.2) : comment affirmer quelque chose de vrai (que l'on peut qualifier de 'proposition' mais que nous appellerons plus modestement 'énoncé') au sujet d'individus

pour lesquels rien n'est exclu *a priori*, et que faire des objets possibles contradictoires (un rond carré) ou inactualisés (Pégase) ? En d'autres termes, le discours des mondes possibles crée des difficultés d'interprétation et semble nécessiter le retour à un discours scolastique, relatif à l'existence d'anges ou de griffons. La position empiriste et naturaliste de Quine le conduit tout naturellement à rejeter par l'absurde une telle sémantique des mondes possibles : parler de tout en logique ne veut pas dire que l'on parle de n'importe quoi. Et cependant, la logique épistémique de Hintikka assume à la fois l'engagement ontologique de Quine *et* l'introduction d'individus possibles.

Le réponse de Hintikka se situe dans le cadre d'une ontologie générale, dont on peut trouver deux versions parallèles dans [Quine 1961a] : à l'ontologie des universaux (philosophie des universaux) et des nombres (philosophie des mathématiques) s'ajoute le problème de l'ontologie des mondes possibles (philosophie de la logique modale), provoqué par l'intrusion des mondes possibles :

Ontologie des universaux : réalisme/ conceptualisme/ nominalisme
Ontologie des nombres : platonisme/ intuitionnisme/ formalisme
Ontologie des mondes possibles: réalisme modal/ conceptualisme modal/ nominalisme modal

Au premier réalisme, selon lequel les idées ont une existence indépendante de l'esprit, [Nef 1991] associe un premier type d'engagement ontologique (défendu par David Lewis), c'est-à-dire l'idée selon laquelle "les mondes possibles existent *per se* ; il y a une réalité modale ; la modalité est objective ; il y a donc des modalités *de re*." Pour le second type d'engagement, celui du conceptualisme modal (Kripke, Stalnaker), "les modalités traduisent le rapport de notre esprit à la réalité ; les mondes possibles sont des concepts objectifs, mais n'existent pas indépendamment de notre esprit ; il n'y a pas de réalité modale". Quant au nominalisme modal, il s'agit d'un désengagement :

> les modalités ne sont que des tournures linguistiques. Ce sont des outils, mais qui ne renvoient pas nécessairement à des concepts objectifs de modalité. Le concept de modalité est obscur, et il faut en limiter l'usage. Il faut se limiter autant que possible aux modalités *de dicto*.[64]

Hintikka se réclame de cette dernière position, d'après laquelle aucun monde possible n'existe indépendamment : celui-ci n'est rien d'autre qu'un artefact utilisé pour exprimer des conjectures. La description précédente de [Hintikka 1962] a montré en effet qu'un monde possible n'est rien d'autre qu'un ensemble de formules, et que chacune de ces formules représente un énoncé décrivant un état de choses possibles. Et bien que l'auteur exige la construction d'un modèle pour expliquer la manière dont le langage nous relie à la réalité, cette liaison n'implique nullement une inflation ontologique car elle porte sur un seul et même monde. Le monde possible est un outil conceptuel destiné à exprimer des alternatives futures ou à définir un espace d'action cohérent, rien de plus ;[65] le logicien épistémique n'aurait donc pas de raison de nourrir quelque scrupule ontologique à la Quine :

---

[64] [Nef 1991], p. 195.

[65] La position de Hintikka sur le statut ontologique des objets irréels est cependant ambiguë : non seulement l'auteur semble admettre l'existence des entités fictives dans [Hintikka 1984] ; mais le lien posé dans [Hintikka

> Il est évident que cette notion de 'monde possible' est une notion inoffensive ; chacun de nous fait constamment des prévisions en vue d'une manière possible dont les choses pourraient s'avérer, pour autant qu'il se trouve savoir.[66]

Le meilleur moyen d'éviter ces malentendus en logique modale, selon Hintikka, serait de rayer la notion de 'monde possible' de son vocabulaire : celle-ci a donné lieu à trop de spéculations fantasques et de débats stériles pour avoir un véritable intérêt. Les questions relatives à l'eccéité des individus (ce qui fait qu'ils sont ce qu'ils sont) ou à la distinction entre existence et actualité, existence et subsistance etc., appartiennent à la métaphysique et, bien qu'elles soient traitées par d'autres logiciens modaux tels que Alvin Plantinga ou David Lewis (théorie des contreparties), elles ne concernent en rien la logique modale épistémique. On peut parler de *déflation ontologique* pour exprimer le traitement nominaliste des modalités. Pour conforter l'austérité ontologique de Quine, Hintikka rappelle que la sémantique des mondes possibles décrit un processus très courant, celui de la *délibération*. Avant décision, l'agent imagine plusieurs cas de figure et cherche à éliminer le maximum d'alternatives en fonction de ce qu'il sait déjà. La sémantique 'relationnelle' apporte sa contribution logique à cette méthode de délibération courante, sans qu'aucune rapport ne soit fait ou à faire avec les problèmes de terminologie des métaphysiciens :

> Afin de parler de ce qu'une certaine personne *a* sait et ne sait pas, nous devons supposer une classe ('espace') de possibilités. Ces possibilités seront appelées des scénarios. Les philosophes les appellent typiquement des mondes possibles. Cet usage est un symptôme de mégalomanes intellectuels. Dans la plupart des applications, les 'mondes possibles' ne sont pas littéralement des mondes au sens d'univers, mais simplement de 'petits mondes', c'est-à-dire autant d'applications du langage en question, typiquement des applications à un recoin relativement petit de notre monde quadri-dimensionnel. Cet espace de scénarios est pour l'essentiel le même que ce que les théoriciens de la probabilité entendent par espace échantillon. On pourrait l'appeler un espace épistémique.[67]

Le problème central de la quantification modale repose en somme sur cette question : quel est le lien entre une modalité et la question de l'existence ?

Tout d'abord, Hintikka réussit à dissocier partiellement engagement ontologique et logique modale en rappelant une distinction cruciale, que nous reprendrons à notre compte dans la suite. La forme logique des énoncés modaux correspond à la tripartition grammaticale entre trois aspects des énoncés du langage ordinaire : les aspects *phrastique*, *neustique* et *tropique* :

> Dans l'usage d'un grand nombre de concepts importants, on peut distinguer ce qui pourrait être appelé la composante descriptive de la composante non-descriptive ou modale. Cette distinction

---

1969b] sur le lien entre l'interprétation dans un modèle et le néo-kantisme le font incliner vers une explication conceptualiste des objets 'inexistants'. Ce problème porte sur les conditions transcendantales du processus de signification, c'est-à-dire sur les facultés dont doit disposer un agent pour assigner un objet à un concept.

[66] [Hintikka 1962], pp. 138-139.

[67] [Hintikka 2003], p. 34. Dans la lignée de [Kripke 1963], les mondes possibles de Hintikka correspondent à des descriptions *partielles* du monde et n'apportent pas une description du monde aussi complète que les descriptions d'état de Carnap. [Hintikka 1975] estime que Carnap aurait limité pour cette raison son analyse des modalités à la version **S5** et n'aurait pas eu l'idée de créer des logiques 'ego-centrées' telles que la logique épistémique, limitées à un point de vue partiel sur le monde.

> peut être caractérisée comme suit : très souvent, ce que l'on fait au moyen d'un concept donné peut être analysé en disant que l'on décrit d'abord un état de choses de la conjoncture possible (un scénario possible), ou plusieurs d'entre eux, et que l'on effectue ensuite un commentaire sur lui ou eux. Par exemple, on pourrait exprimer son approbation ou sa désapprobation de l'état de la conjoncture en question, le désignant comme impossible à se réaliser, dire que quelqu'un sait qu'il a lieu, etc. (...) Cela suggère un programme pour l'analyse logique d'un grand nombre de concepts importants.[68]

L'aspect phrastique est la partie commune et minimale de la communication et qui correspond au 'contenu propositionnel' des formules modales ; l'énoncé, c'est-à-dire les phrases déclaratives du langage ordinaire, désignent ce contenu propositionnel sur lequel portent les opérateurs modaux. L'aspect neustique correspond au jugement que le locuteur cherche à émettre à l'égard du contenu propositionnel qu'il énonce ; l'aspect tropique, enfin, exprime la liaison entre les aspects phrastiques et neustiques par un mode ou déclinaison spécifique du verbe : dans la phrase 'il faut que Pierre assiste à l'école', le mode subjonctif du verbe 'aller' indique par exemple que le contenu propositionnel 'Pierre va à l'école' exprime un jugement impératif, et d'autres opérateurs modaux tels que les contrefactuels ou les impératifs ont également leur propre marque tropique. La logique modale formalise ce triple aspect commun aux jugements quotidiens, l'opérateur modal reflétant l'aspect neustique et s'appliquant au contenu propositionnel, qui reflète quant à lui l'aspect phrastique. Seul l'aspect tropique est éliminé par le formalisme modal, puisque la déclinaison des verbes n'a pas plus d'importance pour la signification des formules que leur forme active ou passive. La distance prise par l'analyse logique vis-à-vis de la grammaire se manifeste également pour d'autres types d'énoncés : l'injonction 'Ferme la porte !', bien qu'elle n'ait pas une forme explicitement modale, peut être restituée comme telle en la paraphrasant ainsi : 'Je veux que tu agisses de telle sorte que la porte soit fermée', où un opérateur exprime un ordre et porte sur le contenu propositionnel 'la porte est fermée'. Cette explication des énoncés comme coordinations d'un neustique avec un phrastique n'est pas nouvelle, puisqu'on la trouve déjà chez les stoïciens avec leur notion de *lexis* ; on la retrouvera par la suite chez les logiciens modaux et leur explication de la déclaration comme application d'un mode à un *dictum* ; on l'examinera tout à l'heure également dans la notion de *signe propositionnel* du *Tractatus* de Wittgenstein. Ainsi, les notions de 'lexis', de 'dictum', de 'contenu propositionnel' ou de 'signe propositionnel' renvoient toutes au signe matériel, c'est-à-dire à la marque linguistique dépourvue d'interprétation. L'analyse logique des énoncés ordinaires nous renverra ainsi à une triptyque opérateur/opérande/opération : un opérateur s'applique à un (ou des) opérande(s) et donne comme résultat une opération ; la confusion entre deux acceptions de la notion d''opération' ou de 'fonction' (tantôt le processus général d'application, tantôt un résultat singulier de ce processus) sera selon nous à l'origine d'une confusion courante en logique philosophique. Bien qu'elle remonte bien avant le 20e siècle, ce type d'analyse devra attendre les progrès de la logique mathématique (fonctionnelle) avant de donner naissance aux logiques impératives, interrogatives (érotétiques), épistémiques, déontiques, etc., chacune

[68] Hintikka [1973a], pp. 4. La distinction entre phrastique et neustique est reprise de l'ouvrage de R. M. Hare : *Language of Morals* (1971), dans lequel l'auteur prétend résoudre les questions d'éthique par une analyse logique des énoncés impératifs. Hintikka s'inspirera à plusieurs endroits de cette procédure philosophique pour justifier les vertus de sa logique épistémique.

d'elles exprimant à sa façon un aspect à la fois *locutoire* (phrastique) et *illocutoire* (neustique) du langage. C'est ce dédoublement des aspects du discours qui permet à la logique modale de prendre ses distances vis-à-vis de tout engagement ontologique ferme : donner son avis sur l'état du monde n'implique pas que cet état soit réel, et toute construction de langage basée sur un élément neustique tendra ainsi à décharger l'import existentiel qu'un terme descriptif manifeste dans une approche référentielle de la signification.

La logique des propositions classiques se compose uniquement d'énoncés, c'est-à-dire de phrases déclaratives et, en tant que tel, un énoncé ne semble comporter qu'un aspect phrastique et aucun aspect neustique. Cependant, un énoncé tel que 'Paris est la capitale de la France' (symbolisé par l'atome propositionnel p en logique classique) exprime un jugement de son locuteur lorsqu'il est prononcé dans un discours ; pour cette raison, il peut être paraphrasé sous la forme 'C'est le cas que Paris est la capitale de la France'. Dans cette nouvelle formulation, un opérateur d'*actualité* @ s'applique au contenu propositionnel et signifie que l'énoncé (la complétive reliée à l'opérateur par le pronom 'que') doit être vrai pour satisfaire l'intention du locuteur (décrire un état de choses réel). La forme @p ainsi obtenue est déjà présente implicitement dans la logique classique, avec l'opérateur implicite et redondant de l'affirmation : $\llcorner\phi$.

L'analyse ci-dessus des énoncés ordinaires permet de mettre en parenthèses le problème de l'engagement ontologique en logique modale, mais elle permet également de rappeler que la logique classique constitue déjà une sorte de logique modale interprétée ; les énoncés ne sont qu'un cas particulier des formules dont l'analyse logique modale est en mesure de déterminer la signification. Nous laissons de côté pour l'instant la question du type de conditions exigées par les autres types d'énoncés non-déclaratifs (interrogatifs, promissifs, etc.).

Toutes les phrases interprétées de forme modale doivent-elles être comprises en termes de conditions de *vérité*, ou d'autres conditions de satisfaction entrent-elles en jeu ? Quoi qu'il advienne de cette question, la distinction entre phrastique et neustique, locutoire et illocutoire semble être au cœur de la logique épistémique, tout au moins de sa sémantique.

Forme logique de l'énoncé de croyance

	Je crois	que Paris est la capitale de la France
Aspects :	neustique	phrastique
Type :	principale	subordonnée complétive
Forme logique :	B_{je}	p

Etant donné l'insistance de Hintikka sur la notion de modèle et le rôle que celui-ci doit remplir pour expliquer le lien du locuteur entre son langage et la réalité, nous accorderons à cette lecture une place essentielle dans la seconde partie de ce travail, consacrée à la relation entre la pensée logique et le sujet pensant. Le développement des logiques philosophiques a

montré que, contrairement aux restrictions de la logique traditionnelle, l'analyse logique ne concerne donc pas uniquement les énoncés ou phrases déclaratives mais peut s'appliquer également aux prières, aux injonctions, aux questions, etc.

La forme modale permet de donner une analyse logique plus proche de la communication des langues naturelles : toute situation de communication ne consiste pas uniquement à décrire le monde ou à produire un discours de type scientifique. Pour [Granger 1979], un énoncé n'est pas analysé de manière complète s'il ne tient pas compte de l'aspect neustique ou illocutoire de la communication :

> dans une langue naturelle, un énoncé complet est *déterminé par ses aspects pragmatiques* beaucoup plus que par la forme syntaxique ou le contenu sémantique correspondant à l'expression (...) Or, si l'on abstrait de l'énoncé complet naturel ses éléments pragmatiques, il reste *moins* qu'un énoncé complet logique.[69]

Nous reviendrons ci-dessous sur les raisons qui ont retardé l'émergence de ces logiques du langage naturel ; mais force est de constater que, quelles que soient les raisons des réticences du début du 20e siècle au sujet de la logique modale, le fait que l'aspect phrastique (déclaratif) soit enchâssé dans l'aspect neustique d'une phrase confirme son appellation d'extension conservatrice : cela montre que la logique modale *étend* les formes logiques classiques tout en se construisant *sur* celles de la logique classique, c'est-à-dire les contenus propositionnels de forme prédicative :

> En d'autres termes, cela sous-entend une approche de la logique non-modale ordinaire comme logique de descriptions de mondes, et une approche de la logique modale (au sens large où elle inclut non seulement l'étude des concepts de possibilité et de nécessité, mais aussi l'étude de tous les autres concepts de comportement similaire) comme étude des propriétés et interrelations de différents mondes possibles.[70]

Le problème reste celui de comprendre la façon dont ces différents états sont reliés au monde réel, en logique épistémique ; le problème ontologique concerne les individus et leurs propriétés, cependant, si bien que l'idée de quantifier sur des mondes n'a de sens que métaphorique : c'est sur les individus de plusieurs états de choses possibles du monde réel que nous quantifions. Mais que faire des attitudes portant sur des individus inexistants, que ce soit des personnages fictifs sur lesquels des agents peuvent tenir des croyances ou sur les illusions perceptives (autres cas d'objets intentionnels non reconnus dans le domaine du monde réel) ?

Une autre solution pour concilier modalité et quantification consiste à décharger un quantificateur de son import existentiel : il serait permis de quantifier sur des modèles sans présupposer que les variables liées au quantificateur existent. Compte tenu de notre incapacité à appliquer la définition extensionnelle des formules modales sans quantifier sur une pluralité de modèles, la logique modale exige apparemment un certain relâchement par rapport au critère d'engagement ontologique. Or c'est précisément cette voie que Hintikka va proposer dans son programme de logique modale : un quantificateur indique que la variable liée

[69] [Granger 1979], pp. 82-83.
[70] [Hintikka 1973a], p. 5.

appartient à un modèle, mais cette appartenance n'est plus forcément synonyme d'*existence*. On retrouve ce procédé de déchargement ontologique des quantificateurs dans les logiques dites 'libres' ou sans présupposition existentielle ;[71] cette réinterprétation des quantificateurs sera revue par la suite dans la section consacrée à la logique épistémique quantifiée, mais on constate ici qu'elle permet d'éluder un tant soit peu la critique ensembliste de Quine relative à l'axiome de compréhension. Une fois que l'on admet la définition des modalités en termes de quantification non existentielle sur des modèles, il reste à définir les individus et à spécifier leurs critères d'identité à travers les modèles (problème de la 'référence croisée', ou 'transquantification').

En résumé, la relation entre logique modale et ontologie nécessite à la fois :

- un certain lien entre vérité logique et vérité ontologique, de façon à maintenir un contact entre le langage et la réalité ; ce contact est décrit dans une sémantique des modèles ;
- une réduction nominaliste des mondes à des ensembles de formules, pour couper court aux spéculations sur l'existence prétendue de ces mondes et délier la sémantique de l'ontologie.

Mais les questions d'engagement ontologique ne sont pas réglées une fois pour toutes par cet affaiblissement des quantificateurs. C'est une chose de distinguer ce qui appartient à un modèle et ce qui existe au sens propre, une fois que l'on a admis un relâchement du lien entre sémantique et ontologie ; c'en est une autre de déterminer ce qui appartient ou non à un modèle. Hintikka est encore tenu de respecter un autre critère d'existence posé par Quine, à savoir le critère d'identité : 'pas d'entité sans identité' signifie que, quel que soit le sens du quantificateur, une valeur d'individu appartenant à un modèle doit être d'abord bien définie pour que les formules dans lesquelles il figure aient un sens précis. Or l'avantage de la sémantique relationnelle, son extensionalisme plural, semble poser d'autres difficultés de définition par ailleurs. Qu'est-ce au juste qu'un *modèle*, et quelle distinction y a-t-il entre une proposition de la logique classique (à modèle unique) et une proposition dans un modèle ? Le problème de la distinction entre énoncés et propositions entre en ligne de compte.

1.3.4 Enoncés et propositions

Dans le *Tractatus* de Wittgenstein, il est courant de concevoir le tableau du monde comme un ensemble de faits exprimés par des propositions atomiques et dont la connexion par des relations logiques donne un ensemble d'états de choses combinés. La signification des connecteurs permet de délimiter le possible et l'impossible dans un monde, et leurs règles sont reprises en logique épistémique afin de spécifier les conditions sémantiques d'une croyance ou d'une connaissance vraie.

Or cette notion de proposition est encore vague : quelles règles définissent les relations possibles entre des propositions, et en vertu de quoi détermine-t-on le lien établi entre une proposition du langage et un fait du monde ? Où se situe la proposition au sein d'un espace

[71] Karel Lambert et Jaakko Hintikka sont considérés comme les deux promoteurs de la 'logique libre' ; sur cette logique sans présupposition existentielle, voir notamment [Hintikka 1969a] : "Existential presuppositions and uniqueness presuppositions", pp. 112-147.

épistémique ? La première question porte sur l'explication nominaliste de la possibilité en termes de consistance ; la seconde porte sur l'identité de la proposition en tant que telle.

Malgré l'intérêt que revêt une réduction nominaliste des mondes possibles à des ensembles consistants de formules, [Lewis 1973] reproche à cette réduction de donner une explication *circulaire* des notions sémantiques et ontologiques. On ne peut comprendre la nature d'un monde possible sans s'engager un tant soit peu sur la nature du monde, pense-t-il : s'il dit vrai, alors le nominaliste modal ne peut éviter quelques hypothèses métaphysiques lorsqu'il réduit les mondes possibles à des formules. Expliquer, au sens de Carnap, c'est définir une notion par d'autres termes, et les mondes possibles ont été décrits plus haut en termes d'états de choses consistants. Or Lewis prétend que, contrairement à la définition nominaliste de Hintikka, la position actualiste (du réalisme modal) est non seulement défendable mais offre une explication satisfaisante des mondes possibles. D'où vient la difficulté pour le nominaliste ?

D'après Lewis, la *possibilité* d'un monde ne peut pas être expliquée par la notion logique de consistance de ses propositions sans entraîner une circularité. Il s'explique ainsi :

> Qu'est-ce que la consistance ? Si un énoncé consistant est un énoncé qui *pourrait* être vrai, ou un énoncé qui n'est pas *nécessairement* faux, alors la théorie est circulaire.[72]

L'idée de Lewis est que, si la notion de monde possible se définit en termes de consistance, cette notion de consistance se définit elle-même en termes de possibilité de vérité. Ce qui est consistant *peut* être vrai, et ce qui est *inconsistant* ne peut être vrai ou est *impossible* à rendre vrai. Au final, le monde possible se définit en termes de possibilité, d'où une circularité dans l'explication nominaliste de Hintikka.

A cette objection, le logicien pourrait redéfinir la consistance comme relative à un système formel clos : ce sont des règles sémantiques qui déterminent les conditions dans lesquelles une formule est vraie ou fausse relativement à d'autres, auquel cas on peut se passer des notions modales pour définir la consistance dans un modèle. On peut donc remplacer les notions de possibilité ou d'impossibilité par celles de vérité ou de fausseté *dans un système*, et la circularité de l'explication disparaît. Un monde possible serait donc un ensemble de formules dans lesquelles la clause de non-contradiction est satisfaite. Si ϕ décrit un état de choses dans ce monde, alors il est *impossible* que $\neg\phi$ y figure également ; cette description des mondes possibles en termes métalogiques d'impossibilité de contradiction est dispensable, de plus, si l'on songe au fait que les propriétés d'un monde possible sont décrites dans une théorie des modèles en termes de clauses récursives de vérité. Il n'y a donc pas de cercle vicieux dans l'explication du monde possible.

Contre la définition formaliste de la possibilité comme consistance ou non-contradiction dans un ensemble de mondes, Lewis objecte que celle-ci reste insuffisante pour décrire le contenu d'un monde possible :

---

[72] [Lewis 1973], p. 183. Nous soulignons.

> Si une proposition consistante est une proposition qui s'avère être vraie pour une assignation d'extensions au vocabulaire non-logique, alors la théorie est *incorrecte* : certaines assignations d'extensions sont impossibles, par exemple une assignation qui assigne des extensions croisées aux termes 'cochon' et 'mouton'. Si un énoncé consistant est un énoncé qui s'avère être vrai pour une assignation possible d'extensions, alors la théorie est de nouveau circulaire.[73]

Lewis exige le recours à des nécessités ou impossibilités *conceptuelles*, en plus de l'acception logique de la non-contradiction ; si la possibilité d'un état de choses exige une analyse conceptuelle de ses termes, alors la logique épistémique semble incapable de donner une définition explicite des mondes possibles sans perdre son caractère formel, puisque certaines assignations d'extensions sont impossibles en vertu de règles d'usage qui sortent du cadre de la logique.[74]

Le problème de l'analyse conceptuelle peut être sinon d'ordre *métaphysique* : il concerne alors l'ensemble des qualités qu'un individu peut posséder sans perdre sa nature. L'analyse conceptuelle correspond à ce que Meinong ou Łukasiewicz avaient évoqué par le thème des objets impossibles : un carré rond est un objet impossible, auquel manque l'actualité mais qui peut être conçu soi-disant dans l'esprit. Nous laissons ces questions d'eccéité ou de compatibilité conceptuelle de côté, parce que hors de propos : la logique épistémique concerne l'analyse formelle des espaces épistémiques, mais qu'elle soit insuffisante pour les besoins de Lewis ne signifie pas qu'elle soit incorrecte ou erronée. La logique étudie les conditions de rationalité du discours, et le logicien épistémique étend ces conditions aux attitudes propositionnelles ; mais le critère formel de consistance des propositions est indépendant du critère conceptuel d'incompatibilité des qualités. [Nolte 1986] insiste sur le caractère autonome de la possibilité logique :

> le moyen le plus évident de représenter les règles sémantiques qui constituent ces sens est celui de fonctions de mondes possibles à extensions. En conséquence, l'explication tombe dans un cercle. Il y *a* un cercle, ici, *mais il n'est pas vicieux*. (...) Puisque une explication devrait procéder du plus familier au moins familier, ces règles comprises en pratique et intuitivement constituent le point de départ approprié'. La question 'Pourquoi ne peut-il pas y avoir de mondes dans lesquels des affirmations de forme $p \wedge \neg p$ soient vraies ?' trouve une réponse correcte et éclairante dans l'explication de la sémantique de '$\neg$' et '$\wedge$'. (...) Le mouvement aura donc consisté à passer d'une compréhension de l'exercice des règles sémantiques à une compréhension de l'exercice des mondes possibles, en retour d'une représentation sémantique des règles dont nous sommes partis. Si cette représentation théorique était le seul moyen de comprendre les règles sémantiques, alors le cercle serait vicieux. Mais puisque notre compréhension initiale est *indépendante* de cette compréhension théorique, le cercle n'est pas vicieux, mais *herméneutique*.[75]

---

[73] *Ibid.*

[74] A moins d'étendre la notion de logique à la logique informelle, ou de formaliser (rendre formel) ces impossibilités conceptuelles. Tout en posant des réserves sur le sens de sa 'logique', le passage suivant de Wittgenstein semble prolonger l'impossibilité conceptuelle de Lewis : "Que, par exemple, deux couleurs soient ensemble en un même lieu du champ visuel est impossible, et même *logiquement* impossible, car c'est la structure logique de la couleur qui l'exclut" (*Tractatus*, § 6.3751). Une formalisation permettrait de montrer ou faire ressortir la contradiction logique d'un énoncé que ne montre pas sa langue naturelle d'origine ; la question est de savoir ce qui justifie une telle transformation de l'énoncé initial. La résolution logique des paradoxes est une raison, pour Russell (voir ci-dessus).

[75] [Nolte 1986], pp. 438-9 (nous soulignons).

En somme, la critique que Lewis adresse à la définition nominaliste des mondes possibles n'est pas fausse, mais elle n'est pas non plus accablante contre le logicien épistémique : celui-ci pose des critères de consistance nécessaires mais non suffisants pour définir un état de choses ; quant à l'analyse conceptuelle, elle est réservée à un aspect de la sémantique qui ne concerne pas les règles de sémantique formelle, sauf si cette incompatibilité conceptuelle peut être transcrite dans un jeu d'opération dans des structures linguistiques. La formalisation pourvoira en partie à cette tâche, comme le montreront les paradoxes de la logique épistémique.

Mais même minimale, une analyse de la rationalité en termes d'ensembles consistants de formules reste superficielle si elle ne permet pas de préciser ce qui la distingue d'un ensemble de *propositions*. Comme nous le verrons dans une section ci-dessous consacrée à Frege, la proposition est ce qu'un individu locuteur est supposé 'saisir' pour exprimer sa pensée. Or si la logique épistémique a pour but de décrire le comportement pratique des agents rationnels dans la vie courante, elle doit parler non pas de propositions mais d'*énoncés* comme véhicules premiers de la signification : les agents n'expriment pas leurs croyances ou leurs connaissances par des propositions, mais par l'entremise d'énoncés. Un énoncé est une prédication prononcée à un lieu et un temps particulier, dans une langue naturelle particulière et par un sujet particulier. La proposition correspond à la signification commune de certains énoncés : deux énoncés peuvent exprimer une même proposition, mais il reste à spécifier les conditions d'individuation de cette dernière. Comment établir que deux énoncés correspondent à une seule et même proposition et expriment la même opinion ? Dit autrement, il s'agit de justifier ce qui permet de passer d'un énoncé pour nous à un énoncé en soi, le *Satz an sich* de Bolzano et que Frege présuppose dans son idéographie. La logique épistémique, si elle porte sur les attitudes propositionnelles, doit s'interroger tout au contraire sur le contenu de ces attitudes et ne pas présupposer qu'elles portent sur des propositions préconçues : il s'agit ici du problème de l'opacité référentielle et des attributions de croyance (*belief reports*), dont la logique mathématique de Frege n'avait pas à se soucier mais qui concerne au premier chef le projet de logique épistémique. L'opacité est le symptôme du fait que les mondes possibles dont parle le nominaliste modal ne sont pas composés de propositions, mais d'énoncés ; c'est la capacité à savoir ce qui réunit plusieurs énoncés distincts qui détermine leur sens commun, et ce sens commun correspond à la notion de proposition.

Il faut donc distinguer énoncés et propositions, et cette distinction a un rapport précis avec la constitution des espaces épistémiques, les mondes épistémiquement possibles. Deux raisons nous y conduisent : les attributions d'attitude à d'autres agents sont souvent ambiguës (opaques), elles sont souvent dépendantes du contexte d'énonciation et exigent un examen prudent des informations que possède réellement un agent ; les propositions sont des entités idéales dont les relations de conséquence ne sont parfaitement applicables que par une pensée pure, contrairement à nos pensées quotidiennes dont les ressources sont limitées et souvent incomplètes. Les espaces épistémiques sont décrits sur le modèle initial des espaces logiques, mais la différence entre ce que nous appelons 'pensée pure' et 'pensée en acte' suppose que les premiers espaces ne sont pas de simples reproductions des seconds. En ce sens, la logique

épistémique correspond à une logique standard plus détaillée, tenant compte de l'aspect incomplet des pensées courantes et exprimant cet aspect par un opérateur modal ; l'intervalle existant entre logique classique et logique modale épistémique existe, il reste à mesurer.

L'approche nominaliste des mondes possibles donne une raison supplémentaire de considérer les objets d'attitude comme des énoncés, et non comme des propositions. De prime abord, une attitude propositionnelle se présente comme une relation à deux places entre un sujet et une proposition, et l'on insiste également sur le rôle grammatical du pronom relatif 'que' comme introduisant une proposition distincte de l'énoncé. Or si l'attitude propositionnelle vise des propositions, cela revient à dire que les mondes possibles sont composés de formules qui représentent des propositions, conformément au langage-tableau du *Tractatus*. Mais une telle intrusion des propositions risque de produire une autre circularité dans l'explication des mondes possibles : lorsque les propositions sont conçues comme des *intensions* (c'est-à-dire des fonctions reliant des mondes possibles à des extensions), elles apparaissent comme des règles déterminant la valeur de vérité d'une classe d'énoncés au sein de plusieurs mondes. En d'autres termes, la proposition 'Pierre est chauve' désigne la classe des mondes possibles dans lesquels l'énoncé 'Pierre est chauve' est vrai. Or qu'est-ce qu'un monde possible dans lequel l'énoncé est vrai et assigne la même extension à ses termes ? Si la sémantique relationnelle propose de définir les mondes possibles comme des ensembles de propositions, l'approche intensionnelle des propositions propose de définir celles-ci comme des ensembles de mondes possibles. Il y a donc une circularité apparente entre les notions de proposition et de monde possible : un monde possible se compose d'un ensemble de propositions, et les propositions se composent d'un ensemble de mondes possibles. On peut répondre à ceci que les mondes possibles dont parle Hintikka ne sont pas les 'mêmes' que ceux qui servent à définir la proposition entendue comme une intension (un *Satz an sich*) : les premiers sont des états de choses conçus par un sujet, les seconds sont des états de choses réels. Or si les espaces épistémiques se composent uniquement d'états de choses considérés par l'agent, la notion intensionnelle de proposition dépasse le cadre de notre réflexion et porte sur des extensions préconçues. La proposition n'est pas prédéterminée dans un espace épistémique ; au contraire, c'est un ensemble d'énoncés considérés par différents agents qui définit l'unité de la proposition. Un tel renversement dans l'ordre des définitions entre énoncés et propositions se retrouve dans la théorie de l'apprentissage du langage de Quine, notamment. Elle caractérise une position *pragmatiste* selon laquelle la signification des expressions d'un langage résulte d'une construction théorique et non d'une collection passive d'entités abstraites préétablies. Hintikka lui-même suit cette attitude, opposée à une définition *réaliste* de la proposition :

> Je comprends parfaitement le genre de raisonnement que les logiciens ont en tête lorsqu'ils parlent de raisonnement en termes d''individus arbitraires', mais la notion d'individu de ce genre n'a aucune valeur explicative en soi. Je n'ai jamais vu, entendu, senti, touché ou embrassé d'individu arbitraire. Le jargon des logiciens réifiant ces entités hasardeuses me semble tout à fait arbitraire, en l'occurrence.[76]

[76] [Hintikka 2003], p. 33.

Si l'on rejette l'ontologie des propositions, l'objet d'une attitude de croyance peut être assimilé sinon à un énoncé, du moins à des ensembles d'individus et relations particuliers qui satisfont un acte propositionnel.

La première solution de rechange évite de recourir à une entité abstraite, mais elle rend l'objet de croyance dépendant d'une représentation linguistique qu'est l'énoncé ; comment garantir qu'un énoncé de langue allemande renvoie à la même croyance qu'un énoncé de langue française, par exemple ? La traduction doit permettre de garantir l'identité des croyances de deux locuteurs venant de deux langues sources différentes ; une fois cette traduction réalisée, la décitation des énoncés d'origine doit permettre de formuler des croyances communes dans un langage formel unique : la logique épistémique, par exemple. Nous reviendrons sur ce problème de traduction au troisième chapitre, mais l'idée est la suivante : la logique épistémique représente une théorie formelle du langage, car elle permet de formaliser les conditions en vertu desquelles les agents d'une langue naturelle peuvent communiquer entre eux par le biais des mêmes propositions. La croyance est l'acte intermédiaire par lequel des énoncés sont rattachés à une même proposition ; la place de la croyance s'inscrirait ainsi entre les notions de signification et de vérité, comme nous le rappellerons plus tard (section 2.3).

Une autre solution consiste à rejeter toute entité linguistique et à décrire les objets de croyance comme des *dispositions* à agir : l'agent croit que ϕ s'il est disposé à agir comme si ϕ avait réellement lieu. Cette dernière approche est non-linguistique et déporte son critère de définition sur l'action des agents, plutôt que sur leurs énoncés ; elle distingue la croyance de l'assertion explicite et permet d'éviter notamment le paradoxe de la décitation de Kripke, qui sera revu dans la section consacrée aux attributions d'attitude. Peut-on construire une sémantique formelle sur une théorie de la disposition non-linguistique, ou l'identité des croyances ferait-elle mieux d'être restreinte à une stricte condition d'assertion explicite ? En d'autres termes : doit-on inclure la théorie de la croyance dans une théorie de l'action, et que devient la logique épistémique si l'on suit cette trajectoire ?

La simplicité de la sémantique relationnelle ne représente donc qu'un avantage de surface : la réduction des intensions (les modalités) à des extensions (des ensembles de formules) est une procédure formelle qui ne met pas au clair la nature des objets des attitudes épistémiques : énoncés, propositions, ou autre chose ? Le problème de l'identité des propositions renvoie à la question de l'identité de *signification* des énoncés courants, et la logique modale épistémique invite à remettre en question le statut des atomes propositionnels ainsi qu'à réinvestir la question de la relation entre signification et vérité. L'un des buts de la thèse sera de comprendre ce qui sépare et unit à la fois la logique classique et la logique épistémique. Par quel moyen passe-t-on de l'une à l'autre, et que représente la transition entre un énoncé non-modal et un énoncé modal ? Le second chapitre qui suit va illustrer les relations existant entre les deux types de logique ; le troisième proposera un cadre théorique commun pour montrer que celles-ci représentent deux niveaux d'analyse distincts d'un même type d'énoncés.

Chapitre 2
Les attitudes en logique

Résumé : La logique épistémique porte sur des modalités, des opérateurs exprimant des attitudes épistémiques. Y a-t-il un sens à formaliser des attitudes ? Certains problèmes liés à l'interprétation de ces modalités épistémiques seront examinés. Ils seront regroupés en trois thèmes généraux :

- Logique et pensée : où se situe la logique épistémique entre logique et psychologie ? (section 2.1)
- Logique et usage : comment interpréter ses constantes logiques ? (section 2.2)
- Logique et contexte : ne va-t-elle pas à l'encontre du principe de neutralité contextuelle qui caractérise la vérité logique ? (section 2.3)

Soit [la logique épistémique] n'a rien à voir en particulier avec la connaissance et n'est donc épistémique que par le nom, soit [elle] a un rapport et, en conséquence, [elle] n'est logique que par le nom.
(Max Hocutt)

Bien évidemment, 'il croit', 'il souhaite', et les autres verbes d'attitude propositionnelle ne sont des mots logiques à aucun point de vue.
(W. V. Quine)

2.1 Logique et pensée

Comme nous l'avons vu, deux innovations formelles ont permis la création d'un système de logique épistémique : l'axiomatique modale de C. I. Lewis, d'une part, l'introduction de la sémantique relationnelle, d'autre part. Pourquoi a-t-il fallu attendre la seconde moitié du 20e siècle pour voir l'émergence d'une logique des attitudes propositionnelles ? Bien que la comparaison entre les opérateurs modaux et ces attitudes soit assez récente, des raisons plus profondes pourraient expliquer le fait que la logique épistémique n'ait pas été envisagée par les logisticiens, c'est-à-dire les logiciens mathématiques qui ont relancé la discipline vers la fin du 19e siècle et le début du 20e siècle.

Dans ce qui suit, nous allons tenter d'expliquer les diverses raisons de la réticence, voire du rejet catégorique de quelques logiciens majeurs vis-à-vis de l'idée d'une logique des attitudes. Le but de cette première section sera de montrer que l'argument de l'analogie structurelle entre modalité et quantification ne suffit pas pour légitimer la logique épistémique. La raison véritable de cette réticence à l'égard des modalités épistémiques semble tenir à une conception plus générale de la logique, largement défendue par quelques logiciens majeurs mais qui n'a quasiment plus cours aujourd'hui. Quelle est cette conception, et en quoi explique-t-elle l'apparition tardive des logiques philosophiques ?

2.1.1 Croyances et psychologie

La logique épistémique, c'est-à-dire la logique de la connaissance et de la croyance, semble paradoxale à double titre. D'une part, si les règles d'inférence fournies par la logique classique constituent un cadre pour nos connaissances, n'est-il pas redondant de vouloir la compléter par une logique de la connaissance, dont le but serait de régir notre attitude à l'égard de ce cadre ? D'autre part, cette logique porte notamment sur le concept de *croyance*. Or parler de croyance, c'est faire allusion à des états d'esprit ou des contenus de pensée. Cette dernière expression est ambiguë, partagée entre la pensée comme activité générale et la pensée comme action particulière : elle désigne tantôt les états mentaux du sujet qui se manifestent lors d'un raisonnement quelconque sur des objets, tantôt les objets mêmes de ce raisonnement. La distinction entre les deux acceptions du terme sera précisée plus bas, dans la section consacrée à Frege. La raison majeure pour laquelle une logique épistémique n'aurait pas de sens, selon des auteurs tels que Frege, Wittgenstein ou Russell, c'est qu'elle porte sur une notion essentiellement *psychologique*, et qu'une confusion entre croyance et vérité, psychologie et logique pourrait conduire à une confusion des genres.

La relation entre les propriétés de la vérité et de la croyance peut être située entre deux tendances extrêmes : le *psychologisme*, où les premières sont déterminées par les secondes ; le *logicisme*, où les premières ne sont pas déterminées par les secondes et sont valables indépendamment. La logique épistémique se situe quelque part dans cet intervalle ; or compte tenu de l'évolution des méthodes de preuve en logique et de l'extension croissante de son domaine d'application au cours du 20e siècle, l'espace intermédiaire qu'occupe la logique épistémique semble être à géométrie variable. Comment considère-t-elle la relation entre vérité et croyance ? Pour mesurer la portée de ce système formel, revenons tout d'abord sur quelques notions de base et voyons ce qui unit et oppose les domaines de la logique et de la psychologie. La notion de *pensée* y est centrale.

Dans la logique propositionnelle classique, c'est-à-dire la logique portant sur les phrases déclaratives ou énoncés, plusieurs pensées singulières se combinent pour former un *raisonnement*, et la combinaison s'effectue au moyen de connecteurs logiques. Par exemple, 'César a conquis la Gaule *et* vaincu les barbares' est une pensée complexe dans laquelle le connecteur est la conjonction de coordination 'et', qui combine les deux énoncés atomiques 'César a conquis la Gaule' et 'César a vaincu les barbares'. Le rôle de la logique est de répertorier l'ensemble des relations qui font que des pensées complexes seront toujours vraies quel que soit leur objet du discours, et cet ensemble est décrit au sein d'une théorie générale de la *conséquence logique* ; cette théorie a pour but de définir les propriétés des connecteurs ou constantes logiques en termes de conséquence valide, c'est-à-dire en termes de ce qui, dans une relation déductive de prémisse à conclusion, permet de transformer les énoncés de départ tout en préservant la vérité des prémisses. L'introduction de la disjonction (de ϕ, inférer $\phi \vee \psi$), l'élimination de la conjonction (de $\phi \wedge \psi$, inférer ϕ) et le syllogisme disjonctif (de $\phi \vee \psi$ et $\neg\phi$, inférer ψ) sont trois exemples de règles de transformation dans lesquelles la vérité des prémisses est préservée dans la conclusion. L'indifférence des conséquences logiques vis-à-vis du contenu (de ce que ϕ et ψ symbolisent) explique l'appellation de logique *formelle*, et le logicien a pour tâche de recenser ces formes valides en faisant abstraction des énoncés pour ne retenir que des marques d'énoncés : on remplace ainsi en logique formelle les énoncés concrets par des lettres schématiques ou formules, afin de mettre en évidence la structure qui les relie. L'inventaire des relations logiques peut s'effectuer par plusieurs méthodes en logique moderne, qu'elles soient d'ordre syntaxique ou sémantique : axiomatique de Hilbert, déduction naturelle et calcul des séquents de Gentzen, dialogique de Lorenzen, tables de vérité, tableaux sémantiques de Beth, sémantique des mondes possibles, etc. (la liste n'est ni close ni exclusive). Quelle que soit la procédure choisie, la logique apparaît dans chaque cas comme une discipline tâchée d'établir les propriétés caractéristiques de la relation de conséquence entre un ensemble de prémisses et une conclusion. A la question : à partir d'un ensemble de prémisses que je considère comme vraies, que m'est-il permis d'inférer, le logicien répond que certaines prémisses impliquent une certaine conclusion et que, pour cette raison, celui qui admet leur vérité doit admettre la vérité de la conclusion. Un exemple d'inférence est la règle d'introduction de la disjonction : de la prémisse ϕ (que je suppose donc vraie), je dois inférer la vérité de $(\phi \vee \psi)$.[1] La stabilité de la discipline est supposée garantie par l'objectivité des relations sémiotiques qui composent le raisonnement : les relations entre des suites de signes $A_1, \ldots, A_n \vdash B$ sont régies par des lois que le sujet ne fait qu'appliquer et auquel il se subordonne, lorsqu'il effectue un acte de raisonnement en déduisant B de la suite des prémisses $A_1, \ldots, A_n$. Or la logique épistémique prétend à une description formelle des relations entre des attitudes psychologiques : si *a* croit que ϕ et si ϕ implique ψ, alors *a* croit que ψ, en vertu du système modal **K**. Si l'on admet un instant cet alignement des croyances sur les modalités logiques, toute référence à une activité mentale lors d'un raisonnement devient subsidiaire dans la mesure où les relations entre les actes de

[1] L'inférence désigne le fait de déduire une conclusion d'une prémisse, c'est-à-dire d'une hypothèse que nous avons admise au préalable ; si B est une conséquence logique de A, alors le sujet qui suppose la vérité de A doit inférer B en vertu de la relation déductive $A \Rightarrow B$. L'inférence indique ainsi un acte de raisonnement effectué conformément aux règles de la conséquence logique. Nous reviendrons plus bas sur l'aspect procédural de la notion d'inférence.

croyance seront déterminées par les relations formelles entre des modalités pures. Et si tel est le cas, cela signifie que le comportement des croyances est aligné sur celui des relations de conséquence logique, donc que la psychologie est déterminée par la logique. Peut-on admettre une telle conclusion, et est-ce vraiment celle-ci que la logique modale épistémique cherche à défendre ? Elle semble en effet suivre cette voie, lorsqu'elle décrit des relations de conséquence entre des opérateurs d'attitudes épistémiques et, à ce titre, elle tient plus du logicisme que du psychologisme. Mais il reste à mesurer le degré de subordination de ces attitudes aux méthodes de logique, ainsi que la valeur ajoutée apportée par une analyse de cette nature.

La distinction souhaitée entre logique et psychologie pourrait être obtenue par d'autres notions : norme et fait, loi universelle et fait particulier, ou type et token. Comparer la pensée logique (les règles de la pensée) à un type et la pensée psychologique (les actes de pensée) à un token, c'est-à-dire une instance ou un cas particulier, cela revient à dire que les actes de pensée particuliers ne font qu'illustrer la concrétisation d'opérations déductives invariables : de même qu'un programme ne fait qu'illustrer ou concrétiser les propriétés générales introduites par son programmeur, de même nos raisonnements particuliers ne feraient qu'appliquer les propriétés générales de la logique pure. La discorde porte ensuite sur le contenu de ces opérations : on les qualifiera de *règles* si ce sont des opérations choisies par convention, de *lois* si ce sont des normes[2] coercitives et imposées à nous. Ceci dit, tout logicien est bien conscient du fait que les règles de déduction ne sont pas suivies à la lettre par les agents réels ; au contraire, ces agents commettent fréquemment des erreurs de raisonnement. Un exemple notoire est la violation du *modus tollens* : $\phi \rightarrow \psi, \neg\psi \Rightarrow \neg\phi$.[3] Cette loi déductive me sert à inférer correctement une conséquence à partir d'une prémisse : si je reconnais le fait que les gens prennent leur parapluie pour sortir lorsqu'il pleut, et si je constate que les gens n'ont pas de parapluie actuellement, alors je peux en inférer qu'il ne pleut pas actuellement. Or une étude des inférences pratiques en psychologie expérimentale montrera que la plupart des gens ne suivent pas spontanément cette opération déductive et lui préfèrent une variante incorrecte : 's'il pleut, les gens prennent leur parapluie pour sortir ; or il ne pleut pas actuellement ; donc les gens ne prennent *pas* leur parapluie actuellement', de forme $\phi \rightarrow \psi, \neg\phi \Rightarrow \neg\psi$.[4] Cet exemple témoigne d'une différence manifeste entre les règles de déduction et nos inférences quotidiennes, qui expriment nos croyances en termes de raisonnements. En référence au vocabulaire de Frege, nous appellerons *pensée pure* l'ensemble des raisonnements logiquement corrects et *pensée en acte* l'ensemble des

[2] La notion de norme peut prêter à confusion : une norme peut être conventionnelle (précaution en éthique, norme d'appellation), auquel cas la différence s'estompe entre norme et prescription. Dans le sens évoqué ici, l'aspect normatif de la logique est comparable à l'idée de loi naturelle ou transcendantale.

[3] Au sens strict, le *modus tollens* est une règle de déduction du métalangage mais s'écrit parfois sous la forme d'une tautologie du langage-objet : la *contraposition*, $\vDash [(\phi \rightarrow \psi) \wedge \neg\psi] \rightarrow \neg\phi$. La contraposition est une implication (un conditionnel toujours vrai), tandis que le *modus tollens* est une règle de conséquence.

[4] Cette inférence incorrecte courante semble due à une confusion des agents sur l'idée d'implication : l'expression 'si ϕ alors ψ' ($\phi \rightarrow \psi$) connote une relation plus forte que le simple conditionnel et se traduirait plutôt par la relation d'implication 'ψ *chaque fois que* ϕ' ($\phi \Rightarrow \psi$). La confusion entre condition *nécessaire* et condition *suffisante* explique cette erreur de raisonnement : dans le conditionnel 's'il pleut, les gens prennent leur parapluie', 'il pleut' est une condition suffisante mais *pas nécessaire* pour que les gens prennent leur parapluie, donc la vérité de 'il ne pleut pas' n'implique pas nécessairement que les gens ne prendront pas leur parapluie (ils peuvent le prendre par précaution, pour se protéger du soleil, par pur dandysme, etc.).

raisonnements, qu'ils soient logiquement corrects ou incorrects. Mais cette double appellation crée d'autres problèmes : quelle sorte de pensée est étudiée par la logique épistémique telle qu'on la conçoit depuis [Hintikka 1962], et quelle sorte de pensée peut-elle aborder dans les limites de ses moyens ? L'évolution constante des techniques de démonstration en logique pose une difficulté pour fixer ces limites, donc une difficulté pour décider précisément de ce que la logique épistémique est capable de décrire. Enfin, quelles opérations faut-il suivre pour produire un raisonnement correct, et le nombre de ces opérations est-il constant ou relatif à des conventions d'usage ? Ces questions ne sont pas limitées au cas de la logique épistémique et renvoient à la philosophie de la logique ; mais selon toute vraisemblance, l'idée d'un traitement logique des verbes psychologiques risque de créer une certaine confusion entre les domaines et les questions générales portant sur la nature de la logique.

Pour le logicien du début du 20e siècle, le thème de la croyance apparaît comme parasite : quelle information nouvelle apporte-t-il au domaine de la logique et, s'il en apporte effectivement, ne confirme-t-il pas l'intrusion de la psychologie en logique ? Le problème posé par la croyance se pose en termes d'alternative. Soit la relation de conséquence valide préexiste aux sujets qui la manifestent dans leurs raisonnements appliqués ; dans ce cas, la question psychologique des croyances est sans effet sur le contenu de la logique. Soit les propriétés fondamentales d'un objet particulier, le concept de *vérité*, ne sont pas garanties indépendamment de l'activité mentale des sujets ; cela implique que ce qui permet de définir les propriétés générales du concept de vérité, la *conséquence logique*, n'est pas une liste définie de lois déductives mais un espace ouvert que l'on peut construire, inventer ou étendre. Cette alternative met en question le rôle et l'influence des actes ou opérations mentales du sujet pensant sur le contenu de ses opérations. Y a-t-il des lois de la pensée qui transcendent tous les sujets pensants ? La correction d'une opération est-elle totalement dépendante au contraire de son auteur ? Si elles ne sont valables ni pour tous ni pour un seul, faut-il considérer les règles de déduction comme des conventions d'usage admises au sein d'une communauté d'agents ; et si tel est le cas, comment délimiter cette communauté et qu'est-ce qui y garantit l'application de ces règles conventionnelles ? La réponse à ces questions dépend d'une théorie préalable de la vérité logique : qu'est-ce qui nous autorise à dire que certaines affirmations peuvent être déduites d'un ensemble de prémisses ; par exemple, qu'est-ce qui me permet d'asserter avec certitude qu'une disjonction est *nécessairement* vraie dès lors que j'ai admis la vérité d'un de ses membres ? Ces questions fondamentales renvoient aux fondements de la logique ; nous ne trancherons pas ce débat, qui dépasse de loin la portée de ce travail ; nous ne ferons que l'effleurer ou l'illustrer à travers le thème de la logique épistémique, en particulier avec le cas spécifique de l'omniscience logique.

L'*universalité* des lois logiques est remise en cause par une logique épistémique, mais celle-ci n'est évidemment pas la seule : toutes les logiques dites enrichies et, *a fortiori*, les logiques déviantes contestent l'idée que les frontières de la logique standard ne seraient ni extensibles ni révisables, et notre problème est celui de comprendre en quoi la logique épistémique peut contester ces limites. Le souci d'universalité, bien qu'atténué de nos jours et remplacé par des vues plus modestes, prédominait au début du 20e siècle. Que s'est-il passé entre-temps, et dans quelle mesure la notion même de logique épistémique incarne un

changement de paradigme en logique ? Quel est ce paradigme, en raison duquel l'idée même de logique épistémique tenait du non-sens au début du siècle ?

La situation décrite ici entre logique et psychologie préfigure un débat que nous associerons à la logique épistémique (voir section 3) : l'opposition entre *réalisme* et *anti-réalisme* en logique. Affirmer le primat des opérations de la pensée sur ses objets, n'est-ce pas réduire la norme au fait et s'exposer aux critiques de l'anti-psychologiste, lorsque celui-ci reproche au psychologiste de commettre une erreur de catégorie et de confondre le fait et la norme, la description et l'explication ? Pour distinguer états mentaux et logique, une réponse courante consiste à dire que logique et psychologie portent sur un thème commun, la pensée, mais l'abordent de deux points de vue distincts : la psychologie dit comment le sujet pensant raisonne *de fait*, tandis que la logique dit comment il *doit* raisonner.

On pourrait reprocher à ce qui précède d'établir sans raison valable un lien entre logique épistémique et psychologie : aucun logicien épistémique en vigueur ne se définirait comme psychologiste ; au contraire, dans la mesure où son rôle est d'analyser des concepts épistémiques par des méthodes de la logique formelle et, en particulier, de la logique modale *pure*. Le problème consiste à savoir si ces méthodes structurelles sont convaincantes. Comment justifier par exemple le fait que le concept de croyance satisfasse la propriété de clôture déductive (distribution du système **K**) ; autrement dit, dans quel sens toute personne croit-elle toutes les conséquences logiques de ce qu'elle croit ? C'est ce que prétend la logique épistémique, tout au moins ses versions de la logique modale *normale*. Rappelons encore une fois la distinction entre *garantir* une démonstration logique et la *justifier* : la validité d'un théorème est garantie par des axiomes et des règles de transformation au sein d'un langage formel ; elle est justifiée lorsque ce formalisme est utilisé à bon escient. L'axiomatique est censée baser les démonstrations sur des bases indubitables, mais les paradoxes de la logique épistémique montrent cependant que des prémisses indiscutables en apparence peuvent conduire à des résultats très discutables.

Face à une formalisation contestée des concepts épistémiques, on peut se demander à l'inverse si ce ne sont pas plutôt les faits psychologiques qui déterminent les formes logiques et, par là, la pensée correcte. Imaginons un instant cette possibilité qui, bien qu'elle ne soit pas du goût des logiciens contemporains, peut être envisagée si nous discutons des fondements épistémologiques de la logique. Pourquoi parler de conditions de vérité de la croyance, plutôt que de conditions de croyance de la vérité ? La position de Brian Ellis, qui se prétend psychologiste et holiste à la fois, reflète plus ou moins ce renversement du rapport entre vérité et croyance, à rebours des clauses de vérité récursives de Tarski (voir plus bas, section 2.1.3.2). Il y a toutefois des raisons de contester un tel renversement.

Un moyen d'éviter ce psychologisme consiste à affirmer que la signification d'une relation logique ne fait pas référence à une relation mentale ou à un contenu de pensée *privé* : les faits exigés pour établir la vérité d'un énoncé sont des *faits de langage* ; or un langage est une structure *publique* qui établit des relations entre des objets, donc le sujet ne fait qu'obéir à ces structures qui transcendent les affections de sa conscience particulière. Et si une relation logique pose un problème d'interprétation informelle lorsqu'elle est appliquée aux phrases d'un langage naturel, il s'agit alors de réviser cette relation plutôt que de faire appel à des raisons d'ordre psychologique. Bien que Frege ne prétendait pas lui-même appliquer la

logique formelle au langage ordinaire, l'un des deux principes fondamentaux de sa philosophie (toute philosophie de la pensée est une philosophie du langage) ne doit pas être remis en cause sous prétexte que certains résultats de la logique épistémique sont soit redondants, soit indéfendables. En d'autres termes, le partisan de la logique philosophique maintiendra l'idée que toute pensée doit être expliquée dans et par le langage, à condition de trouver les relations logiques (les structures formelles) adéquates et de les modifier selon les besoins du discours.

Comment rendre compte du caractère public des significations et de l'usage commun des énoncés, modaux ou non-modaux, si l'on ne s'accorde pas sur un ensemble minimal de règles communes afin de définir le sens des concepts et des énoncés ? Dans la philosophie du langage de Frege, le sens désigne la *valeur cognitive* apportée par un concept, et cette valeur de connaissance transmise par le langage ne se réduit pas à une impression subjective. La logique épistémique n'est pas une logique des croyances considérées comme des impressions privées, ce qui tiendrait de l'absurde ; néanmoins, cette valeur cognitive (le *Sinn*) n'appartient pas au domaine d'étude de la logique de Frege. Pourquoi ? Nous montrerons que la formalisation du sens que permet la logique épistémique étend l'analyse logique plus loin que ne l'admettait Frege, mais sans tomber pour autant dans le piège du psychologisme. Représentation (*Vorstellung*) et cognition (*Sinn*) se rapportent toutes deux à la psychologie, certes, mais à des degrés distincts : introduire la première notion dans la logique conduirait à un psychologisme, contrairement à la seconde.

Tel est l'argument de la variabilité de Frege, présenté ainsi dans [Engel 1996] : si la croyance est uniquement le produit d'une représentation mentale, alors elle constitue un événement privé et indescriptible, donc incommunicable ; or les croyances ne sont pas uniquement des représentations mentales puisqu'elles ont un contenu public, exprimé par un énoncé ; le fait de dire 'je crois qu'il fera beau demain' est l'expression d'une croyance dont le contenu est délivré par l'énoncé 'il fera beau demain', par exemple. Par conséquent, la position du psychologisme sera rejetée par l'absurde si elle réduit la croyance un état mental inexprimable. Cet argument paraît quelque peu forcé, toutefois, s'il consiste à réduire le domaine de la psychologie à une étude de représentations mentales : les arguments anti-psychologistes de Frege reposaient notamment sur l'idée que tout ce qui a trait au psychologique est d'ordre *privé* ; les représentations sont propres à un état mental, et deux sujets pensants n'ont jamais la même idée, la même croyance. Contre cet argument du caractère entièrement subjectif ou privé des croyances, on pourrait répondre que la psychologie expérimentale tente d'établir des lois régulières de l'activité psychologique, à l'image du programme des modèles mentaux de Johnson-Laird qui doit plus à la psychologie qu'à la logique.[5] L'observation de relations régulières en psychologie pourrait donc affaiblir le type d'argument anti-psychologiste de Frege, qui est un argument de la variabilité des états mentaux ; cette observation permettait également de rapprocher la psychologie de la logique et de faire de la logique épistémique une plate-forme commune entre les deux disciplines, sur la base d'un langage formel commun, la logique modale pure. Est-ce la cas ? Les premiers travaux de Hintikka semblent montrer que non.

---

[5] Voir P. N. Johnson-Laird : *Mental Models*, Cambridge University Press (1983) ; au sujet des représentations mentales, voir également J. Fodor : *The Language of Thought*, MIT Press, Cambridge, Mass. (1975).

Parler de la croyance ou de la connaissance dans une logique épistémique, ce n'est pas faire de la psychologie introspective mais constater des relations de conséquence entre des objets de discours. Or si tel est le cas, et si l'argument psychologiste est évacué, faut-il en conclure que la logique épistémique n'est rien d'autre qu'une étude des usages *linguistiques* des verbes psychologiques 'croire' et 'connaître', à supposer que ce dernier verbe soit encore d'ordre psychologique ? Dans ce cas de figure, on est non seulement en droit de se demander ce qu'une analyse des usages ordinaires ou publics des verbes d'attitude épistémique peut apporter à la *théorie de la connaissance*, mais également s'il est possible d'en tirer quelque chose d'essentiel par la logique formelle, c'est-à-dire par une étude de structures.

Comme nous le verrons, la justification des analyses logiques de K_a et B_a n'est pas si unanime qu'il y paraît et se trouve partagée entre deux tendances. Une première tendance, appelons-la la tendance du sens commun, établit des relations de conséquence logique pour K_a et B_a sur la base du sens que prennent ces attitudes épistémiques dans le langage ordinaire. Une seconde tendance, appelons-la la tendance scientifique, établit ces relations non sur les sens du langage ordinaire mais sur ce qu'en disent les 'sciences de la cognition' : la logique épistémique décrit alors des procédures de raisonnement et, en ce sens, se rapproche de la construction de modèles mentaux. La première tendance est proche de l'école d'Oxford mais, à la différence des Oxoniens, procède par un langage formel ; la seconde tendance prédomine dans la logique épistémique actuelle et fait collaborer logique et sciences cognitives. Les deux tendances correspondent plus ou moins à deux périodes distinctes de la jeune histoire de la logique épistémique : une première et une seconde génération, que nous aborderons dans l'ordre mais sans prétendre à un inventaire exhaustif.

Pour revenir à l'histoire de la logique, nous noterons que le début du siècle instaurait un quasi-relation d'identité entre croyance et psychologie. Pourquoi cette assimilation des deux, c'est-à-dire, pourquoi la distinction précédente entre une approche *conceptuelle* (linguistique et structurelle) et une approche *psychologique* (physique, ou physiologique) de la pensée n'avait-elle pas été faite par les principaux logisticiens, ni même par [Quine 1964] (voir la citation du début de section) ?
Avant d'étudier les arguments de plusieurs auteurs majeurs contre l'idée de logique épistémique, nous suggérerons trois hypothèses de travail :

– la distinction entre logique et croyance prolonge une distinction constante entre '*être* vrai' et '*tenir pour* vrai', ou entre fait logique et acte psychologique. Le passage de Frege qui suit est révélateur de cette distinction catégorielle : dans la perspective du psychologisme, "la vérité est ramenée à ce que les individus tiennent pour vrais" ; or il faut entendre par "les lois de la logique, non pas les lois psychologiques de ce qui est tenu pour vrai, mais les lois de ce qui est vrai".[6] La logique épistémique sera considérée comme une sorte de passerelle entre logique et psychologique (de 'logique psychologique'), en raison du lien qu'elle admet entre 'croire' et 'tenir pour vrai'. En ce sens, son analyse s'applique à ce que Frege excluait du domaine de la logique.

[6] G. Frege : *Grundgesetze der Arithmetik, Begriffsschrift abgeleitet* (1893-1903), trad. française de la préface, in J.-P. Belna, *La notion de nombre chez Dedekind, Cantor, Frege*, Vrin (1996), p. 328-9.

– l'apparition tardive de la logique épistémique est liée à l'apparition tardive de la sémantique relationnelle, mais ce retard n'est pas un simple accident ou fait contingent dans l'histoire de la logique ; il révèle un désaccord plus profond entre deux conceptions générales de la logique, et ce désaccord expliquerait une réticence tenace face à l'idée de logique épistémique (voir section 2.1.4.2).

– la logique épistémique oscille entre deux réductionnismes opposés : le psychologisme d'une part, le *logicisme* d'autre part. La logique épistémique 'seconde génération' nous semble bien porter *sur* des opérations de la pensée (considérée en acte);[7] elle en parle en termes de relations logiques qui, elles, obéissent à des lois déductives régulières. Le rôle de la logique épistémique consisterait ainsi à exprimer l'instable (la pensée en acte) en termes stables (la pensée pure), d'où l'impression paradoxale qui s'en dégage si elle prétend décrire des actes illogiques en termes logiques. Le paradoxe de l'omniscience logique sera considéré comme un cas du genre.

L'enjeu global est le suivant. Partant du problème de logique philosophique, qui est le problème de l'interprétation ou traduction adéquate des langues naturelles dans un langage formel, un autre problème philosophique plus profond sous-tend la pratique même de la logique philosophique. Dans quelle mesure une relation logique peut-elle s'étendre aux langues naturelles, ou à quel point peut-on expliquer une langue en termes de structures formelles ? La logique modale épistémique consiste-t-elle à exprimer des opérations de la pensée en acte dans un langage de la pensée pure ? Avant de trancher sur cette question, nous devrons expliquer avec plus de précision ce qui constitue la frontière entre logique et psychologie. Notre conviction de départ est que cette frontière a été progressivement modifiée par l'éclosion de certaines logiques étendues et non-classiques, parmi lesquelles la logique épistémique et la logique intuitionniste. Pour confirmer cette conviction ou pour l'infirmer, la suite du travail devra s'interroger sur les relations qu'entretiennent les trois domaines théoriques que sont logique, psychologie et épistémologie. Peut-on considérer la logique épistémique comme une sorte de logique de l'épistémologie, partant qu'elle est une logique des attitudes épistémiques, et comment articuler les notions de logique et de théorie de la connaissance?

Par opposition à la notion introspective de représentation, la notion de *sens* fregéen sera considérée comme l'exemple d'une annexion des questions de croyance par la logique formelle. Quels aspects de la pensée doivent être pris en compte pour fixer la signification des énoncés de croyance, et quel aspect de la croyance est réellement capté par le formalisme ?

Donner un sens à ces questions suppose que la frontière soit établie entre logique et psychologie ; nous illustrerons dans ce qui suit diverses raisons pour lesquelles des logiciens majeurs n'ont pas envisagé ou admis la possibilité d'une logique épistémique.

---

[7] Ainsi [Dubucs 1991] exprime-t-il une tendance cognitiviste de la logique épistémique, qui n'a plus grand rapport avec l'analyse linguistique des années 1960 et 1970: "Dans la mesure où notre tâche consiste à décrire le traitement d'information des croyants réalistes, l'omniscience logique doit être rejetée, car l'estimation est précisément le type d'activité dont des agents logiquement omniscients n'ont aucun usage. Nous devons donc concevoir une logique *x* telle que : logique épistémique/psychologie populaire = *x*/science cognitive" (p. 41).

2.1.2 Quelques adversaires d'une logique épistémique modale

Les attitudes propositionnelles posent doublement problème en logique : en tant que verbes psychologiques, d'une part ; en tant que modalités, d'autre part. On trouve dans l'histoire de la logique des opposants à la logique modale pure. Or cette logique enrichie constitue aujourd'hui une activité florissante de la logique 'philosophique'. La grammaire de Lambek ou la logique hybride sont deux exemples d'application de la logique modale qui n'ont que peu de rapport apparent avec des notions de philosophie, mais celles-ci ont pour point commun avec la logique modale épistémique de chercher à structurer le discours ordinaire. Cet état de fait ne répond toutefois pas à une question plus générale : dans quelle mesure la logique modale est-elle une logique ? Une certaine confusion actuelle entre *langage logique* et *langage formel* consiste à présenter le premier comme un ensemble d'opérations déterminé par la forme des relations, dont les méthodes sont récursives et peuvent être retranscrites au sein d'une algèbre pure (dont les objets sont quelconques). Tout langage formel est-il logique dès qu'il est réductible à des opérations telles que l'idempotence, la distributivité, etc. ? Cette question peut trouver une réponse à une condition précise : que la logique soit un pur langage, c'est-à-dire une discipline sans objet particulier. C'est cette dernière position que nous allons retrouver dans les critiques des auteurs qui suivent.

"Peut-il donc y avoir une logique modale ?", demande Fitting. Comme la majorité des logiciens actuels, celui-ci répond à cette question oratoire qu'

> Il n'y a pas de raison *a priori* de penser que non. Le moyen le plus efficace de montrer qu'il peut y avoir une logique modale est d'en construire une de fait.[8]

Mais construire un langage formel ne suffit pas à garantir que celui-ci soit de nature logique, tant que la question précédente ne trouve pas satisfaction. D'après Frege, les modalités ne changent rien au contenu propositionnel et, en tant que telles, elles ne sont pas constitutives des lois logiques. A ce type d'objection, Fitting répond que la logique modale normale ne change rien aux propriétés de conséquence de la logique standard, certes, mais pour la bonne raison qu'elle est une simple extension de cette dernière. Les axiomes de possibilité et de simplification : $\phi \rightarrow \Diamond\phi$ et $\Box\phi \rightarrow \phi$, sont connus depuis l'Antiquité et semblent démentir les propos de Frege. Mais encore une fois, la difficulté porte sur la nature logique de ces structures, sur ce qui fait qu'un langage formel est également un langage logique. Si l'on se contente du point de vue 'formaliste', en vertu duquel la logique est un ensemble *quelconque* d'opérations exprimées à l'intérieur d'un langage formel, alors le langage modal fera évidemment partie des langages de type logique ; mais si le langage logique est, comme l'entend Frege, une théorie de la *vérité pure* (considérée en soi) qui ne correspond pas à n'importe quelle opération mais à un ensemble de lois *spécifiques*, alors le caractère formel d'une opération ne suffit plus pour en faire une opération proprement logique. La distinction entre vérité absolue (vérité logique) et vérité dans une structure (vérité formelle) sera reprise plus loin, pour expliquer que les structures modales de Kripke ne

[8] M. Fitting & R. L. Mendelsohn : *First-Order Modal Logic*, Synthese Library, Vol. 77, Kluwer Academic Publishers, 1998, p. 5.

suffisent pas selon certains pour faire d'une théorie formelle des croyances une logique à part entière ([Kielkopf 1971], [Hocutt 1972]).

Dans ce qui suit, nous allons examiner les arguments de plusieurs logiciens contre l'idée de logique modale et interpréter leurs positions à partir de la distinction ci-dessus : langage formel et langage logique. Frege, Wittgenstein, Russell et Quine partagent (à des degrés divers) une conception scientifique de la logique qui se présente alors comme un 'ortho-langage' ; plus que de mettre de l'ordre dans de possibles structures cachées du langage ordinaire, il s'agit de *corriger* ou rectifier les ambiguïtés des langues naturelles afin de construire un seul et unique langage idéal. L'enrégimentation dans un langage canonique suppose que celui-ci contient des canons de la pensée correcte (et pas seulement claire) ; dans cette perspective, Frege présenta son idéographie comme une étude des propriétés les plus générales du fait d'être vrai [*Wahrsein*], le Vrai majuscule opposé à l'étude des vérités particulières dont s'occupent les théories scientifiques (physique, histoire, psychologie, etc.).[9] Définir la vérité et ses propriétés fondamentales, n'est-ce pas distinguer le bon du mauvais raisonnement ? De ce point de vue, les axiomes modaux ci-dessus ne font pas figure de lois logiques parce que $\Box$ et $\Diamond$ ne déploient pas des propriétés fondamentales du Vrai ; d'autre part, ces modalités sont certes réductibles à des relations purement formelles, mais elles représentent une régression vers les imperfections du parler ordinaire. Une telle extension des constantes logiques (de ce qui sert à définir les propriétés fondamentales du vrai) s'effectuerait au prix d'une relativisation des propriétés du Vrai : parler des modes de la vérité (nécessaire, possible, contingente) revient à en dire plus que sur la vérité elle-même. Par opposition à cette spécification du Vrai, nous qualifierons d'*universalisme logique* cette attitude philosophique consistant à limiter le domaine de la logique aux propriétés du vrai les plus générales. Mais la notion même de généralité reste à motiver. A quel moment passe-t-on d'une relation proprement logique à une relation simplement formelle ; autrement dit, quel degré de complexité des structures est autorisé pour parler encore d'un langage formel comme d'une logique ?

Quoi qu'il en soit, cette position universaliste se retrouve également dans le Wittgenstein du *Tractatus* : le langage logique a une mission thérapeutique, celle de corriger les erreurs de pensée dues à l'ambiguïté des langues naturelles ; mais cette correction ne s'obtient pas par une construction *arbitraire* d'opérations bien définies. L'idée est que les lois de la logique doivent s'imposer à nous et distinguer les propositions dépourvues de sens des autres, ces autres étant les 'propositions' vraies ou fausses inscrites dans l'espace logique. A cet égard, le critère grammatical précédent de la distinction entre phrastique et neustique ne suffit pas pour garantir la nature logique de la logique modale. La logique de l'universaliste doit au contraire s'affranchir des structures grammaticales et déterminer un espace de discours qui délimite le pensable et l'impensable.

---

[9] "La psychologie est concernée par la vérité uniquement au sens ou les autres sciences le sont, à savoir que son but est d'étendre le domaine des vérités ; mais l'étude de la propriété 'vrai' n'appartient pas à son domaine de recherche comme les propriétés 'lourd', 'chaud', etc., appartiennent à celui de la physique. Cette étude relève de la logique", dit Frege. D'où la présentation de la logique comme étude des conditions de préservation générales de la vérité, sans plus de détails sur les circonstances dans lesquelles une prémisse est considérée comme vraie ou non.

Justifier l'actuelle logique modale épistémique exige donc un premier examen de ces auteurs universalistes, afin d'évaluer l'hypothèse d'un 'ortho-langage' et de décider si, oui ou non, la distinction entre langage logique et langage formel est justifiable. La paraphrase des expressions d'un langage formel dans une langue naturelle dépend de cette hypothèse : si la logique a un statut canonique, alors l'idée même de créer des logiques 'faites pour' les langues naturelles est absurde puisque cela reviendrait à reformuler le langage ordinaire plutôt qu'à le corriger. Pour montrer qu'elle n'est pas un langage canonique, il s'agira de montrer que la fonction correctrice de la logique a été dévaluée entre-temps et remplacée par une fonction plus modeste de construction de structures formelles. Mais accepter cet affaiblissement de la fonction logique suppose que l'on rejette la définition de la logique comme étude des propriétés universelles et fondamentales du Vrai. Deux conceptions générales s'opposent à ce sujet : la *logique comme éthique*, selon laquelle il n'y a qu'une logique susceptible de bien conduire la pensée *a priori* ; une *logique comme physique*, selon laquelle les propriétés du Vrai ne sont pas préétablies mais décrites par des structures révisables.

Examinons d'abord les arguments déployés par les universalistes, ainsi que leurs raisons de refuser une logique modale épistémique.

2.1.2.1 Frege

Pour résumer ce qui suit, la logique épistémique était non pas impossible selon Frege mais inutile, pour la raison que la logique épistémique porte son attention sur des *jugements* et non sur des *propositions*. Pour Frege, la logique porte uniquement sur des propositions, c'est-à-dire des sortes d'énoncés abstraits ou éternels et dont la valeur sémantique ne change pas avec le contexte d'énonciation ; quant aux jugements, leur étude reviendrait non pas à la logique mais à la psychologie. Pourquoi cette division du travail ?

D'après Frege, la logique a pour but d'organiser les *pensées* et de déterminer leurs critères de correction. Mais comme nous l'avons déjà signalé, la notion de pensée est ambiguë, prise entre les eaux de la logique et de la psychologie. Chez Frege, la pensée ne désigne absolument pas un acte mental ou une impression subjective quelconque. Pour préciser les termes, il propose une distinction entre *proposition* et *jugement* : la pensée [*Gedanke*] correspond à une entité objective, et celle-ci n'a rien à voir avec la réaction subjective du locuteur qui accompagne l'acte d'énonciation d'un jugement. Il s'en explique ainsi :

> J'entends par pensée [*Gedanke*] non pas l'acte subjectif de penser mais son contenu objectif, lequel peut être la propriété commune de plusieurs sujets.[10]

Conformément à sa distinction systématique entre sens et référence (établie à partir de [Frege 1892]), chaque expression se divise en deux catégories d'objets. Lorsque l'expression est un énoncé [*Satz*], c'est-à-dire une phrase déclarative, le sens de cette expression est une pensée [*Gedanke*] et sa référence [*Bedeutung*] est une valeur de vérité [*Wahrheitswert*]. La pensée telle que Frege la définit ne représente donc pas un état mental particulier, qu'il assimile à la représentation [*Vorstellung*], mais le sens d'un énoncé, et ce sens a pour

[10] [Frege 1971], p. 108, note 1.

caractéristique de pouvoir être commun à plusieurs sujets : plusieurs sujets peuvent avoir une même pensée, par opposition aux représentations de la psychologie. Lorsque deux sujets expriment leur jugement par les énoncés respectifs 'La neige est blanche' et 'Snow is white', nous sommes en présence de deux énoncés différents ; mais c'est une seule et même pensée qui est exprimée et qui correspond à une proposition commune : la pensée que la neige est blanche, exprimée par la forme complétive '*que* la neige est blanche' (une *that-clause*, en anglais). La proposition est une entité extra-linguistique qui transcende ses expressions verbales et justifie le recours à une langue artificielle, l'*idéographie*, dont le but est d'éviter l'ambiguïté des langues naturelles en symbolisant uniquement la pensée qui leur est commune. De ce fait, décrire la logique comme l'étude des lois de la pensée ne conduit absolument pas à un psychologisme : la 'logique des propositions' porte sur les significations objectives des énoncés et fait abstraction de leur usage circonstanciel. La logique ne s'occupe pas des énoncés proprement dits mais de leur contenu, c'est-à-dire de la proposition. Pour éviter la confusion entre proposition grammaticale (un énoncé de forme sujet-complément) et proposition logique, on appelle également 'contenu propositionnel' la pensée ou proposition qui intéresse le logicien.[11] La proposition reflète la pensée des sujets ; ces pensées sont communes à plusieurs locuteurs et correspondent à ce que Bolzano appelait la proposition en soi [*Satz an Sich*], dont la signification est indépendante des formes d'énonciation dans lesquelles elle est incarnée au sein des langues naturelles. Certaines modifications grammaticales n'ont aucune influence sur l'identité de la pensée : c'est le cas de la distinction entre voix active et voix passive. Dans 'César a envahi la Gaule' et 'La Gaule a été envahie par César', par exemple, nous avons deux énoncés distincts mais une seule et même pensée. Autrement dit, les deux auteurs de ces énoncés pensent à la même 'chose' (la même proposition) lorsqu'ils expriment leur croyance par le biais de leur expression verbale. Mais d'autres modifications grammaticales influent en revanche sur l'identité de la pensée exprimée. Ainsi, une modification du mode de *nomination* du terme sujet peut changer la pensée qui est exprimée par l'énoncé : dans 'Jules César a envahi la Gaule' et 'L'auteur des *Commentaires de la guerre des Gaules* a envahi la Gaule', les deux énoncés n'expriment pas la même pensée bien que leur terme sujet porte sur une seule et même référence. La raison pour laquelle ces deux pensées sont distinctes se trouve dans le principe de *compositionalité* de Frege : la valeur sémantique d'un énoncé est déterminée par la valeur sémantique de ses parties. Ce principe se décline de deux manières : le sens d'un énoncé est déterminé par le sens de ses parties, et la référence d'un énoncé est déterminé par la référence de ses parties.[12] Or puisque le sens se définit comme le mode de donation de la référence [*Gegebenheitsweise*], un changement dans le mode de donation du terme sujet entraîne un changement dans la proposition : 'Jules César' et 'l'auteur des *Commentaires de la guerre des Gaules*' représente deux différents sens du terme sujet, donc le sens de l'énoncé dans lequel

[11] Un dictionnaire francais-allemand courant traduit *Satz* par 'énoncé', 'phrase' ou 'proposition', mais au sens de la proposition grammaticale et non de la proposition logique. Le *Satz* de Frege prend pour sens le *Satz an Sich* de Bolzano, c'est-à-dire ce qui est exprimé 'en soi' (en propre) par un énoncé.

[12] Puisque la référence d'un énoncé est une valeur de vérité, et son sens une proposition, on a les règles de signification suivantes : la valeur de vérité qu'un énoncé dénote (sa référence) est déterminée par la référence des expressions qui y figurent, et la *proposition* qu'il exprime (son sens) est déterminée par le sens de ses expressions. Le sens d'une expression (la manière de présenter une référence) est indépendant de la référence de l'expression, et inversement : César reste César quelle que soit le nom qu'on lui associe.

ils figurent est différent. La référence de ces deux modes de dénotation reste identique, en revanche, donc la référence de l'énoncé ne change pas : les énoncés 'Jules César a envahi la Gaule' et 'L'auteur des Commentaires sur la guerre des Gaules' ont la même référence, c'est-à-dire la même valeur de vérité, mais ils n'ont pas le même sens, c'est-à-dire qu'ils n'expriment pas la même proposition. En d'autres termes, l'ordre des expressions dans un énoncé n'est pas aussi important que le choix de ces expressions. L'ambiguïté des expressions vernaculaires justifie l'importance de la logique de Frege considérée comme un ortho-langage, puisque son écriture idéographique a pour fonction de dépouiller la pensée ordinaire de son 'habillage' grammatical superflu au sein des langues naturelles. Pour garantir l'objectivité des lois de la pensée et le passage d'une pensée inexacte à une pensée scientifique, Frege postulera un troisième domaine [*Dritten Reich*] d'entités abstraites, le domaine des pensées que saisissent les locuteurs et qu'ils expriment diversement en termes d'énoncés particuliers. Quant à la pensée étudiée par la psychologie, la pensée en acte, sa correction dépendra de sa capacité à saisir ou reconnaître les lois de la pensée pure.

Pour comprendre le lien établi entre une pensée objective et un jugement subjectif, Frege dit ceci :

> On distinguera donc :
> 1. La saisie de la pensée – l'acte de penser.
> 2. La reconnaissance de la vérité d'une pensée – le jugement.
> 3. La manifestation de ce jugement – l'assertion [*die Kundgebung dieses Urteils – das Behaupten.*[13]

La pensée n'est donc pas l'acte psychologique de penser, mais l'objet visé par cet acte ; quant à notre tripartition entre énoncé, proposition et valeur de vérité, elle illustre à la tripartition établie par Frege entre un nom propre (un signe pourvu de sens), son sens et sa référence. La logique fregéenne a pour but de s'abstraire des contingences du langage ordinaire. Les ambiguïtés d'usage des expressions peuvent porter à confusion sur la signification des énoncés : deux locuteurs veulent-ils parler du même individu, lorsque l'un parle de 'Jules César' et l'autre de 'l'auteur des *Commentaires sur la guerre des Gaules*'? Rien ne le garantit, sinon la capacité des deux locuteurs à associer ces deux noms propres au même individu. Cet inconvénient vernaculaire est corrigé au sein du langage idéographique : dans celui-ci, à une pensée correspond un et un seul symbole formel. On retrouvera plus bas ce souci d'univocité du formalisme dans [Wittgenstein 1969]. Lorsque l'on parle de logique des propositions, la proposition dont il est question n'est donc pas un énoncé quelconque mais son sens, c'est-à-dire l'entité extra-linguistique que cet énoncé exprime. Par conséquent, ce que Frege appelle la 'pensée' correspond à ce que nous appelons ici une proposition, mais cette proposition abstraite (la *Gedanke*) ne doit surtout pas être confondue avec son mode d'expression naturel (le *Satz*). Ce sur quoi Frege reste plus discret, c'est la question de ce qui permet d'établir un lien entre un énoncé et une proposition. Si la pensée de Frege n'est rien d'autre que le sens d'un énoncé, c'est-à-dire le sens d'une phrase déclarative susceptible d'être vraie ou fausse, c'est parce que le sens se définit comme le mode de dénotation d'une expression. De ce fait, le sens d'un énoncé est effectivement public puisqu'il s'exprime par le

[13] [Frege 1971], p. 175-6.

biais d'expressions du langage ordinaire. Mais il reste à préciser une chose : le critère d'identité des propositions, c'est-à-dire ce qui permet d'affirmer que deux énoncés distincts expriment bien la même proposition. Comment démontrer que deux locuteurs ont 'la même pensée en tête' ou, pour le dire autrement, qu'ils saisissent la même 'chose' par des moyens d'expression différents ? L'absence d'un critère d'identité propositionnelle a été reprochée plus tôt aux logiciens des modalités aléthique (voir Introduction), lorsque ceux-ci prétendaient expliquer la nécessité par la synonymie puis par l'identité de signification entre deux énoncés : mais comment identifier la signification d'un énoncé, et quel est l'intérêt explicatif d'une modalité aléthique si celle-ci dépend précisément de ce que nous ignorons encore, à savoir la nature du lien entre un énoncé et la proposition qu'il exprime (sa signification) ? Nous rappellerons ce problème de la notion de signification, dans le cadre de la philosophie du langage de Quine. Précisons toutefois que si Frege ne parlait de logique qu'en termes de propositions et non d'énoncés, c'est parce que son idéographie portait sur une catégorie d'expressions dépourvues d'ambiguïté, à savoir les expressions du langage mathématique, et que la question à venir de l'opacité référentielle n'avait pas lieu d'être dans ses réflexions sur les fondements de l'arithmétique.

Quant à la position de Frege sur les modalités, elle est semblable à celle que l'on retrouvera chez Russell ou Quine. Le passage qui suit présente la nécessité comme l'expression d'un *jugement* de certitude d'une conclusion relativement à ses prémisses :

> Le jugement apodictique diffère du jugement assertorique en ce qu'il suggère l'existence de jugements universels dont la proposition puisse être inférée, tandis que dans le cas du jugement assertorique pareille suggestion fait défaut. En disant qu'une proposition est nécessaire je donne une indication sur les raisons en faveur de mon jugement. *Mais puisque celle-ci n'affecte pas le contenu conceptuel du jugement, la forme du jugement apodictique n'a pas d'importance pour nous.* Si une proposition est présentée comme possible, soit le locuteur suspend son jugement en laissant entendre qu'il ne connaît pas de lois dont résulterait la négation de la proposition, soit il dit que la généralisation de la négation est fausse.[14]

Si l'on considère une tautologie comme une sorte de jugement apodictique, cela veut dire qu'elle sera connue avec certitude dans toutes les situations que le locuteur envisage : pour tout modèle dans lequel les prémisses sont vraies, la conclusion est vraie également et ne peut pas être autrement sans produire une contradiction. La règle de nécessitation du système modal **K** pourrait symboliser la citation ci-dessus, mais la limitation technique du langage logique de Frege et sa conception de la logique comme science du Vrai l'empêchaient de suivre une telle démarche.

Frege aurait-il admis la sémantique relationnelle, s'il eût choisi une autre idéographie ? Il semble que non : ce n'est pas que la forme apodictique est inexprimable, c'est qu'elle n'a "pas d'importance pour nous" et n'apporte rien à l'idéographie, qui n'est pas une construction

[14] G. Frege, *Begriffschrift* (1879), p. 13. Le jugement *apodictique* (*a* est nécessairement F) correspond à l'une des trois modalités du jugement de Kant dans sa *Critique de la Raison Pure*, qu'il distingue des jugements *assertoriques* (déclaratifs : *a* est F) et *hypothétiques* (conditionnels : si *a* est F, alors *a* est G). La différence avec Frege est que Kant assimile les jugements à des propositions, d'où une certaine confusion entre pensée comme objet logique et pensée comme acte psychologique. [Lemmon 1977] explique dans le même sens que "le point de vue de Frege dans la *Begriffschrift* est que les notions modales telles que 'doit' sont en réalité épistémiques et font référence à la connaissance humaine ; en tant que telles, elles sont sans rapport avec la logique pure" (p. 5).

arbitraire mais constitue l'armature de toute pensée vraie. Nous relèverons plus loin des remarques de Hintikka montrant que le rejet par Frege des modalités était plus profond. L'argument d'une limitation technique du formalisme de l'époque n'est pas un argument recevable : l'école algébriste de Boole et Schröder, d'une part, l'axiomatique formaliste de Hilbert d'autre part auraient pu servir de base théorique pour construire un langage modal à part entière, mais la raison pour laquelle Frege ne suivit pas Hilbert sur son approche formaliste marque une rupture plus profonde entre deux approches de la logique. La raison de ce rejet est liée encore une fois à l'approche universaliste de Frege, une approche que Hintikka opposera plus loin à une filiation entre la doctrine formaliste de Hilbert et la future théorie des modèles.

Dans [Frege 1919a], le logicien de Iéna présente la logique comme une sorte de métathéorie de la vérité. Les sciences naturelles portent sur *ce qui* est vrai, la science logique porte sur *ce qu'*est le vrai :

> De même que le terme 'beau' renvoie à l'esthétique et le terme 'bon' à l'éthique, le terme 'vrai' renvoie à la logique. Certes, toutes les sciences ont la vérité pour but, mais la logique s'en occupe un peu comme la physique a trait à la pesanteur ou à la chaleur. Découvrir des vérités est la tâche de toutes les sciences, mais c'est à la logique qu'il appartient de connaître les lois de l'être vrai.[15]

Ces lois logiques délimitent les règles de la pensée pure, la pensée valable pour tout locuteur et dans tous les contextes de discours. Elles sont consignées au sein de son idéographie et correspondent à des lois fondamentales ou axiomes auxquelles toutes les autres vérités sont réductibles. Frege en compte neuf. Trois axiomes définissent le vrai en termes de négation et de conditionnalité (conditionnel, ou implication matérielle) : $(\phi \rightarrow \psi) \rightarrow (\neg\psi \rightarrow \neg\phi)$, $(\neg\neg\phi \rightarrow \phi)$, et $(\phi \rightarrow \neg\neg\phi)$. Trois axiomes le font en termes de conditionnalité : $(\phi \rightarrow (\psi \rightarrow \phi))$, $((\phi \rightarrow (\psi \rightarrow \varphi) \rightarrow ((\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \varphi))$, et $((\phi \rightarrow (\psi \rightarrow \varphi) \rightarrow (\psi \rightarrow (\phi \rightarrow \varphi))$. Deux sont relatifs à la notion d'identité : $((a = b) \rightarrow (F(a) = F(b))$, et $(a = a)$. Le dernier axiome concerne la généralité (le quantificateur universel) : $\forall xF(x) \rightarrow F(c)$. Les deux axiomes précédents seront revisités dans la logique épistémique, où ils conduisent à des cas de paradoxes.

Dans le cadre frégéen, l'idéographie désigne une écriture conceptuelle qui est plus qu'une simple structure de signes. C'est un langage formulaire de la pensée pure où chacun des axiomes est supposé exprimer une propriété fondamentale de la notion de vérité. A partir de cet inventaire des lois de la vérité, ce langage formulaire permet de donner les conditions de préservation de la vérité et distingue deux acceptions de la notion de pensée : à la pensée pure qui décrit ces lois éternelles de l'être vrai, Frege subordonne la pensée appliquée en acte, dont la condition minimale pour poursuivre la vérité est de suivre les propriétés indépendantes de la pensée pure. De cette séparation entre les lois logiques de la pensée pure et les actes psychologiques de la pensée en acte, Frege tire une série d'oppositions conceptuelles : *conséquence logique* et *inférence*, *vérité* et *reconnaissance de la vérité*, *contenu jugeable* et *jugement*, *simple considération* et *assertion*. Dans chacun de ces cas, la logique se définit par les premiers termes et les seconds expriment simplement l'application concrète des lois de la

[15] [Frege 1971], p. 170.

logique par un sujet pensant. L'inférence [*Schluß*] désigne une relation entre des jugements, c'est-à-dire des énoncés reconnus ou tenus pour vrais par le sujet :

> Une inférence (...) est un acte de jugement tiré de jugements effectués précédemment, sur la base de lois logiques. Chaque prémisse est une certaine proposition qui a été reconnue comme vraie, et donc une certaine proposition est reconnue également comme vraie dans le jugement de conclusion.[16]

En d'autres termes, l'inférence est une preuve déductive qui repose sur les lois logiques, et Frege considère ces lois comme des propriétés inaliénables de la vérité.

Distincte de l'inférence, la conséquence logique [*Folgerung*] exprime les lois de l'être vrai et décrit une relation de propositions simplement considérée, sans l'acte psychologique de jugement à l'appui. Dans la lignée du *Wissenschaftslehre* de Bolzano, la relation de conséquence entre les prémisses $\phi_1, \ldots, \phi_n$ et la conclusion B est considérée comme valable en soi, indépendamment du locuteur qui l'applique lors d'une inférence ; pour exprimer le rôle secondaire du sujet pensant par rapport aux lois de la pensée pure, Frege parle de *re*connaissance de la vérité, et l'activité recognitionnelle du sujet sous-entend que celui-ci ne fait que manifester ou reproduire des lois qui lui préexistent. Lorsque les parties d'une conséquence logique sont reconnues comme vraies, la conséquence devient une inférence et les propositions de base deviennent des jugements. Si l'on symbolise l'acte de jugement en préfixant la lettre 'J' aux propositions, on obtient l'inférence de $J(\phi_1), \ldots, J(\phi_n)$ à la conclusion $J(\psi)$: cette inférence produit un jugement *valide*, mais elle n'ajoute rien au contenu des lois logiques et ne fait qu'exprimer leur re-cognition par le sujet.[17]

Deux choses intéressantes peuvent être notées ici, en rapport avec la logique épistémique.

Premièrement, la reconnaissance de la vérité d'une proposition est obtenue à partir d'une preuve matérielle ou donnée probatoire (une 'évidence'), et la langue allemande traduit par *Erkenntnis* ce résultat d'un jugement vrai. Cette définition permet de préciser le rapport de subordination entre logique et théorie de la connaissance, ou *Erkenntnislehre* : la première détermine des relations de conséquence entre des séquences de propositions, et non de jugements. D'après Bolzano, nous pouvons certifier la validité de la relation de conséquence *à*

---

[16] G. Frege : "Über die Grundlagen der Geometrie", *Jahresberichte der Deutschen Mathematiker-Vereinigung* **15**(1906), pp. 377-403, en part. p. 387. Cité in G. Sundholm : "Inference versus Consequence", in *LOGICA Yearbook 1997*, Filosofia Publishers, Czech Academy of Science, Prague, pp. 26-36.
Il y a une ambiguïté possible ici entre *reconnaître* une vérité et *tenir pour* vrai : un jugement doit-il être relié à une pensée vraie pour être un jugement? Pour éviter cette absurdité, on supposera plutôt que Frege entend par 'reconnaissance de vérité' la même chose que le simple fait de 'tenir pour vrai'. Sans cela, toute pensée serait une pensée vraie *eo ipso* et la notion de 'pensée fausse' deviendrait une contradiction dans les termes. Nous reviendrons au Chapitre 3 sur cette ambiguïté, caractéristique de la notion d'*assertion.*

[17] On distingue pour cette raison les relations de *déduction* (ou *conséquence syntaxique* : $A \vdash B$) et de *conséquence logique* (ou *conséquence sémantique* : $A \vDash B$), d'une part, et la relation d'*inférence* d'autre part : les deux premières relations représentent une relation objective entre un antécédent A et un conséquent B, respectivement d'ordre syntaxique et sémantique ; la seconde relation représente un acte subjectif consistant à tirer la conséquence B de la prémisse A (de A, j'infère B), par des moyens soit syntaxiques soit sémantiques. Inférer, c'est donc appliquer un schéma logique ; dans la mesure où un schéma préexiste à son instanciation, Frege suppose que les lois de l'être vrai sont des propriétés de conséquence logique indépendantes des sujets.

l'aveugle, c'est-à-dire sans avoir à justifier la vérité des prémisses par une preuve matérielle.[18] Cela signifie que les prémisses d'une relation de conséquence n'ont pas à être reconnues, c'est-à-dire établies comme vraies, pour que la conclusion soit garantie comme découlant des prémisses. Frege exprime ceci en distinguant l'assertion d'une proposition, qui donne un jugement (*Urteil*), et la simple considération ou assomption (*Annahme*) de cette proposition, qui donne un contenu jugeable (*beurteilbare Inhalt*). La différence entre les deux réside dans l'engagement ou non du locuteur à l'égard de la vérité des prémisses. Quant à la théorie de la connaissance, elle détermine les moyens d'obtenir une preuve matérielle pour chaque proposition individuelle. Dans le cas de la relation de conséquence entre les prémisses ϕ, ψ et la conclusion $\phi \wedge \psi$, ϕ et ψ n'ont pas besoin d'être reconnues comme vraies et sont simplement supposées comme telles, donc les lois logiques sont appliquées sans théorie de la connaissance à l'appui. La vérité logique est purement hypothétique : *si* l'on suppose que ϕ et ψ sont vraies, *alors* le passage à la conclusion $\phi \wedge \psi$ est garanti par les propriétés du Vrai. Ainsi, la différence entre logique et théorie de la connaissance correspond à peu près à la différence courante entre vérité formelle et vérité matérielle. Or si une logique épistémique concerne l'analyse logique de concepts appartenant à la théorie de la connaissance, la distinction entre attitude propositionnelle, croyance vraie justifiée et jugement semble remise en cause.

Deuxièmement, l'une des difficultés de principe de la logique épistémique concerne la question de savoir si la reconnaissance de la vérité par le sujet obéit à des lois aussi régulières que la logique formelle. Qu'est-ce qui, partant d'un ensemble de prémisses, permet à ce sujet de passer à la conclusion, et les conditions de ce passage sont-elles données par la logique elle-même ? Bien que la logique modale épistémique procède en représentant les propriétés des attitudes épistémiques en termes de vérité logique, nous verrons qu'une autre façon d'aborder le problème consisterait au contraire à représenter les propriétés mêmes de cette vérité logique en termes de compétence épistémique. Cette compétence est-elle de nature psychologique, ou concerne-t-elle plutôt la doctrine intuitionniste de la vérité comme vérifiabilité ? Quoi qu'il en soit, il est encore une fois question d'une inversion des rapports explicatifs entre vérité et reconnaissance de la vérité, ici ; les solutions proposées par Hintikka à l'omniscience logique iront dans ce sens : décrire les compétences cognitives du sujet pour garantir l'application d'une inférence. Il semble difficile de réduire les thèmes de la logique épistémique à un unique problème : l'individuation des individus, l'omniscience logique, l'introspection positive, les systèmes de révision de croyance, etc. n'ont de commun que les concepts épistémiques sur lesquels ils portent, mais nous chercherons pour notre part à en tirer une problématique commune : dans quelle mesure la relation logique de conséquence recouvre-t-elle la relation épistémique d'inférence ?

Un problème de recouvrement de ce genre concerne la distinction entre la *valeur de vérité* d'une proposition et sa *valeur cognitive*, ou informative. Dans [Frege 1892], Frege introduit la

[18] La déduction 'à l'aveugle' fut évoquée dans la *Wissenschaftslehre* de Bolzano (§155(2) : notion de déductibilité). En un sens, le jugement correct aveugle est une contrepartie de ce que les épistémologues appellent une connaissance basique ou directe : elle ne repose sur aucune preuve préalable et s'impose d'elle-même. Les lois logiques n'ont pas à être démontrées, puisqu'elles sont les principes de toute démonstration.

distinction entre ces deux types de valeurs pour expliquer la différence entre les deux égalités 'a = a' et 'a = b'. Cette différence ne réside pas dans la valeur des termes d'individus *a* et *b*, puisque l'égalité indique que *a* et *b* désignent la même chose ; elle réside dans la manière dont un seul et même objet est exprimé, et c'est le mode de donation (*Gegebenheitsweise*) des référence qui permet d'identifie le sens (*Sinn*) de son expression. 'Le maître d'Aristote' et 'Platon' sont deux noms propres (au sens large de Frege, qui entendait par 'nom propre' n'importe quel terme doué de sens)[19] de même référence mais dont le sens diffère, puisque le même individu Platon n'est pas présenté de la même manière. Sur la base de cette distinction systématique entre sens et référence, Frege va proposer une solution à un problème bien connu des futurs logiciens épistémiques : le problème de l'*opacité référentielle*. En vertu de la loi de substitutivité des identiques, une expression propositionnelle conserve la même valeur de vérité lorsqu'à une de ses expressions individuelles est substitué une autre expression de même référence. Par voie de conséquence, cela signifie que les deux expressions propositionnelles ϕ et ψ obtenues après substitution auront même valeur de vérité dans tous les contextes de discours C (symbolisés ci-contre par des opérateurs) : $(\phi \equiv \psi) \rightarrow (C(\phi) \equiv C(\psi))$. On retrouve cette propriété fondamentale de la logique consignée dans l'idéographie de Frege, sous la forme : $((a = b) \rightarrow (F(a) = F(b))$. Or cette loi semble échouer dans certains contextes de discours, parmi lesquels le contexte épistémique : même si Jean sait que 2 + 3 = 5, il ne sait pas forcément que le nombre 5 désigne la même chose que le nombre des joueurs dans une équipe de basket-ball ; or si l'on applique la loi de substitutivité, la co-référence des noms propres '5' et 'le nombre des joueurs dans une équipe de basket-ball' (notée : a = b) signifie que les propositions 'Jean sait que 2 + 3 = 5' (notée : C(a), où C symbolise le contexte de connaissance) et 'Jean sait que 2 + 3 = le nombre des joueurs d'une équipe de basket-ball' (notée : C(b)) doivent avoir la même valeur de vérité (notée : F(a) = F(b)). Or la deuxième formule peut être fausse si Jean ne connaît rien aux choses du sport. Dans ce cas, le critère de l'identité de référence ne suffit pas pour garantir la valeur sémantique d'une formule propositionnelle, et la loi de substitutivité est falsifiée. On obtient ainsi le contre-exemple suivant, où '1' et '0' symbolisent la valeur sémantique de chaque formule ('vrai' et 'faux') :

F(a)	1
<u>(a = b)</u>	1
∴F(b)	0

Pour préserver cette soi-disant 'loi de l'être vrai' et montrer que ce paradoxe n'en est pas un, Frege propose de distinguer d'abord les discours directs (*gerade*) des discours indirects (*ungerade*), puis de délimiter le domaine d'étude de la logique.

D'après Frege, l'échec apparent des lois d'extensionalité ci-dessus vient d'une confusion dans l'interprétation des énoncés, selon qu'ils dénotent une proposition ou une valeur de vérité : un même signe peut avoir des fonctions différentes selon le contexte de discours dans lequel il est employé, et son sens habituel peut devenir sa dénotation (ou référence) dans certains cas. Frege considère que l'usage *habituel* des signes a une fonction d'ordre déclaratif,

[19] "La désignation d'un objet singulier peut consister en plusieurs mots ou plusieurs signes. A des fins de brièveté, on appellera nom propre toute désignation de ce type" ([Frege 1971], p. 104).

c'est-à-dire visant à décrire des faits du monde. Mais la fonction des signes peut changer, lorsque l'on passe d'une fonction *référentielle* du discours à d'autres fonctions. L'auteur cite le cas de la paraphrase, ou citation indirecte :

> Si on emploie les mots de la manière habituelle, c'est de leur dénotation qu'on parle. Mais il se peut qu'on veuille parler des mots eux-mêmes ou de leur sens. Le premier cas est celui où, par exemple, on cite au style direct des paroles d'un autre. Les mots prononcés dénotent alors les mots d'autrui et ce sont ces derniers qui ont la dénotation habituelle. Nous avons ainsi affaire à des signes de signes (...) Pour parler bref, nous dirons qu'au style indirect, les mots sont employés *indirectement* ou encore que leur dénotation est *indirecte*. Nous distinguons dès lors la dénotation *habituelle* d'un mot de sa dénotation *indirecte*, et son sens *habituel* de son sens *indirect*. La dénotation indirecte d'un mot est son sens habituel.[20]

La solution de Frege à l'échec ci-dessus de la substitutivité consiste à traiter le contexte de croyance comme un contexte de discours indirect, c'est-à-dire comme un contexte inhabituel où le sens habituel devient la référence. Dans le contexte déclaratif ci-dessus, celui du nombre des joueurs de basket-ball, les *énoncés* 'que 2 + 3 = 5' et 'que 2 + 3 = le nombre des joueurs d'une équipe de basket-ball' dénotent la même référence (le vrai) mais expriment deux propositions distincte. Or dans un contexte épistémique, le sens des énoncés devient la référence et cette référence ne dénote plus leur valeur *de vérité* mais leur valeur *cognitive* (l'information qu'ils contiennent).[21] La valeur cognitive portée par un énoncé représente un des différents sens possibles d'une proposition, et ce sens exprime une pensée. Mais l'identité de valeur de vérité entre deux 'propositions' n'implique pas l'identité de leurs valeurs cognitives : on peut penser qu'une égalité arithmétique est vraie sans penser par là même qu'il faille être cinq pour composer une équipe de basket-ball. L'idée de Frege est donc que le contexte épistémique (croyance ou connaissance) est comparable à un discours indirect, dans lequel l'objet du discours n'est plus *ce* qui est dit mais le dire lui-même.

L'objection à la loi de substitutivité des identiques est donc annulée par le changement de référence qui s'est produit entre les deux contextes de discours, que l'on distinguerait ici en termes respectifs de contextes non-modaux et modaux : dans un contexte épistémique, la proposition de départ se transforme en un énoncé et la loi de substitutivité des identiques ne s'applique donc plus puisque l'objet de discours n'est plus alors le Vrai, mais la pensée. La modalité de croyance a donc pour effet de modifier le sens et la référence des termes de départ. Ce changement sauve en quelque sorte la loi logique de substitutivité des identiques, parce que les expressions co-référentielles dans un contexte non-modal (déclaratif) ne sont plus co-référentielles un contexte modal : la substitution n'a pas lieu d'être entre deux pensées distinctes. Seules des références élémentaires déterminent des références complexes, et seules des sens élémentaires déterminent des sens complexes ; le principe de compositionalité de Frege distingue ainsi en deux catégories le rôle d'un nom propre selon sa fonction dans un contexte :

---

[20] [Frege 1971], p. 105.

[21] La valeur cognitive de l'énoncé '2 + 3 = 5' est supposée nulle : c'est un énoncé *a priori* qui ne nous apprend rien sur le monde ; en revanche, la valeur cognitive de l'énoncé '2 + 3 = le nombre des joueurs de basket-ball' est non nulle puisqu'elle nous apprend quelque chose sur le sport. En termes de sémantique relationnelle, on dirait d'un énoncé sans valeur cognitive ou *a priori* qu'il est vrai dans *tous* les mondes possibles, contrairement aux énoncés empiriques ou *a posteriori* qui peuvent être faux dans certains mondes possibles.

> si l'on veut déterminer le sens de la proposition, seul entre en compte le sens de cette partie, nullement sa dénotation. La pensée demeure identique que le nom d''Ulysse' ait ou non une dénotation.[22]

Une fois que le sens et la dénotation (ou référence) d'un nom propre ont été définis, que le changement des fonctions sémantiques d'un nom propre selon le contexte a été expliqué, puis que le principe de compositionalité a été admis, Frege a donc bloqué l'inférence incriminée qui semblait conduire au rejet d'une de ses lois de l'être vrai. Pour corriger ce 'faux problème' due à une confusion sémiotique entre proposition et pensée, marquons en caractère gras la fonction *intentionnelle* (exprimer une pensée) des énoncés afin de montrer à proprement parler le lieu où se situe l'erreur de raisonnement. Dans le syllogisme suivant caractéristique du phénomène d'opacité référentielle, où **F(a)** symbolise la croyance que *a* est F et où (a = b) symbolise le fait que *a* est identique à *b* :

F(a)	1
(a = b)	1
∴ **F(b)**	?

la conclusion ne peut pas être inférée des deux prémisses puisque la substitution entre deux signes sémantiquement distincts, l'un dénotant la proposition **a** et l'autre la valeur de vérité *b*, n'est pas permise par la compositionalité. La comparaison des contextes de croyance aux contextes de citation sera reprise par Quine, mais celui-ci montrera par ailleurs qu'une telle comparaison a ses limites.

Si nous perdons de vue la proposition dans un contexte de croyance, cela implique que la valeur de vérité n'est plus l'*objet* visé dans ce type de contexte. Or la logique selon Frege porte uniquement sur les lois de l'être vrai et, en tant que telle, le type de valeur à laquelle elle s'intéresse est uniquement celui de la valeur de vérité, quel que puisse être son porteur. Faut-il en conclure que la logique ne *peut* pas parler de tout, ou qu'elle ne s'intéresse pas à la signification de n'importe quelle type de phrase ? La distinction vue plus haut entre proposition et énoncé, contenu jugeable et jugement, conséquence logique et inférence doit nous rappeler que la logique de Frege a une vocation scientifique : servir le but scientifique de poursuivre la vérité, à partir d'une définition des propriétés fondamentales du Vrai. Compte tenu de cette fonction spécifique, l'échec ci-dessus de la substitutivité des identiques est réglé en montrant que le passage entre les prémisses et la conclusion repose sur un usage incorrect de la substitution.

L'analyse fregéenne d'un nom propre en termes de sens et de référence permet donc de distinguer deux types d'information véhiculées par la proposition : sa valeur de vérité en rapport aux objets, d'un côté, sa valeur cognitive en rapport aux sujets, de l'autre. La valeur informative peut changer d'un énoncé à un autre, bien que ces deux énoncés sont considérés comme exprimant une même valeur de vérité : les deux termes généraux 'le nombre des

[22] [Frege 1971], p. 109.

joueurs d'une équipe de basket-ball' et '5' ont même référence, mais le premier est exprimé par une proposition susceptible d'apprendre quelque chose à l'interlocuteur. Encore une fois, Frege considère que cette valeur cognitive n'est pas le souci de la logique :

> la langue scientifique, comme le *Begriffschrift* de Frege, ne doit contenir que des énoncés qui sont interprétables sans allusion à ce que Frege appelle 'le jeu réglé du locuteur et de l'auditeur'. Frege procédait à cette exclusion des facteurs *pragmatiques* et *contextuels* parce qu'il voulait faire une théorie de la preuve mathématique.[23]

Ce problème de compétence linguistique des locuteurs est un problème de *compréhension*, et la logique ne concerne que la *signification* des propositions. Ces significations transcendent leurs porteurs, elles sont déterminées par des règles de structure qui sont supposées 'refléter' la structure des pensées.

La restriction imposée par Frege sur l'analyse logique du langage signifie que les facteurs pragmatiques de compétence linguistique sont sans effet sur les lois logiques ; mais cela ne réduit pas la question de la compétence à une question psychologique, puisque la représentation et le sens sont deux choses clairement distinctes :

> La dénotation d'un nom propre est l'objet même que nous désignons par ce nom ; la représentation que nous y joignons est entièrement subjective ; *entre les deux* gît le sens, qui n'est pas subjectif comme l'est la représentation, mais qui n'est pas non plus l'objet lui-même. La comparaison suivante éclairera peut-être ces rapports. On peut observer la lune au moyen d'un télescope. Je compare la lune elle-même à la dénotation; c'est l'objet de l'observation dont dépendent l'image réelle produite dans la lunette par l'objectif et l'image rétinienne de l'observateur. Je compare la première image au sens, et la seconde à la représentation ou intuition. L'image dans la lunette est partielle sans doute, elle dépend du point de vue de l'observation, mais elle est *objective* dans la mesure où elle est offerte à plusieurs observateurs.[24]

D'autre part, le sens est public puisqu'il peut être transmis de locuteur à locuteur et augmenter l'information que possède un sujet sur les propriétés du monde qui l'entoure. La logique épistémique donnera l'occasion de confirmer cette distinction entre représentation subjective et sens objectif, dans le cadre du problème de l'individuation des croyances. Mais surtout, nous examinerons si cette logique enrichie (par des modalités) est en mesure d'enrichir la théorie de la signification de Frege et, notamment, de préciser les conditions dans lesquelles une même pensée peut être saisie par plusieurs locuteurs. D'après Frege, le thème du sens n'est pas un thème scientifique puisque la reconnaissance du lien entre un mode de dénotation et sa dénotation est une affaire de compétence personnelle :

> Le sens d'un nom propre est donné à quiconque connaît suffisamment la langue ou l'ensemble des désignations dont il fait partie.[25]

---

[23] [Gochet 1978], p. 201 ; nous soulignons.
[24] [Frege 1971], p. 104; nous soulignons.
[25] [Frege 1971], *ibid.* Cette contingence est une raison pour Frege de dissocier de son idéographie tout ce qui a trait aux croyances : "Tant que la dénotation demeure la même, ces fluctuations de sens sont supportables, encore qu'elles doivent être évitées dans l'exposé systématique d'une science démonstrative, et rejetée d'un langage parfait". (*ibid.*, note 1).

Nous avons ici deux problèmes distincts, que nous examinerons plus loin dans le cadre de la logique modale épistémique :
- un problème *vertical* (relation langage-monde), qui concerne la relation de dénotation entre les signes du langage et leurs objets ; cette relation est liée au thème de l'opacité référentielle et concerne l'identité des extensions dans un espace épistémique ;
- un problème *horizontal* (relation langage-langage), qui concerne la structure des pensées et les règles déterminant cette structure dans le langage; cette relation est liée au thème de la rationalité et concerne la définition des constantes logiques dans des 'mondes épistémiquement possibles'.
Les deux thèmes adressent respectivement deux questions au logicien épistémique :
- comment individuer des croyances en cas d'opacité référentielle, c'est-à-dire comment reconnaître que deux croyances sont les mêmes ? L'identité de leur objet devrait apporter la solution, qui exige que la forme logique des attitudes épistémiques soit précisée. Avec quoi le sujet est-il en relation lors d'une croyance : une proposition, un énoncé, ou autre chose ?
- l'étude de la rationalité limitée des agents exige-t-elle une redéfinition de l'être vrai, autrement dit : faut-il réviser la définition des constantes logiques pour concevoir une logique modale épistémique ?

2.1.2.2 Russell

Russell prétend qu'un examen des croyances est indispensable pour définir nos raisonnements quotidiens et comprendre ce que les philosophes entendent par la notion de vérité :

> La vie intellectuelle globale se compose de croyances, ainsi que du passage de la croyance à une autre que l'on appelle 'raisonnement'. Les croyances apportent la connaissance et l'erreur ; elles sont les véhicules de la vérité et de la fausseté. La psychologie, la théorie de la connaissance et la métaphysique gravitent autour de la croyance, et notre approche philosophique dépend en large partie de l'idée que nous nous faisons de la croyance.[26]

Cela étant, l'auteur considère qu'une étude de ces croyances n'a aucun rapport avec la logique, laquelle concerne uniquement le vrai et le faux. En d'autres termes, Russell associe croyance et psychologie tout en réservant le domaine de la logique au thème de la *proposition*, vraie ou fausse par définition. Une certaine théorie de la vérité sous-tend son rejet primordial d'une logique épistémique ; si la croyance a un intérêt philosophique, c'est parce que Russell se sert de son examen pour discerner la théorie de la vérité que défendent plusieurs philosophes : la théorie de la *vérité-cohérence* est une cible principale du Britannique, qui défend pour sa part une théorie de la vérité-correspondance en caractérisant la proposition par sa relation de correspondance ou de non-correspondance avec un fait.

La suite va tenter de montrer que la définition de la croyance telle qu'on la trouve en logique épistémique est inapplicable, voire inconcevable au sein de la philosophie de la logique de Russell ; pour ce faire, considérons tout d'abord ce que l'auteur entend par les

[26] B. Russell : *Analysis of Mind*, London (1921), p. 231. Cité in [Rescher 1960], p. ?.

notions de *vérité* et de *fait*. C'est à partir de ces deux notions basiques qu'il définit les notions consécutives de proposition et de croyance.

L'analogie structurelle entre modalités et quantificateurs a été évoquée précédemment pour jeter les bases d'une logique des attitudes. Or il semble que Russell n'était pas sensible à cet argument formel, probablement parce que l'identité de structures ne suffisait pas à ses yeux pour parler de 'vérités logiques'. La vérité logique présuppose la notion de vérité, mais celle-ci renvoie à une correspondance avec des faits et non à une simple similarité entre des formes d'énoncés. Le terme de vérité n'est pas un marque-place ou une valeur simplement algébrique, ici, et un symbole doit être dans une certaine relation avec un 'fait' pour être qualifiable de vrai. On retrouve dans les écrits de Russell la même critique que celle aperçue plus tôt chez Frege, concernant les modalités aléthiques ; en effet, Russell assimile plus ou moins la 'nécessité' d'une proposition au sentiment de certitude qu'un énoncé inspire chez le locuteur, et l''impossibilité' au sentiment de certitude de sa fausseté. Or ces deux sentiments opposés décrivent un degré optimal de croyance qui, comme nous allons le voir, est strictement dissocié de la logique selon Russell :

> Il n'y a pas de notion logique fondamentale de *nécessité*, ni par conséquent de *possibilité*. Si cette conclusion est valide, le thème de la modalité devrait être banni de la logique, puisque les propositions sont *simplement* vraies et fausses.[27]

Les critiques de Russell contre les notions de nécessité et la possibilité visaient principalement le 'raisonnement symbolique' de Hugh MacColl, l'un des promoteurs de la logique modale interprétée avec Peirce. Selon MacColl, il n'y a pas deux mais cinq classes de 'propositions' [*statements*] distinctes : une proposition est vraie, fausse, variable (parfois vraie et parfois fausse), certaine ou impossible. Ces propriétés sémantiques ne s'excluent pas mutuellement : une proposition est *impossible* lorsque sa contradictoire est *certaine* ; si une proposition est *certaine*, alors elle est *vraie* également ; si elle est *impossible*, alors elle est *fausse* également, etc. Une telle distinction sémantique préfigure la distinction actuelle des modalités formelles : $\Box\phi$, ϕ, $\neg\phi$, $\nabla\phi = (\Diamond\phi \wedge \Diamond\neg\phi)$, et $\Box\neg\phi$ symbolisent respectivement le fait pour une proposition ϕ d'être nécessaire, vraie, fausse, variable et impossible. Or Russell reproche à cette distinction de reposer sur une confusion entre deux aspects de la proposition : son expression verbale, que nous appelons l'*énoncé*, et le contenu même de cette expression, c'est-à-dire le contenu propositionnel ou *proposition* au sens propre (la pensée fregéenne). Si l'énoncé 'la porte est ouverte' est variable plutôt que vrai, c'est parce que sa valeur de vérité peut changer selon le contexte de discours et le locuteur qui le prononce; par opposition, un énoncé tel que '2 + 2 = 4' est nécessaire parce qu'il est vrai dans n'importe quel contexte d'énonciation et pour n'importe quel locuteur, et '$2 + 2 \neq 4$' est donc impossible. La réponse de Russell à cette interprétation modale des 'propositions' (ou pseudo-propositions) comporte deux arguments.[28] Premièrement, l'interprétation modale porte sur des expressions *incomplètes*, c'est-à-dire des énoncés qui ne restituent pas clairement le contenu de la proposition qu'il vise à exprimer ; dans l'expression 'la porte est ouverte', cet énoncé exprime

[27] [Russell 1905b], p. 520 ; nous soulignons.
[28] La critique de Russell se trouve dans son article "Review of 'Symbolic Logic and its Applications'", *Mind* **15**(58) (avril 1906), pp. 255-260.

une pensée incomplète et n'est complétée que par une formule plus précise de type 'la porte est fermée à l'endroit x au temps t'. Dans ce cas, l'énoncé enrichi exprime une pensée complète et n'est plus variable mais soit vrai, soit faux. Deuxièmement, l'interprétation modale provient selon Russell d'une confusion entre une *proposition* et une *fonction propositionnelle*. Une expression telle que 'x est un avocat' n'est pas nécessaire parce qu'elle n'est pas vraie pour tout x, mais elle est variable parce que vraie pour certains x et fausse pour d'autres ; une expression telle que 'la porte est fermée au temps x' est variable également, parce qu'elle est vraie pour certains x mais pas pour tout x. A partir de cette analyse des modalités de MacColl, Russell affirme sa position classique en maintenant une distinction entre l'expression d'une proposition et le contenu d'une proposition : parce que la logique des propositions s'intéresse uniquement au contenu de ces propositions, lesquelles doivent être strictement distinguées de leur moyen d'expression qu'est l'énoncé, une proposition est soit vraie soit fausse et le mode de vérité ne concerne pas la logique en tant que telle. Et puisque la certitude représente la version épistémique de la nécessité (est certain ce qu'un sujet considère comme nécessairement vrai), une logique épistémique est légitime pour MacColl mais ne peut l'être pour Russell.

Russell a exposé notamment sa définition de la croyance dans [Russell 1923a] et [Russell 1923b], où deux questions cruciales sont posées : qu'est-ce qui distingue les attitudes propositionnelles (fonctions de signification) des fonctions de vérité, d'une part ; que faut-il comprendre par l'expression 'A croit p', d'autre part. Dans l'Introduction, nous avons déjà signalé que Russell avait bien conscience des difficultés créées en logique par les attitudes propositionnelles : ce ne pas sont des fonctions de vérité, puisque la croyance à une proposition ϕ (symbolisée par $f\phi$) n'implique pas la croyance à toutes les propositions ψ équivalentes à ψ. Si l'on définit la relation d'équivalence comme une relation d'identité de valeurs de vérité : $(\phi \equiv \psi) \Leftrightarrow (f\phi \equiv f\psi)$, celle-ci ne permet donc pas d'identifier des croyances, d'où une application restreinte de la loi de vérifonctionnalité. Russell en conclut qu'une fonction linguistique n'est pas toujours une fonction de vérité : toute fonction de vérité est une fonction de signification, puisqu'elle contribue à la signification des propositions, mais la réciproque n'est pas toujours vraie. La question est de savoir ce qui distingue le fait d'être vrai du fait d'être significatif, ou pourvu de sens. Pour expliquer l'irréductibilité des attitudes propositionnelles à des fonctions de vérité, Russell déclare que la proposition ne remplit pas un rôle *logique* dans une attitude :

> il y a quelque chose à dire en faveur de l'idée selon laquelle les propositions peuvent se présenter de deux manières, (a) comme des assertions, (b) comme des faits, et selon laquelle les déclarations [*statements*] dans lesquelles elles se présentent comme des assertions sont toujours des fonctions de vérité, bien que celles dans lesquelles elles se présentent comme des faits peuvent ne pas l'être. Une telle théorie exigera une analyse logique de ce que l'on entend par 'déclaration portant sur un fait'; et aussi de ce que l'on entend, *en logique*, par 'assertion'.[29]

[29] [Russell 1923a], p. 156. Russell distingue ici l'assertion de la déclaration [*statement*], c'est-à-dire la 'pseudo-proposition' de MacColl, parce qu'une déclaration peut être fausse alors qu'il conçoit l'assertion comme une déclaration vraie par définition. Le même problème s'était posé dans la section précédente consacrée à Frege (voir note 13), au sujet de la notion de jugement comme 'reconnaissance de la vérité'. Lorsqu'on la considère comme un simple acte de discours, l'assertion peut être assimilée au jugement de Frege et traitée comme synonyme de la simple déclaration, en revanche.

Que faut-il entendre par un 'fait', et pourquoi une proposition n'est-elle plus une fonction de vérité lorsque l'on s'intéresse à elle en tant que fait en non en tant que qu'assertion ? Tout d'abord, Russell entend par 'fait' la réalisation d'un événement quelconque. Une croyance est un fait, un fait psychologique. Une déclaration, c'est-à-dire la production d'un énoncé, est un fait, un fait linguistique.

> Qu'entend-on par 'A croit p'? Il y a une manifestation dans A, appelée 'croyance'. C'est un fait comme un autre.[30]

Mais lorsqu'une fonction de signification est une fonction *de vérité*, la proposition à laquelle est attribuée une valeur vérité ne représente pas un fait quelconque. Dans l'énoncé 'Platon est Grec', par exemple, l'individu nommé Platon est un constituant de la proposition que Platon est Grec, et le fait que Platon est Grec (la pensée fregéenne) est exprimé par cet énoncé ; mais deux types de faits distincts sont produits lors d'une énonciation : le fait contenu dans l'énoncé, c'est-à-dire la proposition fregéenne, et le fait d'énoncer ce contenu. Le second fait, le fait d'avoir exprimé la proposition que Platon est Grec par le biais de l'énoncé 'Platon est Grec', est d'ordre linguistique et concerne notre discours sur les choses ; le premier fait, le fait que Platon est Grec, est d'ordre ontologique et concerne les choses elles-mêmes. Russell explique ainsi que, dans une fonction de vérité, la proposition n'est pas considérée comme le *fait linguistique* qu'est l'acte d'énoncer. Elle n'est pas considérée non plus comme le *symbole* du fait qui rend la proposition vraie ou fausse, car réduire la proposition à un symbole impliquerait que deux expressions issues de deux langues naturelles distinctes ne représentent pas la même proposition : si 'Platon est Grec' et 'Plato is Greek' n'exprimaient pas la même proposition, alors ces expressions ne seraient pas vraies ou fausses conjointement. Et puisque, dans une fonction *de vérité* $f\phi$, la proposition ϕ est le contenu propositionnel de l'énoncé, à savoir *ce* qui est rendu vrai ou faux par la réalisation d'un fait ontologique, la proposition ϕ ne dépend pas du langage dans lequel elle est exprimée mais est un constituant du fait qui rend $f\phi$ vraie ou fausse. Seules les propositions sont des fonctions de vérité, autrement dit, et les énoncés qui les expriment sont des fonctions de signification de ces propositions.

Contrairement aux fonctions de vérité, dont les constituants sont en relation avec *ce* qui les rend vrai ou faux (les états de choses ou faits ontologiques), la croyance porte sur un fait psychologique parce que ses constituants ne sont pas définis en termes de ce qui les rend vrais ou faux. L'ambiguïté vient de la notion de fait, qui peut être d'ordre social, psychologique, linguistique, ou ontologique ; un fait peut être logique ou psychologique, selon Russell, mais seul le fait exprimant une relation de correspondance avec la réalité peut être qualifié de vrai ou de faux. La croyance n'est pas évacuée pour autant de l'analyse des fonctions de vérité et de leurs propositions, mais elle correspond à ce que Russell appelle une *fonction de signification* et n'équivaut pas à une fonction de vérité : une croyance est un fait psychologique qui porte sur un contenu propositionnel vrai ou faux, mais la vérité ou fausseté de ce contenu (de la proposition fregéenne) ne dépend pas de la croyance, qui est un fait

[30] [Russell 1923b], p. 159.

strictement psychologique. Une croyance se manifeste dans des circonstances diverses : si le sujet croyant a une raison de considérer quelque chose comme vrai, s'il dispose d'une preuve en faveur de cette vérité, etc. Mais on voit bien ici que les circonstances d'une croyance ne déterminent pas elles-mêmes ce qui est vrai ou faux dans le monde. Pour cette raison, Russell déclare que

> la logique ignore par convention ces circonstances, toutefois ; elle suppose que, pour toute croyance donnée, il y a un certain fait défini qui la rend vraie (ou fausse).[31]

A partir de cette distinction entre croyance et vérité, on peut redéfinir la notion de proposition et compléter ce que nous avions dit plus tôt au sujet de la relation entre proposition et énoncé : la proposition désigne la classe de toutes les croyances ou leurs expressions linguistiques qui sont rendues vraies ou fausses par un fait.

> Une proposition est donc une classe de faits, psychologiques ou linguistiques, définis comme ceux qui ont une certaine relation (qui peut être assertion ou négation [*denial*], selon les cas) à un certain fait.[32]

Cette définition de la proposition évite le recours aux entités abstraites que l'on trouvait chez Frege, les pensées, et caractérise la proposition comme une classe de *faits linguistiques* rendus vrais par un même fait ontologique, ou état de choses ; et puisque ces faits linguistiques expriment la totalité des croyances qui caractérisent une proposition, ce sont donc ce que nous appelons des énoncés ou phrases déclaratives. Lors de l'expression d'une croyance par le biais d'un énoncé, une pensée commune est supposée 'saisie' par plusieurs sujets croyants et se manifeste sous la forme linguistique d'un ou plusieurs énoncés ; mais l'entité idéale que représente la proposition n'est définie chez Russell que de manière indirecte, par le biais de la classe de tous les énoncés qui l'expriment. Ce n'est donc pas la proposition en tant que telle qui est exprimée dans une croyance mais seulement un membre de cette proposition, et ce membre qu'est l'énoncé appartient à une classe de croyances (ou de leurs expressions symboliques). On retrouve ici la même distinction que celle aperçue chez Frege entre une pensée et l'entité linguistique qui sert à l'exprimer; dans les deux cas, la pensée est une entité extra-linguistique et s'exprime par le biais d'un énoncé.

Récapitulons : une croyance est un fait psychologique ; la vérité est une relation entre une proposition et un fait ontologique, indépendant du sujet croyant. Les croyances sont exprimées par le biais d'expressions symboliques, et celles-ci expriment une proposition. Plusieurs énoncés sont reliés à une seule et même proposition lorsqu'ils parlent du même fait ontologique, celui qui les rend tous vrais ou faux. Lorsque l'*énonciation*, c'est-à-dire l'acte d'exprimer une proposition par le biais d'un énoncé, est considérée comme un fait à part

---

[31] [Russell 1923a], p. 157.

[32] *Ibid.* Nous traduisons 'denial' par négation, puisque Russell parle ici de la déclaration fausse. Mais de même que pour la notion d'assertion, 'denial' a également un autre sens (pragmatique) de rejet ou dénégation, qui correspond alors au refus d'une déclaration. La différence entre les deux sens réside dans l'engagement ou non du locuteur sur la valeur de vérité de sa déclaration. L'*attitude* de dénégation sera revue en détail dans le dernier chapitre (section 3.3.2).

entière, la vérité ou fausseté n'est plus en question et la logique n'est donc plus concernée puisque celle-ci s'intéresse uniquement aux conditions de vérité et de fausseté des propositions. La distinction entre attitudes épistémiques et fonctions de vérité réside dans une distinction de la nature des faits qui les concernent, psychologiques ou linguistique pour les premières, ontologiques pour les secondes :

> Une fonction de vérité de p est rendue vraie (ou fausse) par le même fait qui rend p vraie (ou fausse); en vertu de la définition ci-dessus de la proposition, c'est donc la même proposition que p, ou la même proposition que non-p. De même que deux fonctions de vérité de plusieurs arguments sont identiques si elles sont équivalentes (...) Lorsqu'une croyance est vraie, si *f* est le fait qui la rend vraie nous disons que la croyance a la relation A (assertion) à *f*. Lorsqu'une croyance est fausse, appelons N (négation) la relation.[33]

Comme indiqué plus haut, l'analyse de la proposition est d'ordre logique lorsqu'elle est considérée du point de vue de son assertion : c'est le cas qu'une proposition est rendue vraie par un fait ; elle n'est plus d'ordre logique lorsque la proposition est considérée du point de vue de son expression : c'est le cas qu'un énoncé est produit pour exprimer cette proposition. Dans ce dernier cas, on parle d'une proposition au sens de la proposition grammaticale ou énoncé, qui est l'expression d'une croyance. Mais ce sens courant de la 'proposition' n'est pas celui qui intéresse Frege et Russell : lorsque nous parlons ici de 'proposition', nous faisons allusion uniquement à la proposition logique, c'est-à-dire au contenu d'une proposition grammaticale.[34] A cet égard, l'erreur du logicien épistémique serait de confondre les deux types de cas et de parler indifféremment de vérité lorsque l'on dit que la manifestation d'une croyance 'correspond' à un fait (psychologique).

Pour définir la forme d'une proposition et la forme d'une croyance α, [Russell 1923a] utilise un symbolisme ensembliste constitué du fait *f*, de l'assertion $A^{\rightarrow}$ (rendre vrai) et de la négation $N^{\rightarrow}$ (rendre faux). On obtient les explications suivantes :

Proposition

$\phi =_{df} A^{\rightarrow}`f$ ou $N^{\rightarrow}`f$, pour un certain *f*

Croyance

Jones croit $\phi =_{df} \exists!\alpha \cap A^{\rightarrow}`f$, si $\phi = A^{\rightarrow}`f$

$\exists!\alpha \cap N^{\rightarrow}`f$, si $\phi = N^{\rightarrow}`f$

Dans [Russell 1923b], l'auteur précise sa définition de la proposition en termes de classes de croyances (et de leurs expressions linguistiques) α rendues vraies ou fausses par un fait, remplaçant les prédicats d'assertion A et de négation N par ceux de vérification V et de

[33] [Russell 1923a], p. 158. L'assertion telle que Russell l'entend présuppose donc bien la vérité de la déclaration, contrairement à la notion de jugement chez Frege : un jugement peut être faux. Si l'on reprend la distinction de [Hanson 1952] entre déclarations de fait (toujours vraies) et déclarations factuelles (falsifiables), l'assertion de Russell s'exprime par le biais d'une déclaration de fait.

[34] Lorsque l'on parle d'un 'contenu propositionnel', l'expression 'propositionnel' renvoie à la proposition au sens grammatical du terme, et le contenu propositionnel correspond à la proposition au sens logique du terme.

falsification F. Quoi qu'il en soit, la proposition est dans les deux cas une classe de croyances dont l'appartenance au même ensemble est déterminée par un fait ontologique commun.

Pour revenir sur une question précédente : comment établir l'identité de deux croyances, notons que la réponse de Russell ne portera évidemment pas sur ce qui les distingue en tant que faits psychologiques ou linguistiques : le 'fait propositionnel' ne compte pas, puisque l'identité est liée à la notion de vérité et se rapporte ainsi au fait qui rend vraie une proposition. Deux croyances sont identiques si leur expression linguistique porte sur la même proposition, donc si elles appartiennent toutes deux à la classe des énoncés qui sont rendus vrais ou faux par le même fait.[35] Il y a une conséquence essentielle dans cette organisation entre les notions de logique et de psychologie, distinguées par le critère de vérité : tout fait n'est pas vrai, et seul un certain type de fait peut être qualifié ainsi. C'est le *fait logique*, et ce fait est indissociable du fait ontologique dans la mesure où il porte sur la relation de correspondance ou non-correspondance entre une proposition, exprimée par une classe d'énoncés, et *ce que* cette proposition dénote. Le fait que Pierre croit que Platon est Grec n'est pas 'vrai' en tant que tel : le fait que Pierre produise ici ce qui s'avère être une assertion (une proposition vraie) est un fait linguistique portant sur un énoncé mais pas sur *ce* qui rend cet énoncé vrai ou faux ; pour cette raison, le fait d'énoncer quelque chose de vrai (de l'asserter) est psychologique et n'a pas de valeur de vérité.

La distinction établie par Russell entre un fait primaire (logique) et un fait secondaire (non-logique) préfigure en un sens la clause d'introspection du système modal **S4**, présenté plus tôt. Pour distinguer les deux types de faits, l'auteur recourt en effet à une forme d'itération :

> Ainsi la proposition qui apparaît lorsque nous *assertons* une proposition est tout à fait différente de celle qui apparaît lorsque nous *assertons que nous assertons* une proposition ; celle-ci est un fait psychologique ou linguistique, celle-là une classe de ces faits.[36]

Mais il ne faut pas confondre la clause d'itération de Hintikka et l'itération des assertions décrite ci-dessus : l'assertion correspond chez Russell à ce que lui-même entend par une assertion logique : ce n'est pas un *acte* d'assertion, comparable à ce que décrit la barre de jugement de Frege, mais un *fait* logique relatif à la correspondance entre une proposition et un fait qui la rend vraie. Asserter que l'on asserte quelque chose décrit en revanche un acte commis par le locuteur : dans ce cas, Pierre *dit* que l'énoncé qu'il exprime se rapporte à une proposition vraie, et la croyance devient alors un constituant du fait *psychologique* qu'est l'énonciation. L'itération de la croyance change ainsi le statut de l'assertion qui, lorsqu'elle est redoublée, ne porte plus sur un fait logique mais sur un fait psychologique. On retrouve ici l'idée du changement de référence de Frege dans un contexte de discours indirect, lorsque l'énoncé n'est plus considéré pour sa valeur de vérité mais pour sa valeur cognitive, ou informative. Russell distingue clairement cette valeur du domaine des faits logiques, limités pour leur part à l'étude de la relation entre langage et monde.

---

[35] L'expression 'croyance vraie' signifie de manière elliptique que son énoncé est rendu vrai par le fait qu'il exprime. La vérité est "une qualité entièrement dépendante du rapport existant entre la croyance et les objets extérieurs" (B. Russell, *Problèmes de Philosophie*, "Le vrai et le faux", Petite Bibliothèque Payot, 17e éd., 1965, p. 144).

[36] [Russell 1923a], p. 158. Nous soulignons.

L'analyse logique de Russell rend donc absurde la notion même de 'logique épistémique', à moins de confondre les notions d'être vrai et d'être le cas. Un fait psychologique peut être le cas, mais un seul type de cas intéresse le Russell logicien, dont la distinction entre différents ordres de faits interdit toute confusion entre plusieurs genres de vérité. Cette acception précise de la notion de vérité a au moins deux conséquences pour la sémantique modale. Revenons dans l'ordre sur ce qui distingue une proposition et une attitude propositionnelle, puis sur la notion de vérité selon Russell.

Propositions et attitudes propositionnelles :

L'explication de Russell semble confirmer ce que nous avons dit précédemment, au sujet de la définition nominaliste des ensembles épistémiques : lorsque la logique modale épistémique définit une croyance comme un ensemble de *formules*, ces formules représentent dans un langage formel des *énoncés* des langues naturelles et non des *propositions*. Dans la mesure où ces énoncés expriment un ensemble de croyances communes au sein d'un langage, ce n'est donc pas un ensemble de 'propositions' qui définit la croyance; bien au contraire :

> Lorsque nous considérons une croyance, la proposition au sens qui vient d'être défini ne se présente pas *dans notre croyance* ; ce qui se présente, c'est un membre de la classe qu'est la proposition. Même chose lorsque nous utilisons des mots pour asserter une proposition.[37]

La définition russellienne de la proposition ressemble en un sens à la définition de l'opérateur modal épistémique : de même que la proposition a été interprétée plus haut comme une *classe* de faits psychologiques (de croyances) ou linguistiques (les énoncés qui les expriment), la croyance de ϕ est interprétée en logique modale épistémique comme la *classe* des mondes possibles où ϕ est vraie. Dans les deux cas, l'expression à définir est présentée en termes de classes. Mais contrairement à la croyance de Russell, où ce fait psychologique est corrélé à un fait linguistique qui est son énoncé, la croyance se définit en termes de classe de *formules*, chez Hintikka. D'une part, cette classe de formules n'est pas déterminée par le fait qui les rend vraies. On parle de classes aussi bien chez Russell que chez Hintikka, mais le contenu de ces classes est donc distinct. D'autre part, la réduction nominaliste des mondes possibles à des ensembles modèles crée une autre difficulté : une croyance porte sur une formule, mais cette formule représente-t-elle une proposition, un énoncé, ou autre chose ? Dans l'esprit de Russell (et Frege), une croyance est exprimée par un énoncé mais désigne une proposition ; or si, dans une formule modale de type $B\phi$, l'objet ϕ de la croyance B est exprimé par un énoncé, cet énoncé n'est pas vrai ou faux à proprement parler mais exprime quelque chose, une proposition, qui est rendue vraie ou fausse par un fait (ontologique). Un énoncé n'est donc vrai ou faux que par l'entremise de la proposition qu'il exprime, et une croyance vraie signifie uniquement une croyance dont le contenu ou idée[38] exprimé par un

[37] [Russell 1923a], p. 157. Nous soulignons.

[38] [Russell 1923b] suggère la notion d'idée pour exprimer le contenu d'une croyance, qui n'est pas une proposition : "Prenez 'A croit que B a tué C'. Cela dit que dans A il y a des manifestations appelées 'idées' de B, de tuer et de C, liées d'une certaine manière. Ici la proposition n'apparaît pas, et 'A croit p' n'est pas une fonction de p" (p. 159). On notera que, dans la formulation anglaise de l'attitude propositionnelle, il est écrit 'A croit p' et non 'A croit *que* p' : l'absence de la forme complétive (*that-clause*) est une indication grammaticale

énoncé correspond à un fait. Une lecture de la sémantique relationnelle en termes russelliens conduirait donc à une impasse : si les constituants d'un ensemble modèle sont des énoncés, on ne peut pas les qualifier eux-mêmes de vrais ou faux sans commettre une confusion des genres ; si ces constituants sont des propositions, il n'y a donc plus seulement un mais plusieurs faits ontologiques à prendre en compte pour évaluer une seule et même croyance. On aboutit à une absurdité dans les deux cas, et ce pour une raison claire : chez Russell, la croyance sert à définir la proposition mais n'est pas elle-même définissable ; une croyance correspond tout au plus à un énoncé et une classe de croyances constitue une proposition, mais jamais la croyance n'apparaît chez Russell comme un terme à définir, contrairement à la logique modale épistémique. Les classes dont parle Russell et Hintikka ne sont pas de même nature : elles permettent d'identifier une *proposition* chez Russell et une *attitude propositionnelle* chez Hintikka, de sorte que le point de vue que l'un examine est strictement différent du point de vue de l'autre. Dans une attitude propositionnelle, la classe à laquelle il est fait allusion détermine l'ensemble des formules que le *sujet* est disposé à croire ; dans une proposition, la classe en question détermine l'ensemble des croyances qui constituent une proposition *objective*. En d'autres termes, Russell et Hintikka s'expriment en termes ensemblistes mais sur des sujet distincts : le premier s'occupe du contenu des attitudes propositionnelles, le second des attitudes elles-mêmes. Mais aucune circularité ne ressort en réalité des définitions de l'attitude propositionnelle et de la proposition. Certes, la proposition se définit comme un ensemble d'attitudes de croyance chez Russell ; mais l'attitude de croyance ne se définit pas comme un ensemble de propositions chez Hintikka. Certes, à une croyance correspond un énoncé chez Russell tandis qu'à une croyance correspondent plusieurs formules (des énoncés, selon nous), chez Hintikka; mais la relation entre croyance et énoncé n'est ambiguë que si l'on confond les points de vue des deux auteurs : une croyance correspond à *un* énoncé lorsqu'il s'agit de définir la proposition, et une croyance correspond à *plusieurs* énoncés lorsqu'il s'agit de définir l'attitude d'un sujet à l'égard d'une proposition.

La notion de vérité selon Russell :

Que veut-on dire lorsque l'on parle de la 'vérité' d'une croyance : qu'une croyance a été effectivement tenue par quelqu'un, ou que le contenu de cette croyance est vrai ? La distinction entre ces deux applications de la notion de vérité correspond à la distinction entre la vérité d'une attitude propositionnelle et la vérité de la proposition, c'est-à-dire du contenu de l'attitude. La distinction marquée par Russell entre logique et psychologie consiste à restreindre l'application de la notion de vérité au second cas : seules les propositions sont vraies, et une croyance n'est vraie que par interposé (c'est son contenu qui est vrai ou faux).
Mais la situation se complique lorsque l'on entre en logique modale et que l'on aborde la notion de vérité du point de vue de la théorie des modèles. Etant donné son caractère *nominaliste*, la sémantique relationnelle de Hintikka n'implique évidemment pas l'existence de faits (ontologiques) multiples. Mais dans le même temps, Hintikka insiste à dire que sa sémantique doit exprimer un lien fondamental entre le langage et la *réalité*. Contrairement à notre recommandation ci-dessus, consistant à ne parler de vrai et de faux que relativement à

---

du fait que le contenu de la croyance n'est pas une proposition, signalée en général par la forme : 'que p'. D'où une confusion possible entre nos expressions courantes et la forme symbolique appropriée de la croyance.

des états de choses (des faits ontologiques), la vérité telle que l'entend le logicien modal peut être comprise comme un simple fait d'*être le cas*, sans précision sur le type de cas en vue (social, psychologique, linguistique, ontologique) : la vérité signifiera tantôt l'appartenance d'une formule à un modèle, tantôt la correspondance d'une proposition à ce modèle singulier qu'est le monde réel (le seul conçu par Russell). La première forme de vérité désigne une sorte de vérité-cohérence, par opposition à la vérité-correspondance de Russell. Cette différence d'usage de la notion de vérité peut expliquer en grande partie les réticences vis-à-vis de la logique modale, lorsque l'usage de la notion de vérité devient plus libéral. La distinction entre attitudes propositionnelles et propositions est sans réserve, chez Russell : dans un énoncé de croyance,

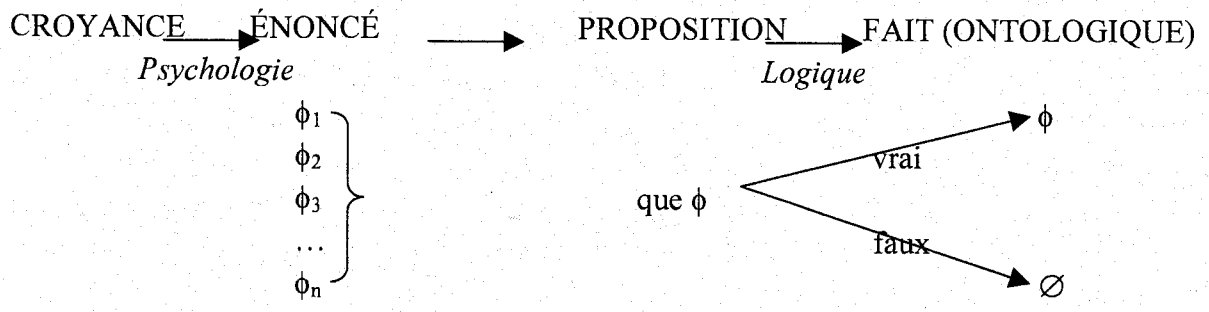
> la référence au fait, *lequel confère la vérité ou la fausseté*, est hors de propos. Une croyance (et son expression linguistique) n'est intéressante pour la logique qu'en qualité de véhicule de la vérité ou de la fausseté. Lorsque j'asserte que Socrate est mortel, la manifestation est différente de ce qu'elle est lorsqu'un autre individu asserte la 'même' proposition; mais cette différence est *sans importance pour la logique*, qui ignore tout ce qui concerne l'assertion à l'exception de ce qu'elle a en commun avec d'autres assertions de la 'même' proposition.[39]

Pour distinguer ce qui est vrai de ce qui ne peut pas l'être, le symbolisme doit donc être défini en plusieurs éléments distincts. Dans un ensemble de formules indistinctes, chacune de ces formules est distinguée selon sa fonction. On entend par *phrase* l'inscription matérielle d'une suite de signes ; deux phrases distinctes peuvent renvoyer à un même énoncé, et une phrase est identifiée par sa production dans un cas particulier. Par exemple : la phrase-là 'Platon est Grec' et cette phrase-ci, 'Platon est Grec' sont deux phrases (occurrences matérielles) distinctes mais produisent un même énoncé. Quant à l'*énoncé*, nous avons dit depuis le début qu'elle correspond à une phrase déclarative exprimant une certaine signification dans une certaine langue ; la signification de l'énoncé se définit en termes de vérité et de fausseté, mais un énoncé appartient à une langue naturelle particulière. La vérité ou fausseté de deux énoncés distincts dépend enfin de la proposition qu'ils expriment tous les deux, mais par le biais d'expressions différentes. Deux énoncés peuvent renvoyer à la même *proposition*, mais ils se distinguent selon qu'ils appartiennent à des langues différentes ou prennent des expressions différentes dans une même langue. Deux exemples : 'Platon est Grec' est un énoncé de la langue française et 'Platon is Greek' est un énoncé de la langue anglaise, mais ils expriment la même proposition. Partant de la proposition comme signification hypostasiée et commune à plusieurs énoncés, on peut reformuler la notion de la proposition en disant qu'elle correspond chez Russell à une classe d'*énoncés*, lesquels sont rendus vrais ou faux par un même fait ontologique. Par extension, une fonction de vérité $f\phi$ est une fonction dans laquelle la formule ϕ symbolise une proposition, et non un énoncé : la valeur sémantique de $f\phi$ ne change pas selon l'expression choisie pour l'énoncé de ϕ. Quant à l'*assertion*, le signe de la barre de jugement de Frege indique le cas où l'assertion est elle-même prise pour un fait. Il faut distinguer alors entre l'assertion ϕ, qui intéresse la logique, et l'*acte* d'asserter ϕ, exprimé par l'assertion d'une assertion et qui n'intéresse plus la logique mais la psychologie :

[39] [Russell 1923a], p. 158. Nous soulignons.

> Qu'entend-on par '⊢p'? Ce qui *se produit*, c'est que j'asserte quelque chose ; ce qui se produit, c'est ce dont on *asserte la production* lorsque j'asserte 'A croit p'. Mais la logique n'est pas concernée par ce qui se produit lorsque j'asserte ; elle n'est concernée que par ce qui est asserté.[40]

Russell ajoute une distinction supplémentaire pour évoquer les expressions linguistiques caractérisant soit une croyance à un fait, soit un fait : la *proposition logique*, qui désigne ce qui renvoie à un fait; la *proposition psychologique*, qui désigne ce qui renvoie à une déclaration de fait ou *énoncé*.
On peut résumer les relations d'ensemble entre énoncé, proposition, valeur de vérité et faits par le schéma suivant, où la flèche désigne la relation d'expression entre deux entités :[41]



Au total, la définition russellienne de la vérité exige une séparation rigoureuse entre logique et psychologie, la première concernant uniquement la relation entre une proposition et un fait ontologique :

> Donc bien que, lorsque j'asserte une proposition, j'asserte nécessairement un membre de la classe qu'est la proposition, en tant que logicien j'ignore en revanche tout ce qui concerne la distinction de ce membre avec les autres. *C'est ce qui rend la logique différente de la psychologie*, et davantage concernée par la vérité ou la fausseté des croyances.[42]

---

[40] [Russell 1923b], p. 159. Et cependant, la suite du passage qui vient d'être cité semble produire une contradiction dans le discours de Russell : "C'est là où une proposition diffère de celle qui est une classe de croyances, que l'on qualifier de *proposition psychologique*. Une proposition logique fait partie d'une croyance individuelle. En assertant 'p ou q', la proposition logique fait partie de la croyance". N'est-ce pas plutôt le contraire qui est le cas, à savoir, une proposition psychologique désigne une croyance exprimée dans un énoncé 'p', qui est un élément de la classe propositionnelle : que p? Jusqu'ici, une proposition n'est pas contenue dans une croyance puisqu'elle est la classe dont les croyances sont les éléments. A défaut de réponse précise sur ce passage ambigu, nous posons (en vertu de tout ce qui précède) qu'une croyance est exprimée par une proposition psychologique ; en ce sens, la proposition psychologique correspond à notre énoncé. Quant à la proposition logique, elle est une classe de propositions psychologiques.

[41] La vérité est décrite ici comme la relation de correspondance entre une proposition ('que ϕ') et un fait (noté 'ϕ'), et la fausseté comme une absence de fait correspondant (noté '$\varnothing$'). Mais à strictement parler, 'ϕ' n'est qu'un signe linguistique et ne représente donc pas un fait mais une expression de ce fait : c'est ce que Tarski affirmera en substance dans sa définition de la vérité comme relation entre deux niveaux de langage (le langage-objet et le métalangage) et non entre le langage et la réalité. Or puisque nous ne pouvons parler d'un fait sans passer par le langage, nous avons représenté par défaut le fait par un signe.

[42] [Russell 1923a], p. 158. Nous soulignons.

Quant aux formules de la logique épistémique, elles symbolisent ce que Russell appellerait des propositions psychologiques et qui sont exprimées par des énoncés. Les expressions modales $K\phi$ ou $B\phi$ sont liées à un fait qui les rend vrais ou faux, mais elles ne sont pas elles-mêmes vraies ou fausses *stricto sensu*. Si en revanche les faits exprimés par des assertions (ou négations) d'assertions sont des faits d'ordre psychologique, de quoi parle le logicien épistémique si son analyse ne porte pas sur des faits logiques ? Pour défendre celui-ci, on répondra que la condition (C.¬) de consistance des ensembles modèles est une formulation du principe de contradiction, et celui-ci *est* une proposition logique pour Russell. Or si les formules modales de croyance sont caractérisées par des principes logiques, pourquoi Russell les maintient-il à l'écart de toute analyse d'ordre logique ?

Le fait que Russell assimile les croyances à des faits psychologiques vient de ce que ces derniers ne se définissent pas en termes de vérité et de fausseté, contrairement aux faits logiques : la production des croyances serait due à des faits accidentels qui ne concernent pas les états de choses eux-mêmes, les faits ontologiques. Or nous avons indiqué précédemment que la logique épistémique est une sorte de prolongement de la théorie de l'image tableau de Wittgenstein : le logicien épistémique, contrairement au logicien tel que le conçoit Russell, n''ignore pas tout' des relations entre les croyances qui sont des membres d'une classe propositionnelle. Elles peuvent être décrites à la rigueur comme des faits psychologiques, mais elles n'en sont pas moins organisées selon un critère minimal de *cohérence* logique, à l'image de (C.¬). On peut parler à cet égard d'une réduction *logiciste* des croyances à des ensembles de propositions non-contradictoires, dans une logique épistémique de type [Hintikka 1962].

Une dernière question concerne la relation entre cohérence et vérité : peut-on parler de la cohérence (la non-contradiction) d'un ensemble d'énoncés sans jamais parler de correspondance avec des faits ontologiques que ces énoncés exprimeraient ? Autrement dit : les faits logiques sont-ils parfaitement indépendants des faits ontologiques ?[43] Pas pour Russell, qui prétend que les partisans de la théorie de la vérité-cohérence présupposent ce qu'ils cherchent à éliminer :

> en réalité, 'cohérence' présuppose la vérité des lois de la logique (...) les lois de la logique fournissent le cadre où vient s'insérer le test de la cohérence, et ces lois mêmes ne peuvent être établies grâce à ce même test (...) Nous voilà donc ramenés à la *correspondance avec les faits*, pour définir la nature de la vérité.[44]

Considérée comme un cas limite de la vérité, le critère de cohérence entre des énoncés reposerait donc sur leur relation de correspondance avec des faits. Deux énoncés sont mutuellement incompatibles s'ils ne peuvent pas correspondre à un même fait, donc la structure des faits de langage dépend fondamentalement de la structure des faits du monde. Pour cette raison, distinguer la logique de Russell et la logique de Hintikka comme désignant respectivement une logique ontologique et une logique épistémique ne serait pas une chose

[43] Nous avons déjà posé la question dans la section 1.3.3, au sujet de la vérité logique : une formule complexe peut-elle être valide sans qu'aucun de ses atomes ne soit vrai ou faux ? Nous avions répondu que non, et Russell fait de même ici.
[44] B. Russell : *Problèmes de Philosophie*, "Le vrai et le faux", pp. 143-4.

correcte : la logique épistémique conserve une base ontologique pour déterminer ses règles de consistance ; on ne peut pas caractériser non plus la logique de Russell comme l'expression d'une théorie de la vérité-correspondance, par opposition à la logique de Hintikka qui serait l'expression d'une théorie de la vérité-cohérence : la différence entre leurs sémantiques ne repose pas sur une différence dans leur théorie de la vérité, car celle-ci porte dans les deux cas sur ce qui donne le droit de qualifier un énoncé de vrai ou de faux. Rappelons que Hintikka est lui-même un partisan de la vérité-correspondance, ce qui prouve que sa caractérisation modale de la vérité logique (vraie dans tous les mondes possibles) est indépendante de sa théorie de la vérité. Ce n'est donc pas la théorie de la vérité qui distingue les logiques de Russell et Hintikka, puisque tous deux adhèrent à l'idée d'une vérité comme relation entre un langage et des faits ; le différend porte plutôt sur l'usage qui est fait de la notion de vérité en logique modale contemporaine : la vérité n'a plus la proposition (*Satz an sich*) pour porteur spécifique ; un énoncé peut être qualifié lui-même de vrai ou de faux, quelle que soit la nature du fait qu'il exprime ; propositions et énoncés sont caractérisées indifféremment en termes de valeurs de vérité, même si leurs conditions de vérité sont distinctes.

Ne risque-t-on pas de perdre de vue le sens primitif de la vérité si celle-ci désigne indistinctement les conditions de satisfaction d'une formule quelconque, qu'elle soit non-modale ou modale ? Au premier chapitre, nous avons signalé que ce que Kripke appelle la 'vérité dans un modèle' ($v(w,\phi) = 1$) signifie chez Hintikka l'appartenance à un monde possible (qu'il appelle 'ensemble modèle'). Or si la sémantique relationnelle assimile 'appartenance à' un monde et 'vérité dans' un monde, ne reproduit-elle pas la confusion des ordres que Russell avait soigneusement évitée par sa distinction entre fonctions de *vérité* et fonctions de *signification* ? Le problème de fond peut être exprimé ainsi, qui concerne la nature du lien sémantique entre les expressions d'un langage et les faits de la réalité. Selon une formule de Quine,

> le prédicat de vérité nous rappelle qu'en dépit d'une montée sémantique qui nous amène à parler des énoncés, notre regard est dirigé vers le monde.[45]

L'est-il encore vraiment en logique modale épistémique, ou ne parle-t-on dans celle-ci de vérité et de fausseté que de façon dévoyée ? La question est importante, puisqu'il s'agit de savoir ce qui justifie les conditions de vérité d'un espace épistémique. Si en effet on suppose que ce n'est pas une relation de correspondance qui établit les relations possibles entre une croyance et une autre, pour quelle raison faut-il maintenir le critère de consistance des logiques modales normales, et quelle raison empêche le logicien de modifier ces règles sémantiques comme bon lui sied ?

Le problème est celui de justifier les règles qui déterminent la signification d'une formule modale. Le prétendu lien entre la structure des énoncés d'un langage et la structure des faits est une hypothèse douteuse, dans la mesure où nous n'aurions jamais accès aux faits eux-mêmes mais seulement à leurs expressions linguistiques. Et si, contrairement à l'approche réaliste de Russell, aucune 'structure ontologique' ne garantit les vérités logiques, alors comment les justifier et à quoi sont censées 'correspondre' les règles sémantiques dans un

[45] [Quine 1975], p. 24.

espace épistémique ? Une solution consiste à traiter les valeurs de vérité comme de simples valeurs algébriques n'exprimant aucune correspondance avec la réalité extérieure. La porte est alors ouverte au relativisme ou au formalisme en logique. Si le seul critère sémantique restant est l'accord avec nos intuitions de langage, alors la logique n'a plus vocation à corriger nos intuitions ordinaires ; au contraire, elle les conforte, et une fameuse formule de Russell semble même aller dans ce sens :

> Une théorie logique peut être testée selon sa capacité à traiter des énigmes, et c'est un projet salutaire, lorsqu'il s'agit de la logique, que d'emplir l'esprit d'autant d'énigmes que possible, puisque celles-ci servent en grande partie les mêmes besoins que les expérimentations en science physique.[46]

Mais l'usage de la notion de vérité n'est cependant pas libre, chez Russell, puisque la vérité exprime encore un lien avec la réalité et ne peut pas être redéfinie arbitrairement dans le seul but de résoudre (mieux, de dissoudre) des problèmes d'analyse du langage.

Nous avons mesuré ici le prix à payer pour l'interprétation d'une logique épistémique dans les termes de la logique selon Russell. Cette logique philosophique pose de sérieuses difficultés en philosophie de la logique dans la mesure où la notion de vérité doit être examinée avec prudence, en ce qui concerne du moins sa relation avec les entités linguistiques et les événements psychologiques. Examinons maintenant des raisons similaires qui rendaient inconcevable le projet d'une logique épistémique chez Wittgenstein.

2.1.2.3 Wittgenstein

Peut-on vraiment considérer l'espace logique du *Tractatus* comme un langage formel clos doté de règles de formation, de règles de transformation, de définitions explicites des connecteurs, etc.? La logique de Wittgenstein semble autant réductible à la logique formelle moderne que la physique d'Aristote à la physique moderne. Il y a donc un risque évident à insérer les positions du premier Wittgenstein dans ce travail. Or il s'avère que le pionnier de la logique modale épistémique, Jaakko Hintikka, a été et reste un commentateur attitré du philosophe Autrichien ; nous le verrons de nouveau plus loin, concernant sa distinction historique entre deux conceptions de la logique dans [Hintikka 1997a]. Il s'agit de montrer dans cette section qu'un détour par Wittgenstein peut rendre compte de certaines réticences profondément ancrées face au symbolisme modal, dont la logique des attitudes propositionnelles est l'expression.

Deux questions sont posées ici : quel rapport y a-t-il entre le *Tractatus* et le thème de la logique modale, en premier lieu ; quel traitement était réservé aux attitudes propositionnelles dans le premier Wittgenstein, en second lieu. Nous nous appuierons ici sur les seules données du *Tractatus* et ne ferons aucune allusion aux textes du second Wittgenstein, notamment ses *Investigations Philosophiques*.[47] Nous y viendrons toutefois au cours du travail, puisque

---

[46] [Russell 1905], pp. 484-5.

[47] Le but de cette section est d'illustrer une certaine conception de la sémantique formelle liée à la fois au réalisme logique (le langage est le reflet du monde) et à l'atomisme logique, voire épistémologique (le langage est constitué d'atomes élémentaires, et ces atomes peuvent exprimer des faits élémentaires ou états de choses du monde). Par opposition, les *Investigations Philosophiques* sont plutôt liées à l'inscrutabilité de la référence (les

Hintikka citera à plusieurs reprises la notion de jeux de langage afin d'étayer sa définition de la logique comme modèle explicatif, dans un premier temps, puis sa sémantique des jeux (dite GTS) dans un second temps.

L'idée générale de cette section concerne la relation entre *signe* et *symbole* : certaines notions de logique peuvent recevoir des traitements totalement incompatibles lorsque les auteurs sont issus de traditions différentes, et le but de ce qui suit est d'illustrer cette incompatibilité à propos des attitudes propositionnelles ; celles-ci apparaissent dans le *Tractatus*, mais sous un angle tel qu'il rendait impossible la construction d'une logique épistémique. Remonter aux origines de cette différence et découvrir les raisons théoriques de cette impossibilité, tel est le but de ce qui suit. En procédant de la sorte, nous reviendrons notamment sur ce qui rapproche et distingue à la fois l'idée de sémantique formelle selon les logiciens. Les distinctions peuvent être de taille, et Hintikka a examiné cette évolution des idées dans l'histoire de la logique. Venons-en tout d'abord au traitement des modalités dans le *Tractatus*, puis à celui des attitudes propositionnelles.

La fonction projective du langage selon Wittgenstein :

Dans plusieurs paragraphes de du *Tractatus*, on trouve des réflexions sur l'espace logique qui sont exprimées en termes modaux ; il y est fait allusion à ce que nous appellerions des modalités aléthiques et épistémiques, ainsi qu'à la notion d'assertion. Pour comprendre le sens donné à ces modalités, il faut comprendre tout d'abord la nature de l'espace logique.

En vertu de la théorie de l'image tableau (*Bildtheorie*) de Wittgenstein, le langage fonctionne comme une représentation imagée de la réalité : "Une image représente une situation possible dans l'espace logique" (§2.202), et les éléments du langage sont combinés de telle sorte que la réalité est décrite selon une méthode de projection des pensées : "Le signe avec lequel nous exprimons une pensée, je l'appelle signe propositionnel. – Et une proposition est un signe propositionnel dans sa relation projective au monde" (§3.12). Il s'agit de concevoir le langage comme un reflet de la réalité constituée de faits élémentaires (les *Tatsachen*), et d'exprimer ces faits par des propositions (des prédications, qui sont des corrélations d'objets). La proposition est donc plus qu'un simple signe linguistique, y compris un énoncé, puisqu'un signe ne devient proposition que s'il reflète une partie de la réalité. Cette conception du langage est appelée plus ou moins la thèse de l'*isomorphisme structurel* entre langage et monde. Si l'on conçoit la sémantique comme l'étude de la relation entre un signe et l'objet qu'il symbolise, l'étude de ces relations fait du *Tractatus* un essai de sémantique. Or cette sémantique est soumise à ce que Hintikka appellera par la suite un problème d'*ineffabilité* : toute pensée est le résultat d'une relation projective, et cette projection donne une image possible d'un fait : "Une image logique des faits est une pensée" (§ 3) ; or nous ne pouvons pas remonter aux sources de ces pensées, c'est-à-dire décrire les règles de projection qui président aux relations instaurées entre les mots d'un langage et les objets du monde. Pour plagier Kant, nous dirions que, pour Wittgenstein, les conditions de possibilité de la pensée ne sont pas connaissables : les faits qui composent le monde sont reflétés par des propositions dans un langage *idéal*, mais non par des énoncés dans le langage

---

objets de référence sont indéterminés) et au *molécularisme* logique (l'énoncé est le primat premier de la signification, et celle-ci est déterminée par des règles de langage).

ordinaire. Or ces propositions ne peuvent pas exprimer le processus de reflet qui s'effectue entre elles et les faits du monde. En d'autres termes, la méthode de projection qui relie le langage à la réalité ne peut pas être *dit*e mais seulement *montrée* à travers ses effets que sont les propositions du langage, pour reprendre la fameuse distinction de l'Autrichien : "Ce qui peut être montré ne peut être dit" (§4.1212). Le caractère inexprimable de la méthode de projection a des répercussions sur des thèmes généraux de la philosophie de la logique tels que la distinction entre langage objet et métalangage, le solipsisme, la sémantique formelle, mais aussi sur l'analyse logique des attitudes propositionnelles. Nous voulons montrer ici que la théorie projective de la représentation sémantique selon Wittgenstein rendait impossible la distinction aujourd'hui courante entre langage objet et métalangage. Le résultat de cette impossibilité conduit à une interprétation non formelle des modalités, notamment.

Conséquence sur l'idée de logique des modalités :

Concernant les modalités dites 'aléthiques', le Wittgenstein du *Tractatus* les utilise pour exprimer la nature de ses propositions. Aucune différence ne semble apparaître par rapport à la notion de nécessité logique, puisque l'auteur parle de nécessité pour évoquer le statut de ce que l'on appellerait imprudemment les faits logiques, c'est-à-dire les vérités et faussetés logiques :

> La tautologie et la contradiction ne sont pas des images de la réalité. Elles ne figurent aucune situation *possible*. Car celle-là permet *toute* situation possible, celle-ci *aucune*. Dans une tautologie, les conditions d'accord avec le monde –les relations de représentation –s'annulent mutuellement, de sorte qu'elle ne constitue en rien une relation de représentation de la réalité (§4.462).[48]

La différence entre la nécessité de Wittgenstein et celle des logiciens modaux réside dans le fait que, contrairement à ces derniers, Wittgenstein refuse de *symboliser* la nécessité des pseudo-faits logiques. Premièrement, si les vérités logiques sont des vérités indépendantes des faits du monde, elles ne sont pas des vérités à strictement parler puisque vérité et fausseté exigent dans le *Tractatus* une relation de correspondance ou de non-correspondance avec des faits : "Une image s'accorde ou ne s'accorde pas avec la réalité ; elle est correcte ou incorrecte, vraie ou fausse" (§ 2.21). Deuxièmement, l'ineffabilité du lien établi entre langage et monde constitue précisément à ce que Wittgenstein considère comme nécessaire : "(...) La connexion entre la connaissance et ce qui est connu est celle de la nécessité logique. ('A sait que p est le cas' n'a pas de sens si p est une tautologie)" (§ 5.1362). Donc seul est nécessaire pour Wittgenstein ce qui n'est pas exprimable et qui constitue le mode de projection des propositions aux faits : le symbolisme de type $\Diamond\phi$ ou $\Box\phi$ est dépourvu de sens (*sinnlos*) ou inapplicable dans l'esprit du *Tractatus*, dans la mesure où toute proposition est vraie ou fausse par essence.[49] En effet, parler de cette essence revient à ne rien dire : c'est parler en dehors

[48] Si l'on met de côté la critique quinienne des dogmes de l'empirisme, la tradition philosophique distingue d'une part les vérités de fait, empiriques ou *a posteriori*, et d'autre part les vérités de raison, logiques ou *a priori*. Les vérités logiques sont indépendantes des faits ; or dans le *Tractatus*, seul un fait permet de rendre une proposition vraie ou fausse. Pour ces deux raisons, parler de faits logiques et de vérités ou faussetés logiques est doublement fautif.

[49] "De même qu'il n'y a qu'une nécessité logique, il n'y a aussi qu'une impossibilité logique" (§ 6.375).

des limites du langage, constituées par les seuls faits du monde, et exprimer autre chose que l'image d'un fait produit un non-sens puisque "Ce qu'une image représente est son sens" (§ 2.221).

Que deviennent les modalités, si le langage ne peut parler que de faits empiriques (bipolaires : soit vrais soit faux) et ne contient pour cette raison que des propositions simples (pas de propositions de second ordre, ou de jugements de valeur) ? Le premier effet de la fonction strictement projective du langage semble être le rejet des aspects neustiques du langage ordinaire, et la réduction à l'aspect phrastique dont traite la logique classique du premier ordre. La théorie du langage qui sous-tend le *Tractatus* permet de comprendre ainsi les difficultés posées par une interprétation de la logique modale, du moins les réticences ressenties face à son application. Les passages qui suivent rappellent l'analogie faite plus tôt (section 1.1.2.2.2) entre l'espace logique de Wittgenstein et le vocabulaire de la logique modale : "Les possibilités de vérité des propositions élémentaires signifient les possibilités d'existence et de non-existence des états de choses" (§ 4.3) ; "Les possibilités de vérité des propositions élémentaires sont les conditions de la vérité et la fausseté des propositions" (§ 4.41) ; l'usage des modalités ne peut donc être que métaphorique,[50] dans le cadre de la théorie projective du langage, mais en aucun cas ces modalités ne peuvent dire quoi que ce soit sur le monde. Pour cette raison, les modalités ne "disent rien".

La limitation du symbolisme à l'aspect descriptif (phrastique) du langage a des effets non seulement sur l'interprétation des modalités aléthiques, mais aussi sur les modalités épistémiques qui représentent des attitudes propositionnelles. Un parallèle a été établi précédemment entre les tables de vérité et l'espace épistémique, et un passage du *Tractatus* semble confirmer tout d'abord ce que nous avions dit plus tôt de la théorie de l'information de Carnap :

> La vérité d'une tautologie est certaine, celle d'une proposition est possible, celle d'une contradiction est impossible. (Certain, possible, impossible : nous avons là une première indication de l'échelle dont nous avons besoin en théorie de la probabilité.) (§ 4.464).

Mais en réalité, la certitude dont il est question ici ne porte absolument pas sur l'état de connaissance d'un sujet particulier : c'est une certitude relative à la constitution de l'espace logique, sans rapport avec la conviction d'un sujet particulier. Par ailleurs, un passage précédent avait montré qu'une connaissance d'un sujet ne peut pas porter sur des 'vérités logiques' : "'A sait que p est le cas' n'a pas de sens si p est une tautologie" (§ 5.1362). A cet égard, la règle de nécessitation de la logique modale ne trouverait pas sa place dans le *Tractatus*, d'autant plus que ce que l'on appelle aujourd'hui les 'vérités logiques' (tautologies) et les 'faussetés logiques' (antilogies) ne sont ni vraies ni fausses *stricto sensu* en vertu du caractère bipolaire, donc *contingent* des propositions. Notons par ailleurs que ces

[50] Un des anciens élèves de Wittgenstein, Georg Henrik von Wright, a consacré un article à l'interprétation des modalités dans le *Tractatus*. D'après lui, le schéma du système modal **S5** : $\Diamond p \rightarrow \Box\Diamond p$ symboliserait le caractère bipolaire de la proposition atomique p. Il s'explique ainsi : " L'idée que ceci ou cela est une possibilité logique exprime une nécessité lorsque c'est vrai ; et une impossibilité lorsque c'est faux, peut-on dire aussi bien. Les possibilités logiques sont les faits (!) de la logique, déclare Wittgenstein en 2.0121c(2)" ([von Wright 1986], p. 209). La possibilité de vérité des propositions est bilatérale, c'est-à-dire contingente : une proposition n'est par définition ni nécessaire, ni impossible.

passages, qui rappellent MacColl par leur description des notions épistémiques en termes de certitude, de connaissance et de probabilité, donnent une raison de rejeter en bloc aussi bien les modalités logiques que les modalités épistémiques : les propositions logiques ne sont pas des faits et, en cela, ils ne sont pas plus représentables dans le *Tractatus* que les attitudes cognitives du sujet. Il ne serait donc pas permis de concevoir l'espace logique de Wittgenstein comme un cadre propice en vue d'une théorie de la probabilité, conformément à l'analogie que Carnap fera plus tard entre descriptions d'état et information. Ce qui suit montre au contraire que le sujet ou agent épistémique ne figure pas dans l'espace logique du *Tractatus*. La difficulté porte sur deux façons d'interpréter le sujet, comme nous allons le voir dans l'analyse tractarienne des attitudes propositionnelles.

La déclaration 5.542 porte sur la forme logique d'une attitude propositionnelle particulière, l'attitude de croyance :

> Il est clair toutefois que 'A croit que p', 'A pense p', 'A dit p' sont de forme '"p" dit p' : et celle-ci ne concerne pas la corrélation d'un fait à un objet, mais la corrélation de faits au moyen d'une corrélation de leurs objets.

La formule en a laissé perplexe plus d'un, et continue d'être discutée. Pour notre part, nous allons tenter de dégager les raisons pour lesquelles Wittgenstein l'a prononcée ainsi, afin d'y retrouver quelques arguments que Hintikka contestera par la suite dans le but de défendre sa logique épistémique.

Le problème central porte sur la théorie des *fonctions de vérité* : toute fonction de signification est-elle une fonction de vérité ? L'analyse précédente de Russell avait distingué deux types d'occurrence d'une proposition : si elle apparaît comme l'expression d'un fait (ontologique), la fonction de signification dont cette proposition fait partie est une fonction de vérité ; si elle apparaît comme un fait à part entière, il n'y a plus de fonction de vérité puisque la proposition parle d'elle-même. C'est notamment le cas avec les attitudes propositionnelles, où le fait mis en valeur est la forme de mots utilisée par le locuteur. Toute fonction de signification ne serait donc pas vérifonctionnelle et, à cet égard, la position de Russell est comparable à ce que Frege disait au sujet du changement de la dénotation dans un contexte de discours indirect. Mais Wittgenstein ne semble pas admettre d'autres fonctions de signification que la fonction de vérité, dans le *Tractatus* : non seulement "Une proposition est une fonction de vérité des propositions élémentaires" (§ 5), mais "Une proposition a une et une seule analyse complète (§ 3.25)". Si tel est le cas, les paraphrases précédentes suggérées par Russell et Frege ne sont pas acceptables parce qu'elles font exception à la théorie des fonctions de vérité, la seule qui soit admise par Wittgenstein ("Dans la forme propositionnelle générale, les propositions ne figurent dans d'autres propositions que comme bases d'opérations de vérité", § 5.54). Il ne s'agit donc pas de délimiter un contexte d'application des propositions vérifonctionnelles et d'en exclure les propositions de type intensionnel ; il s'agit d'aligner *toute* proposition pourvue de sens sur le modèle des fonctions de vérité. La théorie de l'extensionalité ne doit souffrir d'aucune exception, dans la logique du *Tractatus*.

L'analyse de 'A croit p' élimine tout d'abord le sujet A de sa forme logique et le remplace par une *mention* de la proposition p, exprimée par les guillemets de citation "p". Le résultat

semble contre-intuitif, puisqu'il réduit la croyance à une relation impersonnelle entre une expression du langage (le signe propositionnel) et ce qu'elle désigne (le fait propositionnel, qui constitue une pensée). Pourquoi cette exclusion du sujet, et comment analyser une croyance en éliminant l'aspect psychologique de l'attitude ? [Russell 1923] ne l'avait pas éliminé de son analyse, puisqu'il concevait le sujet croyant comme une série de faits psychologiques exprimés par des énoncés. Mais de même qu'elles le seront par la suite dans [Hintikka 1962], ces croyances étaient associées à un sujet particulier ; or l'analyse présente de Wittgenstein efface totalement la présence de celui-ci et ne parle que d'un énoncé quelconque de forme "p". Pour comprendre cet aphorisme sibyllin, plusieurs auteurs ont rendu compte du § 5.542 en deux points : la théorie de l'objet de Wittgenstein ; sa distinction entre sujet empirique et sujet métaphysique.

La *théorie de l'objet simple*, tout d'abord. Dans l'analyse précédente de Russell, où la proposition était identifiée par une série de faits psychologiques, le fait de déterminer une proposition par des faits psychologiques introduit l'esprit de A dans l'analyse des attitudes propositionnelles. En conséquence, la forme logique couramment attribuée à 'A croit que p' correspond à la corrélation d'un fait, le fait propositionnel que p, et d'un objet, l'esprit de A. Or un objet est simple, dit Wittgenstein ("Les objets sont simples", § 2.02), et l'esprit de A est complexe (il est une série de faits psychologiques), donc la forme logique donnée à l'attitude propositionnelle n'est pas correcte. La forme logique exigée pour un état de choses ("Un état de choses est une combinaison d'objets", § 2.01) conduit ainsi Wittgenstein à disqualifier l'attitude propositionnelle en qualité d'état de choses, tout au moins dans sa version courante qui se présente comme la coordination d'un fait et d'un objet. Cette position l'amène également à cette déclaration, tout aussi déconcertante que la précédente : "Cela montre aussi qu'il n'y a pas de chose telle que l'âme –le sujet, etc. –telle que conçu dans la psychologie superficielle contemporaine. En effet, une âme composée ne serait plus une âme" (§ 5.5421). Le prix à payer pour admettre l'analyse logique de Wittgenstein n'est-il pas démesuré, si le rejet des attitudes propositionnelles conduit à un rejet de la psychologie elle-même ? [Favrholdt 1964] constate que ce résultat est une conséquence directe de la théorie des objets du *Tractatus* :

> Les psychologues superficiels prétendent en effet qu'il serait raisonnable de dire que 'A dit p' est une coordination d'un fait au sens wittgensteinien, à savoir un signe propositionnel, et d'un objet, à savoir l'âme pensante et représentante qui, parce qu'elle est simple, peut être qualifiée d'objet. Wittgenstein est contraint de rejeter ce point de vue. D'après la théorie de l'image du *Tractatus*, aucune coordination ne pourrait être établie entre un fait et un objet. Les deux entités en question doivent être également articulées afin d'être coordonnées. Les objets peuvent être coordonnés à des objets (parce qu'ils sont simples), et les faits peuvent être coordonnées à des faits, dans la mesure où ils peuvent être analysés en termes de la même somme d'éléments.[51]

Les principes de combinaison des propositions interdisent donc Wittgenstein de traiter les attitudes propositionnelles de la manière habituelle, celle que l'on retrouve par ailleurs dans la logique modale épistémique. Cela nous amène au second argument pour justifier le § 5.542 : la *distinction entre sujet empirique et sujet métaphysique*.

[51] [Favrholdt 1964], p. 559.

[Hintikka 1958] établit d'abord cette distinction pour revenir sur un malentendu dans la théorie du langage de Wittgenstein : la thèse du *solipsisme*, considérée en général comme un argument en faveur de l'idée de langage privé. Pour dissiper un malentendu dont tout est parti : "Le monde est mon monde : cela est manifeste dans le fait que les limites du *langage* (le seul que je puisse comprendre) signifient les limites de mon monde" (§ 5.62), Hintikka explique que ce dont Wittgenstein s'occupe

> n'est pas le sujet empirique mais le sujet 'métaphysique', examiné en philosophie. En d'autres termes, il s'intéresse seulement à ce que l'on peut dire qui est à moi *nécessairement* ; sans cela, il ferait seulement de la psychologie empirique. Mais la seule nécessité qu'il y a, d'après les autres doctrines du *Tractatus*, c'est la nécessité tautologique vide de la logique.[52]

En temps normal, le solipsisme sous-entend le caractère privé des idées ou leur incommunicabilité, et le passage qui suit semble aller dans ce sens : "Les limites de mon langage signifient les limites de mon monde" (§ 5.6). Or Wittgenstein n'est pas partisan du langage privé, bien au contraire. Pour dissiper ce malentendu sur le solipsisme, Hintikka explique que le moi du *Tractatus* n'est pas un moi psychologique, propre à un sujet pensant particulier. Ce n'est pas l'agent réel de sa future logique épistémique, mais un sujet abstrait qui incarne la totalité des propositions : "Le sujet n'appartient pas au monde : en fait, il constitue la limite du monde" (5.632). La thèse de l'isomorphie doit nous rappeler que les limites du monde sont déterminées par les limites du langage, dans la mesure où la relation projective entre les deux instances constitue à la fois ce qui vaut comme un fait représenté et ce qui vaut comme une proposition représentante. Et puisque l'ensemble des propositions constitue le langage ("La totalité des propositions est le langage", § 4.001), sujet et langage ne font qu'un.

De plus, le sujet métaphysique ne peut pas parler de lui-même au sein du langage qu'il incarne, contrairement à une attitude propositionnelle : "Aucune proposition ne peut faire une déclaration au sujet d'elle-même, parce qu'un signe propositionnel ne peut pas être contenu en lui-même (c'est là le cœur de la 'théorie des types')" (§ 3.332). En vertu de cette auto-référence impossible du sujet, la croyance que p est le cas ne peut pas exprimer la relation entre ce sujet et une proposition et s'en retrouve réduite à la relation entre un signe propositionnel et une proposition : "p" dit p, signifiant que le signe propositionnel exprime la proposition du fait que p. Le solipsisme exprime donc l'impossibilité évidente pour le sujet métaphysique de dépasser les limites du langage, puisque c'est ce dernier qui le constitue ; mais il ne signifie pas l'impossibilité pour un sujet psychologique d'exprimer ses propres pensées. Au contraire, les pensées de Wittgenstein sont tout aussi publiques que celles de Frege (les *Gedanken*), et son solipsisme ne signifie absolument pas que les pensées sont des représentations privées (*Vorstellungen*).[53] Néanmoins, ces pensées publiques sont dissociées

---

[52] [Hintikka 1958], p. 89.

[53] "Ce que l'on considère d'habitude comme étant l'affirmation du solipsisme, c'est l'*impossibilité* d'aller 'au-delà des frontières de moi-même'. Le solipsisme de Wittgenstein est basé sur l'affirmation exactement opposée, selon laquelle les frontières ordinaires du moi sont parfaitement *contingentes* et donc sans rapport 'pour ce qui est d'ordre supérieur'" ([Hintikka 1958], p. 91).
"Il n'y a cependant rien de privé et rien de psychologique concernant la notion de pensée de Wittgenstein. Tout comme la *Gedanke* de Frege, elle peut être partagée par différentes personnes. Dans le *Tractatus*, les signes

dans le *Tractatus* du sujet psychologique qui les saisit, d'où la forme logique proposée en 5.542.

La relation entre "p" et p est donc une relation entre un signe propositionnel et la proposition qu'il symbolise : cette distinction entre un énoncé et la pensée qu'il exprime, bien qu'elle constitue le problème de l'opacité référentielle et concerne à ce titre la logique épistémique, n'est pas exprimée de la même façon dans le *Tractatus*, pour la raison que nous venons d'évoquer. Le signe entre guillemets "p" représente un signe propositionnel qui, dans [Hintikka 1962], désigne l'énoncé d'un monde possible, ou ce qu'il appelait 'ensemble modèle'. Mais il ne faut pas confondre les points de vue des deux auteurs : là où le logicien modal s'intéresse aux pensées des sujets empiriques et porte alors son attention sur les énoncés, Wittgenstein s'en tient à un sujet métaphysique et ne parle de pensées que pour désigner des propositions, c'est-à-dire des entités idéales dont la signification est univoque. L'impossibilité d'exprimer les états du sujet empirique dans le langage tractarien fait que la croyance est réduite à la relation impersonnelle entre un signe propositionnel et une pensée, c'est-à-dire la proposition telle que l'entend Wittgenstein. La déclaration du § 5.542 décrit donc bien la coordination de deux faits, "p" et p, au moyen de la coordination de leurs objets, les signes du langage et les objets du monde.[54] Mais les affirmations du *Tractatus* interdisent qu'il en soit dit davantage au sujet du contenu des pensées. La relation entre l'usage d'un signe propositionnel et la pensée que vise un sujet empirique, par exemple, concerne la logique épistémique mais se trouve reléguée dans le domaine de la psychologie, chez Wittgenstein aussi bien que chez Russell et Frege.[55] Le point commun aux trois auteurs est donc la séparation stricte entre le point de vue du sujet et le point de vue logique, qui concerne les règles de constitution du langage.

Mais [Favrholdt 1964] rappelle cependant que le sujet pens*ant* est présent implicitement dans la relation exprimée en 5.542 entre "p" et p :

> '"p" dit p' ne dit rien de plus que p. Il déclare que le signe propositionnel *est pensé*, et ceci est la même chose que d'*asserter* la proposition p. En conséquence, d'après Wittgenstein, dans les

---

propositionnels sont totalement publics, et de même pour ce qui 'n'est pas exprimé dans les signes' ; car si cela était privé, cela ne pourrait pas être 'montré par leur application' (3.262). Et les limites du sujet métaphysique ne peuvent pas être les limites de pensées réelles, car il n'y a rien de nécessaire concernant *cette* limite-*là*. Les seules limites nécessaires sont les limites générales des pensées possibles, i.e. les limites du langage en général. C'est là la base du 'solipsisme' de Wittgenstein." ([Hintikka 1958], pp. 90-1).

[54] "Le passage 5.542 peut être reformulé désormais de la manière suivante : Mais il est clair que 'A croit que p', 'A pense p', 'A dit p' sont de forme '"p" dit p' ; et ici, nous n'avons pas la coordination d'un signe propositionnel p et d'un objet A, que d'aucuns conçoivent comme l'âme ou sujet non-composé de la personne A, mais une coordination de deux faits, à savoir le signe propositionnel p et la pensée 'p' qui, lorsqu'elle se produit dans l'ego empirique de A, transforme le signe propositionnel p en proposition p. Les deux faits sont coordonnés au moyen d'une coordination de leurs objets" ([Favrholdt 1964], p. 561).

[55] Logique et psychologie traitent deux questions distinctes posées à un même thème : la *pensée*. [Hintikka 1958] relève que "Dans une lettre à Bertrand Russell, datée du 19.8.1919 (voir *Notebooks*, p. 129), Wittgenstein écrit : '(...) Mais une *Gedanke* est un *Tatsache* : quels sont ses constituants et composants, et quelle est leur relation à ceux du *Tatsache* imagé ?' Je ne sais pas ce que sont les constituants d'une pensée, mais je sais qu'elle doit avoir des constituants qui correspondent à des mots du langage. Une fois encore, le type de relation des constituants de la pensée et du fait imagé est sans importance. Cela serait un problème de psychologie à découvrir' (p. 557). Nous avons déjà distingué par pensée pure et pensée en acte ces deux questions, l'une relative aux règles de la pensée (*Gedanke*) et l'autre à ses manifestations concrètes (*Denkens*).

énoncés comme 'A dit p' p ne figure pas dans une proposition d'une façon particulière qui serait en conflit avec sa théorie générale des fonctions de vérité.[56]

En d'autres termes, '"p" dit p' prend tout son sens s'il est interprété non pas comme une relation entre un signe et ce qu'il symbolise, mais comme une action commise par un sujet empirique pour effectuer cette relation. La relation n'est pas statique, elle est dynamique : l'idée n'est pas seulement que le signe "p" sert à symboliser la proposition (l'image) p, mais que quelqu'un se sert de "p" dans cette *intention*. Le signe propositionnel *devient* une pensée lorsqu'il est utilisé dans l'intention de représenter son image, et l'intention annonce clairement l'intervention du sujet psychologique (croyant) lors de l'utilisation d'un signe.[57] On retrouve ici la distinction sémiotique usuelle entre le signe, son sens et son référent : le signe propositionnel "p" indique une direction, en vertu de la relation projective qui lui associe une image (sens), et cette direction sera vraie ou fausse si elle correspond ou non à un fait du monde (référent). La croyance correspond dans cet ordre d'idées à l'utilisation de "p" par un sujet empirique en vue d'indiquer que l'image (la proposition) p représente un fait. Le sujet empirique, bien qu'il ne soit pas exprimé dans la forme logique de l'attitude propositionnelle, est donc présent en filigrane, pour établir la relation de projection entre son langage et la réalité. La compétence linguistique détermine le succès ou l'insuccès de la connexion mais, encore une fois, celle-ci ne concerne pas la logique du *Tractatus* ; cette compétence est inexprimable, puisqu'elle ne constitue pas un fait que le langage serait en mesure de symboliser par une proposition. Pour le dire autrement, l'espace logique ne traite que la partie visible de l'iceberg que constitue la pensée ; quant à la partie immergée, elle ne concerne pas la logique.[58]

Nous finirons par deux remarques sur la notion d'assertion, d'une part, ainsi que sur le statut de la logique selon Wittgenstein, d'autre part. Nous tenterons de montrer dans cet essai que la logique épistémique correspond à un compromis entre les deux thèmes : elle formalise et transforme (calcule sur) des expressions de croyance, dont l'assertion est la manifestation publique. Mais la conception tractarienne de la logique empêche une telle procédure.

Puisque l'*assertion* désigne la manifestation d'une croyance par le biais d'une déclaration, elle n'ajoute rien aux propositions qui servent à l'extérioriser et se situe derechef dans le domaine de la psychologie. L'idée d'une 'logique doxastique' est absurde, dans le cadre de la théorie du langage tractarien :

---

[56] [Favrholdt 1964], p. 560. Nous soulignons.

[57] "Le signe propositionnel appliqué, pensé, est la pensée" (§ 3.5), et "La pensée est la proposition pourvue de sens" (§ 4). Le sujet empirique établit donc la liaison entre "p" et p ou donne un sens à leur relation de projection, à proprement parler. Mais le succès ou l'échec de cette projection est un problème d'ordre psychologique et non logique, en conséquence de quoi la question de l'opacité référentielle n'est pas un problème de logique mais de psychologie.

[58] "La pensée est un concept fondamental du *Tractatus* que l'on peut déjà apercevoir dans la préface, où il est dit que 'Le livre va (...) tracer une limite pour l'acte de penser, ou pas à l'acte de penser mais à l'expression des pensées, plutôt'. La limite ne peut être tracée que dans le langage ; en conséquence de quoi Wittgenstein, ayant montré que les pensées sont des images et que les pensées peuvent être exprimées dans des propositions, traite seulement de propositions dans le *Tractatus*", signale [Hintikka 1958] (p. 557).

> La pensée produit une proposition à partir du signe propositionnel p, et cela revient à la même chose que d'asserter p. Si p n'est pas pensé, il reste un signe propositionnel, et l'expression '⊢ p' est dans ce cas absurde ; vous ne pouvez pas à la fois asserter, c'est-à-dire penser p, et ne pas penser p. Donc le signe d'assertion est totalement dépourvu de sens du point de vue logique (voir 4.442).[59]

Quant aux *règles de la logique*, elles déterminent les limites au sein desquelles les sujets s'expriment et peuvent s'exprimer : contradictions et tautologies ne disent rien mais représentent des formes propositionnelles que le sujet ne peut pas penser, soit parce qu'elles ne visent aucune image particulière (tiers exclu), soit parce qu'elles annulent toute projection (non-contradiction). La nature projective du langage selon Wittgenstein fait de sa théorie logique une sorte de cadre transcendantal de la pensée. La logique du *Tractatus* est-elle une logique transcendantale ? Certains passages semblent converger dans ce sens, dans la mesure où les limites du langage qu'elle décrit constituent les limites du monde ; on pourrait interpréter cette idée comme une réminiscence de la logique transcendantale de Kant, définie comme l'étude des conditions de possibilité *a priori* de l'expérience en vertu des catégories de l'entendement. Hormis la notion d'entendement, le criticisme de Kant se retrouve ici dans l'impossibilité pour le sujet empirique de connaître les limites du langage ; le sujet pense *dans* le langage, mais il ne peut en sortir pour observer depuis l'extérieur ce qui distingue les pensées logiques des pensées illogiques.[60]

La logique se confond donc avec les méthodes de projection des propositions du langage en faits de la réalité, mais ces méthodes sont elles-mêmes inexprimables, sinon dans leurs effets de 'surface' (la totalité des propositions du langage). Hintikka reprendra à son compte cette déclaration d'impossibilité du premier Wittgenstein, qu'il décrira comme caractéristique d'une certaine conception de la logique. Dans celle-ci, la sémantique formelle n'est pas descriptible : les règles d'application d'un ensemble de formules dans un modèle, telles qu'elles peuvent été décrites dans [Hintikka 1962], étaient pour ainsi dire inconcevables dans l'esprit tractarien. Si l'on tient compte de cette limitation des moyens symboliques, notamment de l'absence de métalangage où sont consignées les règles sémantiques, d'autres affirmations du *Tractatus* prennent un sens qu'a sans doute rendu obscur la pratique actuelle de la logique, plus formelle (abstraite) que symbolique (concrète) : "Nous ne pouvons pas dire quelque chose d'illogique, car sinon nous devrions penser illogiquement" (§ 3.03), notamment, signifie que personne ne peut exprimer une pensée en dehors des règles du

---

[59] [Favrholdt 1964], p. 560. Le fait que croyances et assertions se situent en amont de la logique est expliqué ainsi par Wittgenstein : "La 'barre de jugement' de Frege '⊢' est tout à fait insignifiante logiquement : dans les travaux de Frege (et Russell) elle indique simplement que ces auteurs tiennent pour vraies les propositions marquées de ce signe. '⊢' n'est donc pas plus une partie composante de la proposition que le nombre de la proposition, par exemple. Il est tout à fait impossible pour une proposition de déclarer d'elle-même qu'elle est vraie" (§ 4.442).

[60] Quelques déclarations peuvent corroborer cette comparaison avec Kant. La logique est un cadre de la pensée : "Il est aussi impossible de représenter dans le langage quelque chose qui 'contredit la logique' que, dans la géométrie, de représenter par ses coordonnées une figure qui contredit les lois de l'espace ; ou que de donner les coordonnées d'un point qui n'existe pas." (§ 3.032). La logique est *a priori* : "Notre principe est que toute question susceptible d'être en général décidée par la logique, devrait être établie sans autre apport" (Et si nous nous trouvons en situation de devoir résoudre un tel problème en observant le monde, cela montre que nous nous sommes engagés dans une voie fondamentalement erronée.)" (*Tractatus*, § 5.551). La forme logique n'est pas librement créée : "Et d'ailleurs, est-il vraiment possible qu'en logique j'aie affaire à des formes que je puis inventer ? C'est plutôt à ce qui me permet de les inventer que je dois avoir affaire" (§ 5.555).

langage. Si un sujet A croit à une contradiction, par exemple, [Favrholdt 1964] rappelle que la distinction entre signe propositionnel et pensée rend cette croyance dépourvue de sens (sa projection est impossible, elle n'indique aucune direction concevable) :

> Ne serait-il pas possible que A dise 'p ∧ ¬p', violant ainsi les lois de la logique ? La réponse est : non. (...) A peut penser 'p', ou il peut penser '¬p'. Dans le premier cas, la première liaison de p ∧ ¬p deviendra une *proposition*, mais la dernière partie (¬p) restera un *signe propositionnel*, parce qu'elle n'est pas *pensée*. (...) Par conséquent, si l'on ne peut pas penser quelque chose d'illogique, on ne peut pas non plus présenter quoi que ce soit dans le langage qui 'contredise la logique'. Car le langage n'est pas les faits physiques que nous appelons signes propositionnels, mais les faits dans leur relation projective à d'autres faits.[61]

Dans la mesure où elle constitue une simple extension de la logique classique propositionnelle, la logique modale épistémique est en total accord avec celle-ci lorsqu'elle interdit deux propositions contradictoires d'appartenir à un même modèle. La formalisation de ce fait d'impossibilité, qui donne la condition sémantique (C.¬), ne dit rien de plus que Wittgenstein n'ait dit ici ; mais elle le dit autrement, dans les termes d'une sémantique formelle.

D'autres thèmes de la logique épistémique concernent également la relation entre un signe propositionnel et la pensée qu'elle symbolise : l'opacité référentielle, le paradoxe de Moore, et tout ce qui est lié en général à la théorie de l'intentionnalité. Toute la question reste de savoir ce que la logique épistémique apporte de plus qu'une analyse telle que le § 5.542, par exemple. L'opposition de Wittgenstein à l'idée même de logique épistémique tient donc à sa théorie de l'image-tableau, qui ne porte que sur des propositions déclaratives liées à la description de la réalité. Cette interdiction de principe se traduit en termes d'*ineffabilité de la sémantique*, c'est-à-dire des relations entre le langage et la réalité, puisque le langage tractarien ne contient que des propositions (qui sont le résultat de cette relation). Elle se traduit également par une distinction lourde de sens entre *symbolisme* et *formalisme* : le langage symbolise le monde, uniquement, et la mise en forme (formalisation) des images du monde dans le langage ne doit pas porter sur autre chose que le tableau de la réalité. Cet arrière-plan théorique explique sans doute l'influence de la logique classique du premier ordre au début du 20e siècle, pour autant que cela concerne l'ontologie du *Tractatus*. Dans la logique mathématique de Frege et Russell, le résultat est semblable puisque leur utilisation de la logique conduit à l'application inconditionnelle de la théorie des fonctions de vérité. Notre conclusion provisoire est donc que toute proposition peut être analysée extensionnellement tant que sa fonction consiste à décrire une partie de la réalité (usage *référentiel* du langage), et que la théorie des fonctions de vérité ne convient pas pour décrire la relation d'application entre langage et réalité (usage *intentionnel* du langage). Nous reviendrons plus loin sur l'arrière-plan théorique que suppose une logique de l'usage intentionnel du langage, dont la logique épistémique est un exemple notoire.

L'universalité du langage et l'ineffabilité de la sémantique sont deux conditions que Hintikka exclura de sa propre conception de la logique. Le problème que pose la logique épistémique est d'ordre sémantique, ici, car elle ne semble concevable que lorsque les

[61] [Favrholdt 1964], pp. 561-2.

postulats précédents (universalité du langage, ineffabilité de la sémantique) ont été abandonnés. La théorie de l'image ou *Bildtheorie* de Wittgenstein est apparue comme une conséquence de ces postulats, et son maintien intégral par le logicien épistémique aboutirait à une situation paradoxale. En effet, une logique des attitudes épistémiques porte sur l'*acte* de penser plus que sur la pensée qui en résulte, mais elle propose d'expliquer cet acte par le biais des expressions de la logique formelle. Or si l'acte de penser se décrit en termes de ses expressions, qui sont des images, cela semble signifier qu'il s'introduit lui-même dans l'image et que, en vertu de l'universalité du langage, la logique épistémique aboutit à l'idée absurde de vouloir expliquer le processus de reflet ou de projection d'une image par cette image elle-même. C'est ce résultat que nous semblons avoir admis précédemment, dans la sémantique des attitudes propositionnelles. Décrire une attitude propositionnelle en termes de propositions, s'il constitue un paradoxe apparent, devra être résolu néanmoins par la suite. A travers cette justification, nous tenterons de donner un meilleur aperçu de la conception actuelle de la logique, clairement distincte de l'apparence transcendantale qu'elle dégage dans le *Tractatus*. La réduction du rôle symbolique de la logique à un rôle formel sera mise en avant pour rendre compte de la logique épistémique et, surtout, de la raison pour laquelle la pratique de la logique modale nous semble correcte malgré les explications ci-dessus.

Un dernier auteur mérite d'être considéré pour son opposition à la logique des croyances, dans la mesure où ses positions vis-à-vis de la logique sont plus souples que les précédents mais pas moins classiques : le cas de W. V. Quine sera présenté ci-dessous comme un exemple de défiance à l'égard de la logique philosophique et de ses résultats formels. Comment exclure les modalités épistémiques de la liste des constantes logiques sans buter de nouveau sur le problème initial de cette thèse, à savoir, la non-vérifonctionnalité des énoncés modaux ? L'incapacité de la théorie des fonctions de vérité à comprendre ce type de discours ne montre-t-elle pas l'insuffisance de la logique classique et la nécessité de s'en remettre à une logique modale ? C'est parce que Quine a tenté de montrer le contraire jusqu'à sa mort que son cas est intéressant. Observons-le à présent.

2.1.2.4 Quine

Quine est considéré aujourd'hui comme l'adversaire majeur de toutes les formes de logique modale, même si ses attaques ont porté principalement sur les modalités de type aléthique et épistémique.[62] Ces deux modalités sont liées à deux thèmes récurrents dans la philosophie de Quine, à savoir : l'analyticité, d'une part, la théorie de la connaissance d'autre part. La logique épistémique n'échappe donc pas à la règle car, selon l'auteur, elle porte en elle les deux défauts qu'il associe généralement au langage des modalités : l'inutilité d'un côté, l'imprécision de l'autre. Explicitons le contenu de ces deux critiques.

[62] Pour une recension des articles de Quine contre la logique modale, voir [Quine 1943, 1947, 1953, 1956, 1960, 1961d, 1964, 1976, 1977, 1986, 1993]. Les articles constructifs, destinés à trouver une analyse logique extensionnelle pour les énoncés modaux, sont [Quine 1956, 1960, 1968, 1993]. Une allusion est faite à la logique temporelle dans [Quine 1975] et [Quine 1976] ; ses objections aux modalités ne portent en fait que sur les énoncés susceptibles d'intéresser le discours scientifique. Aucun de ses articles ne porte sur la logique déontique, par exemple.

Concernant le reproche d'inutilité de la logique modale, Quine considérait notamment la modalité aléthique et son *implication stricte* comme le produit d'une confusion entre deux niveaux de discours : le langage objet d'un côté, au sein duquel le logicien effectue des calculs sur des formules ; le métalangage de l'autre, où les règles de calcul et le domaine de discours sont fixés. Selon la distinction utilisée par Quine entre l'*usage* d'une expression et sa *mention*, la logique modale cherche à utiliser des expressions métalinguistiques au sein du langage objet ; il impute notamment la responsabilité du 'péché de la logique modale'[63] à son ancien professeur de philosophie, C. I. Lewis, lequel est accusé d'avoir confondu à travers le connecteur de l'implication stricte deux formes d'implication logiquement distinctes : l'implication matérielle ou *conditionnel*, qui est un connecteur propositionnel reliant deux énoncés utilisés dans le langage-objet de la logique classique, et l'implication formelle (ou implication logique), qui est une relation de conséquence entre deux noms d'énoncés mentionnés ; la critique s'adressait également au Russell des *Principia Mathematica*, dont la notion d'implication formelle anticipait cette confusion dans les ordres de discours. D'après Quine, l'implication stricte est un connecteur inutile parce que son but est déjà rempli par la notion métalinguistique d'implication formelle. Or même si cette implication modale trouve déjà son expression dans le métalangage classique, elle ne semble pas régler une difficulté sémantique propre au conditionnel : les paradoxes de l'implication matérielle. Comment les éviter, sinon en remplaçant le conditionnel par un connecteur plus apte à exprimer la relation déductive entre un antécédent et son conséquent ?

Le 'péché' de la logique modale semble se justifier ainsi : les paradoxes de l'implication matérielle surgissent lorsque l'on prend le conditionnel 'si ... alors', symbolisé par la flèche $\rightarrow$ du langage objet, pour ce que Quine associe exclusivement à la notion métalinguistique distincte d'implication, symbolisée par la flèche $\Rightarrow$ du métalangage. Autrement dit, la confusion porte sur les formules $(\phi \Rightarrow \psi)$, c'est-à-dire $\vDash (\phi \rightarrow \psi)$, et $(\phi \rightarrow \psi)$. L'implication est un conditionnel *toujours* vrai, et la situation paradoxale apparaît pour certaines assignations de valeurs de vérité. A la question : peut-on admettre une relation *conditionnelle* lorsque l'antécédent est faux ou le conséquent est vrai, ceux qui répondaient par la négative ont encouragé l'utilisation de l'implication stricte puis de la logique relevante.[64] Quant à Quine, son objection consiste à dissocier un conditionnel de l'idée de dépendance sémantique entre antécédent et conséquent : cette dépendance concerne l'implication, non les conditionnels. Un conditionnel n'est pas une *preuve* dans laquelle la vérité de la conclusion

---

[63] La motivation théorique de l'implication stricte remonte bien avant le début du 20e siècle, cela dit (voir section 1.1.1) ; contre l'implication matérielle ou philonienne, du nom de son auteur Philon de Mégare, les stoïciens avaient répondu par leur théorie de la conséquence comme *incompatibilité* : p implique q si et seulement si la fausseté de q est incompatible avec la vérité de p. Une théorie médiévale du même type, associée à Pierre Abélard, avait pour but d'exprimer la relation de conséquence comme une relation de contenance entre le sujet et le prédicat. Cette idée de conséquence comme fait de contenir sa conclusion se retrouve dans la notion de jugement analytique, où la conclusion est supposée contenue dans la prémisse.

[64] Dans une table de vérité, l'implication $(p \Rightarrow q)$ correspond à un conditionnel *valide* en vertu duquel l'implication matérielle $(p \rightarrow q)$ est une tautologie : $v(p \rightarrow q) = 1$ quel que soit $v(p)$ et $v(q)$. Les paradoxes de l'implication matérielle apparaissent dans les cas où $v(p) = 0$ et $v(q) = 1$: le faux implique n'importe quoi (*ex falso sequitur quodlibet*), et le vrai est impliqué par n'importe quoi (*verum sequitur ad quodlibet*). Cela étant, ces deux exemples d''implication' paraissent moins choquants lorsqu'on les interprète en termes ensemblistes d'inclusion : la classe vide est incluse dans toutes les classes, et la classe universelle inclut toutes les classes.

résulte toujours de la prémisse,[65] et l'assimilation des deux est à l'origine des logiques modales, inutiles et même nuisibles si elles sont nées d'une confusion théorique. Par ailleurs, l'implication stricte reproduit à son échelle d'autres formes de paradoxes : de l'impossible il s'ensuit n'importe quoi, et le nécessaire s'ensuit de n'importe quoi. Le mérite de la logique modale s'en trouve sérieusement écorné, si la solution de l'implication stricte est non seulement superflue mais reproduit des difficultés semblables à celle qu'elle prétend résoudre.

On retrouve la même sorte de critique pour d'autres interprétations de la logique modale : de même que Quine reproche aux logiciens modaux de confondre conditionnel et inférence, il reproche aux intuitionnistes de confondre vérité et croyance justifiée. Contre la réduction par Dummett de la classe des propositions vraies aux propositions assertables, il reprend le même argument que ci-dessus et parle d'une confusion inutile entre vérité et connaissance :

> L'argument qui est presque le pire est une confusion entre le savoir et le vrai. Il est certain qu'il y a un vaste domaine qui est intermédiaire entre les énoncés que nous savons ou croyons être vrais et ceux que nous savons ou croyons être faux ; mais nous pouvons quand même affirmer que chacun de ces énoncés de la zone intermédiaire est ou bien vrai sans que nous le sachions, ou bien faux sans que nous le sachions. Il est possible que la discorde vienne en partie d'une confusion entre (a) savoir que quelque chose est vrai ou faux, et (b) savoir que quelque chose est vrai ou savoir qu'il est faux.[66]

Or de même que l'enjeu de la notion de nécessité ne se limite pas à un problème de définition de l'implication, la notion de croyance justifiée ne se limite pas non plus à la définition de la vérité. Après tout, la logique modale épistémique n'est pas pieds et poings liée avec les enjeux de l'intuitionnisme, la preuve étant que la confusion dénoncée ci-dessus n'apparaît pas du tout dans [Hintikka 1962] : confondre (a) et (b), ce serait admettre la règle de clôture modale sous disjonction $\Box(p \vee q) \Rightarrow (\Box p \vee \Box q)$. Puisque cette formule n'est pas valide en logique modale épistémique (seule sa version faible l'est : $\Diamond(p \vee q) \Rightarrow (\Diamond p \vee \Diamond q)$), le reproche fait par Quine à la théorie de l'assertabilité de Dummett ne concerne pas tant la validité de la formule ci-dessus que la distinction introduite par ce symbolisme modal entre vérité et connaissance de la vérité. En réalité, les reproches adressés par Quine à celle-ci se situent sur un autre terrain que celui de la confusion des ordres de discours. Nous en venons à l'argument de l'imprécision des logiques modales.

Ce défaut d'imprécision allégué concerne globalement un problème de référence : quels sont les objets d'une attitude épistémique, et à quoi s'engage-t-on lorsque l'on construit une sémantique adaptée à ce genre de langage ? La majeure partie des critiques de Quine contre la logique modale, épistémique ou non, repose sur son critère d'engagement ontologique ainsi

---

[65] Dans le passage qui suit, on voit clairement que l'implication stricte désigne ce que Frege, tout comme Quine, qualifiait d'inférence : "Rien ne peut être inféré de prémisses fausses. Une simple pensée, qui n'a pas été reconnue comme vraie, ne peut pas être une prémisse. (...) De simples hypothèses ne peuvent pas être des prémisses." (G. Frege, "Lettre à Jourdain", in *Wissenschaftliche Briefwechslung*, G. Gabriel *et al.* (eds.), Felix Meiner, Hamburg, 1976, p. 118).

[66] [Quine 1975], p. 126. Quine accuse donc l'intuitionniste de confondre deux formes logiques distinctes, à savoir : (a) $K(p \vee \neg p)$ et (b) $Kp \vee K\neg p$. La logique modale épistémique de [Hintikka 1962] confirme Quine puisque, dans une logique modale normale, seule (a) est valide. Mais cela ne fait évidemment pas d'une logique modale un outil indispensable pour défendre la position de Quine.

que sur son interprétation référentielle (ou objectuelle) de la quantification. D'après lui, c'est sur le terrain de l'ontologie et du domaine de référence que reposent les difficultés de la logique modale en général. Puisque sa critique de la logique modale comme provenant d'une confusion des niveaux de discours est provisoire (Quine concédera la portée limitée de son grief, voir note 288), et l'argumentation concerna désormais le statut des attitudes épistémiques dans la notation canonique de Quine.

Naturalisme et *universalisme* sont deux qualificatifs qui peuvent être associés à la logique selon Quine. La logique a pour but de décrire les structures ultimes de la réalité : un énoncé logiquement vrai est un énoncé dont la vérité repose uniquement sur la signification des mots logiques et non sur son lexique ; en d'autres termes, seuls les mots logiques ont une occurrence essentielle dans une vérité logique et le contenu matériel des énoncés n'importe pas. La substitution uniforme est un critère pour juger si une vérité est logique ou ne l'est pas, ce qui devrait garantir une application sans restriction des formules abstraites et parfaitement hypothétiques de la logique.

Cela dit, le caractère abstrait et général des lois logiques repose sur un choix préalable quant à l'extension de la classe des mots logiques. Qu'est-ce qu'un mot logique et, surtout, qu'est-ce qu'il n'est pas et ne peut pas être ? La distinction entre les mots logiques et le lexique au sein d'un langage se faisait en logique traditionnelle en termes de catégorèmes et syncatégorèmes : les premiers sont des termes ayant un référent, tandis que les seconds ne font sens que dans leur combinaison à d'autres termes.[67] Mais une telle opposition ne servirait pas à Quine comme argument pour interdire la logique modale : les modalités sont considérées comme des syncatégorèmes. Dans une position extrême, la frontière entre mots logiques et mots non-logiques a été radicalement contestée par Tarski, lorsque celui-ci affirma que les frontières de la théorie sont arbitraires ou choisies par convention.[68] Sans aller jusqu'à cette conclusion, la position de Quine sur la question montre que sa conception de la logique n'a en tout cas rien d'apriorique : une extension du vocabulaire logique est permise, en témoignent les opérateurs unaires de la logique modale. Mais l'Américain s'oppose à une extension indéfinie des mots logiques, qu'il limite pour sa part aux connecteurs de la logique des fonctions de vérité (négation, conjonction, disjonction, conditionnel et quantificateurs) :

> Les vérités logiques au sens étroit sont des vérités qui ne font appel, essentiellement, qu'à des mots logiques. Or, à l'intérieur même de cette catégorie, il est possible d'interpréter la logique

---

[67] [Granger 1979] compare cette distinction traditionnelle à la distinction linguistique moderne entre *désignateurs* et *formateurs*, respectivement. Bien que cette distinction appartienne à un "fonds universel du langage", l'auteur prétend à leur "indistinction originaire" dans les langues naturelles : "même dans nos langues les plus familières, la séparation radicale de signes purement 'syntaxiques' et de signes purement 'sémantiques' est illusoire, et n'est effectivement réalisée que dans les systèmes formels" (p. 190).

[68] "On ne connaît aucun fondement objectif qui nous permette d'établir une frontière nette entre les deux groupes de termes. Il semble être possible d'inclure, parmi les termes logiques, certains de ceux qui sont habituellement considérés par les logiciens comme extra-logiques sans tomber dans des conséquences contraires à l'usage ordinaire. Dans le cas extrême, nous pourrions considérer tous les termes du langage comme logiques." (A. Tarski, *Logic, Semantics, Metamathematics*, 1956, p. 418). Une autre position sensiblement analogue est le Principe de Tolérance de Carnap, dont la doctrine linguistique de la vérité logique fut critiquée par [Quine 1954]. Pour une distinction entre les positions de Carnap et Tarski sur les critères de vérité logique, voir Achille Varzi, "On logical relativity", *Philosophical Issues* **12**(2002), pp. 197-219.

plus ou moins étroitement, en assignant des limites plus ou moins étroites au vocabulaire des mots logiques.[69]

Dans la suite de cette section, il s'agit d'expliquer ce qui incite Quine à rejeter les modalités dans sa philosophie de la logique :

> Dans ce secteur, la frontière se borne à séparer ceux qui acceptent de faire entrer 'nécessairement' dans un discours philosophique sérieux et ceux qui s'y refusent. Qu'on me permette une fois encore d'évoquer ma *sympathie* pour la position négative.[70]

Qu'est-ce qu'un discours philosophique "sérieux", et comment expliquer de telles affinités exclusives chez Quine vis-à-vis de la logique classique du premier ordre ?

Cette affinité est liée à la mission que Quine assigne à sa théorie logique : dans la lignée du Cercle de Vienne, une notation canonique a un rôle de première main, celui de déterminer par des règles de combinaison de signes les énoncés courants qui ont un sens et ceux qui n'en ont pas. La syntaxe logique effectue cette première tâche, puis la théorie de la connaissance prend le relais en déterminant parmi les énoncés bien formés ceux pour lesquels la science est en mesure d'apporter des critères de vérité ou de fausseté. Quine opte pour une méthode de type physicaliste, dans laquelle la signifiance d'un énoncé dépend de sa conformité au discours de la science physique, qu'il considère comme un modèle scientifique suffisant.[71] Cela étant posé, la structure ultime de la réalité dont parle Quine n'a rien d'a priori et se démarque à cet égard du réalisme logique des premiers écrits de Russell ou Wittgenstein. La réalité est l'image construite qui résulte des théories scientifiques, et la logique ne fait que décrire un cadre général pour organiser ce discours au sein du langage. Contrairement à la connotation transcendantale de l'idée de cadre *a priori* de la logique (voir note 103), Quine admet que des hypothèses scientifiques tirées de l'expérience puissent interférer avec les lois logiques :

> Une révision même de la loi du tiers exclu a été proposée comme un moyen de simplifier la mécanique quantique ; et quelle différence y a-t-il en principe entre un changement de ce genre et le changement par lequel Kepler a succédé à Ptolémée, Einstein à Newton, ou Darwin à Aristote ?[72]

---

[69] [Quine 1964], p. 193.

[70] [Quine 1964], p. 193-4 ; nous soulignons. L'auteur exprime également dans [Quine 1977] son point de vue en termes d'affinités électives : face à la logique modale aléthique, il dit se trouver "dans la position d'un chef juif préparant un jambon pour une clientèle respectable. L'analyticité, l'essence et la modalité ne font pas partie de ma nourriture" (p. 116).

[71] La procédure de traduction logique est donc semblable aux empiristes logiques, mais la différence réside dans l'analyse sémantique : pour évaluer dans une seconde étape de l'analyse un énoncé bien formé, le critère du vérificationnisme est abandonné par Quine. Voir [Quine 1961b] et son rejet du second dogme de l'empirisme, dans lequel la distinction entre les questions de langage et les questions de fait est supprimée Les suspicions de Quine vis-à-vis des modalités aléthiques sont liées en partie à cet arrière-plan épistémologique.

[72] [Quine 1961b], p. 43. On trouve d'autres passages évoquant le rejet de la distinction catégorielle entre vérités logiques (analytiques) et vérités empiriques (synthétiques) : la totalité des énoncés vrais est un "savoir gris pâle, blanc par les faits et noir par les conventions (...) je n'ai pas trouvé de raisons substantielles pour en conclure qu'on y trouve des fils parfaitement noirs, ou des fils parfaitement blancs" ([Quine 1954], p. 406). Une des conséquences de ce rejet est le pluralisme *de principe* de systèmes logiques : "La logique n'est en principe pas plus fermée aux révisions que la mécanique quantique ou la théorie de la relativité" ([Quine 1975], p. 149) ; "En principe, la logique n'est pas moins ouverte à la révision que la mécanique des quanta ou la théorie de la relativité. L'objectif est dans chaque cas un système du monde" ([Quine 1975], p. 149) ; "Je suis réaliste en

La logique de Quine se présente ainsi comme une logique révisable en principe mais non révisée en fait, préservée selon les besoins et les intérêts du discours scientifique et qui, pour cette raison, n'est pas universelle à proprement parler. Pour illustrer cette révisabilité de principe des lois logiques, [Quine 1961b] fit appel notamment à sa métaphore du champ de force pour expliquer la place de la logique au sein de nos connaissances : il y a une interdépendance entre toutes ces vérités –conséquence de son rejet de la distinction entre vérités analytiques et vérités synthétiques –mais les vérités logiques se situent au centre de champ de force et sont les moins affectées par le mouvement constant du modèle scientifique.[73] Le rejet de la distinction entre vérités analytiques et vérités synthétiques devrait favoriser le cas des modalités, mais l'auteur invoque un principe de *mutilation minimum* pour exprimer son conservatisme en logique : parce qu'il a pour fonction de représenter la classe des vérités les plus générales, le langage logique doit être modifié en dernier recours. Les énoncés modaux représentent-ils un cas de force majeure ? Pas pour Quine.

Le cadre logique de Quine est la logique classique du premier ordre avec identité ; ce langage formel lui semble nécessaire –contrairement aux logiques enrichies– et suffisant –contrairement aux logiques déviantes, ou affaiblies– pour servir de cadre linguistique au discours des sciences, auquel il subordonne toute forme d'énoncé pour décider de sa vérité ou sa fausseté. Si Quine a choisi cette logique pour paraphraser toute sorte d'énoncé, c'est qu'elle présente selon lui des avantages que les logiques non-classiques n'ont pas. La simplicité, tout d'abord : "La logique néo-classique des fonctions de vérité à deux valeurs et des quantificateurs est un chef d'œuvre de simplicité et d'efficacité" ;[74] la clarté, ensuite : "la logique intuitionniste n'a pas le caractère familier, la commodité, la simplicité et la beauté de notre logique : comme la logique de Birkhoff et von Neumann, la logique intuitionniste n'a pas la transparence d'une logique des fonctions de vérité à deux valeurs".[75]

Pour être admissible dans sa notation canonique, un énoncé doit donc être analysable en termes de fonctions de vérité. Or nous le savons, toute logique des modalités a un caractère intensionnel et ne satisfait pas cette condition. L'attitude de Quine sur ce point n'est pas uniforme : après avoir proposé plusieurs analyses logiques des modalités aléthiques et épistémiques, l'auteur laissera progressivement les premières de côté et s'intéressera plus

---

logique, comme Popper, mais je mets la logique et la physique au même niveau" (in "Comments on Popper", Yourgrau, Breck, Allen (ed.), *Physics, Logic and History*, N. Y. Plenum, 1970, p. 31).

[73] "Ou encore, pour changer d'image, la science totale est comparable à un champ de force dont les conditions limites sont l'expérience. Un conflit avec l'expérience à la périphérie donne lieu à des réajustements à l'intérieur du champ. Les valeurs de vérité doivent être redistribuées pour certains de nos énoncés. Une réévaluation des énoncés entraîne une réévaluation de certains autres, du fait de leur interconnexions logiques –les lois logiques étant à leur tour simplement d'autres énoncés du système, certains autres éléments du champ." ([Quine 1961b], p. 42).

[74] [Quine 1964], p. 203. A d'autres endroits, Quine reproche notamment aux logiques déviantes de compliquer inutilement le comportement logique de la négation ; ce qui, pour lui, fait de la négation classique une constante simple et familière est sa propriété de *symétrie* : la négation du vrai est le faux, et inversement. Il y a évidemment d'autres types de négation, et la propriété d'involution de la négation classique ne constitue qu'un aspect des diverses négations étudiées en logique algébrique. Le choix de Quine est d'ordre quasi-esthétique, ici.

[75] [Quine 1975], p. 129-30. L'auteur fait allusion ici à la mécanique quantique, dont les expérimentations sur les particules micro-corpusculaires exigent apparemment une révision de certaines lois logiques (non-contradiction, associativité, distributivité). L'enjeu pour Quine est de montrer que sa paraphrase peut venir à bout de ces difficultés.

sérieusement au cas des attitudes propositionnelles. Nous reviendrons plus tard sur la difficulté créée par l'analyse logique des modalités, notamment dans un contexte quantifié, et examinerons le verdict de Quine relativement à leurs résultats (voir section 2.3.1.2).

Au final, les exigences philosophiques de Quine l'ont incité à rejeter à double titre la logique épistémique : non seulement les attitudes propositionnelles n'entrent pas dans le cadre de sa notation canonique, mais les opérateurs épistémiques sont des notions trop 'colorées' pour maintenir le caractère 'obvie' des vérités logiques. Bien que l'évidence des théorèmes logiques ne soit pas synonyme chez lui d'évidence apriorique ou de vérité valable dans n'importe quel langage, Quine n'était cependant pas favorable à l'idée d'une logique composée de règles conventionnelles ou inventées. L'évidence des vérités logiques doit s'enraciner dans la pratique du langage dont la logique est un cadre, une garantie minimale du discours pourvu de sens. C'est ce manque d'évidence dans la *pratique* du langage que Quine reproche aux vérités des logiques modales, parmi lesquelles des théorèmes aussi convaincants soient-ils que la clause de succès (le schéma **T**-valide $\Box\phi \rightarrow \phi$). Si l'on applique la définition grammaticale de la vérité logique selon Quine, les vérités logiques épistémiques sont valables puisqu'elles persistent après substitution des termes du lexique et reposent sur l'occurrence essentielle de mots logiques. La différence entre Quine et Hintikka réside non pas dans leur définition de la vérité logique, mais dans leur choix de la liste des mots logiques.

Y a-t-il un argument décisif pour justifier l'austérité de la philosophie de la logique de Quine ? Nous avons proposé les critères de simplicité du langage et de réductibilité au discours du langage scientifique. Ajoutons également l'argument de l'universalité des mots logiques classiques, par opposition aux mots des logiques modales : une vérité logique classique est valable dans toutes les structures de Lewis **S1-S5**, contrairement aux théorèmes valides dans certaines structures seulement. Mais cet argument de l'universalité ne saurait convaincre un partisan des logiques affaiblies, dans lesquelles certains théorèmes classiques ne valent plus (le tiers exclu, notamment). Seule une conjonction des critères d'universalité, de simplicité et de scientificité semble justifier la position conservatrice de Quine en logique. Chacun de ces trois arguments séparés est insuffisant pour justifier le rejet de la logique modale, malgré les positions défendues par d'autres philosophes logiciens tels que Gonseth et Granger : leur distinction entre théorie et langage suppose que les logiques modales portent sur un objet d'étude restreint, tout comme la théorie des nombres ou la géométrie. Ainsi [Granger 1979] :

> le développement des logiques modales, des logiques déontiques, des logiques temporelles, des logiques de la causalité fait apparaître les difficultés que rencontre toute tentative d'ignorer les frontières entre l'apophantique pure et simple et les théories d'objets. Dans quelle mesure est-il possible de construire une *théorie des propositions* modales, déontiques, temporelles, causales qui mérite pleinement le nom de logique? Ne s'agit-il pas alors bien plutôt de la mise en forme de systèmes de propriétés caractérisant des mondes physiques ou mentaux ?[76]

[76] [Granger 1979] répète ailleurs sa critique des logiques modales, qu'il considère comme des théories d'objets particuliers : "à considérer de près le déploiement de ces calculs, depuis Lewis jusqu'aux récentes interprétations sémantiques, il nous semble que leur caractère proprement logique de mise en forme des règles d'un symbolisme propositionnel s'est progressivement effacé devant un caractère *ontologique* : plutôt que de véritables logiques, ils seraient alors des théories proprement mathématiques portant sur de nouveaux *objets* –les temps et les modes

Mais qu'est-ce qu'une apophantique pure, sinon un langage idéal dont les vérités logiques seraient incontestables ? Cette affirmation suppose que les conditions de vérité d'un énoncé sont invariantes et permet de distinguer en conséquence les mots logiques, les plus abstraits, des autres termes du lexique. Une telle distinction est-elle graduelle, ou catégorielle ? Quine refuse la distinction catégorielle, nous l'avons vu, et nous constaterons également que la définition des mots logiques classiques est contestée dans les logiques relevantes (pour l'implication), paraconsistante et intuitionniste (pour la négation), quantique (pour la conjonction), etc. Il n'y aurait finalement de différence que dans le choix des langages formels pour définir ces constantes : la réduction de plusieurs concepts philosophiques à des opérateurs formels montre que les notions de connaissance, de temps ou d'obligation ne sont plus des objets 'colorés' d'expérience lorsqu'ils sont définis au sein de structures abstraites. A cet égard, le reproche adressé à la logique modale semble injustifié : le connecteur de négation est certes plus général que l'opérateur de croyance, donc davantage abstrait de l'expérience que ce dernier lorsqu'il est utilisé dans les langues naturelles, mais les deux notions sont traitées à la même échelle au sein de langages formels.

Considérant ce que nous avons vu dans cette section, la différence essentielle entre partisans et opposants à la logique modale repose globalement sur la portée d'application qu'ils accordent aux vérités logiques : relative à un contexte ou un langage formel pour les uns, insensible au contexte de discours et universelle, pour les autres. Mais le cas de Quine a montré qu'universalisme ne rime pas forcément avec apriorisme, puisque son rejet de la distinction catégorielle entre vérités logiques et vérités empiriques le rapproche davantage des partisans du pluralisme en logique. Et cependant, pluraliste il ne le fut pas. L'évidence était garantie par les trois premiers philosophes précités (Frege, Russell, Wittgenstein), mais sans explication sur la nature des vérités logiques. A défaut de trouver un argument décisif pour distinguer la nature des termes logiques et extra-logiques, la condition pour admettre une logique épistémique modale repose donc sur d'autres arguments que la signification des connecteurs logiques : ceux-ci peuvent varier en principe, reconnaît Quine, mais son mot d'ordre méthodologique est la simplicité théorique et, pour cette raison, le dissuade d'admettre le vocabulaire de la logique modale.

Quoi qu'il en soit, nous disposons désormais de données suffisantes pour montrer que la logique épistémique ne brouille pas les frontières entre logique et psychologie, en dépit de possibles apparences du contraire. Comment expliquer ce qui distingue la logique de la

---

– tout comme la géométrie et l'arithmétique construisant et décrivant des nombres et des espaces." (p. 87) Cette critique s'applique aussi bien à la logique des prédicats du premier ordre : l'objet formel n'est plus quelconque dès lors qu'il est décrit dans un modèle, par le biais de structures d'interprétation. Le parallèle établi jusqu'ici entre logique du premier ordre et logique modale le confirme.

La définition de la logique comme physique de l'objet quelconque, due à Gonseth, repose sur la même distinction catégorielle entre la logique comme langage pur et les autres sciences, qui sont des théories spécifiques : "Les lois qu'elle [la Logique] formule trouvent leurs réalisations naturelle dans le domaine des objets concrets ; elles y prennent la signification de lois naturelles très primitives. (...) [elle] n'est qu'un chapitre de la physique, un des tout premiers, si ce n'est pas le premier : celui qui traite de l'existence des objets de nature quelconque" (F. Gonseth, "La logique en tant que physique de l'objet quelconque", in *Actualités Scientifiques et Industrielles* n° 393, Paris, 1936, p. 10).

psychologie, et quelle place une logique de la croyance occupe-t-elle entre les deux disciplines ? Pour montrer que la logique épistémique n'est pas réductible à un psychologisme logique, pas plus qu'à un logicisme psychologique, revenons une fois pour toutes sur cette possible accusation, largement due au fait que l'opérateur modal de croyance désigne un état mental.

2.1.3 Contre l'accusation de psychologisme

Le psychologisme affirme que les opérations logiques sont fondées sur des faits psychologiques ; quel est ce sens de ce réductionnisme, et en quoi le psychologisme logique concerne-t-il la logique épistémique ? Ce débat concernera avant tout la *logique doxastique*, c'est-à-dire la logique des croyances, puisque le verbe 'croire' renvoie clairement à un état psychologique.[77] Pour rejeter toute accusation de psychologisme à son égard, et bien qu'elle porte partiellement en effet sur des actes de pensée (les croyances), nous allons dissocier deux niveaux d'étude des actes de pensée, réservant à la psychologie l'étude de leurs *causes* et à la logique l'étude de leur *expression*. Si la logique sert simplement à exprimer des actes de pensée, elle ne produit aucune théorie causale à leur sujet et ses calculs formels ne feront que suivre les clauses de vérité de la logique formelle de Tarski. En ce sens, la notion de vérité demeurera première dans une analyse logique. Pour montrer que la logique épistémique formalise les relations entre des actes de pensée, et rien de plus, la comparaison entre 'croire' et 'tenir pour vrai' sera mise en évidence. Pour défendre cette distinction entre logique épistémique et psychologisme, nous nous arrêterons notamment sur le cas de la sémantique des croyances de [Ellis 1976] et montrerons que ses prises de position 'psychologistes' ne sont pas conformes à notre acception du terme. La confusion porte avant tout sur ce que l'on entend par l'appellation de 'lois de la pensée'.

2.1.3.1 Une confusion sur la notion d'explication

Nous ne prétendrons pas donner une définition définitive du psychologisme logique, mais discuter sa relation avec la logique épistémique : peut-on associer la logique épistémique à une forme quelconque de psychologisme logique ? Si l'on part de la définition provisoire selon laquelle le psychologisme logique considère les opérations logiques comme l'expression de *lois de la pensée*, le résultat risque de ne pas être plus clair si l'on songe dans le même temps à la notion de pensée selon Frege. Distinguer chez ce dernier la logique de la psychologie, c'est distinguer entre l'étude de la vérité pour elle-même et l'étude de la vérité pour nous, les sujets pensants. La psychologie empirique et la psychologie cognitive, par exemple, décrivent les processus au cours desquels un agent se forme une opinion et finit par reconnaître quelque chose comme vrai ou faux, mais elles ne disent rien au sujet les propriétés mêmes du vrai, et c'est sur cette base que Frege a justifié sa campagne anti-psychologiste : la distinction entre *opinion vraie* et *vérité*, que nous avons distinguée plus tôt en termes de

[77] La chose n'est pas aussi claire pour le verbe 'savoir' : la connaissance renvoie-t-elle à un état mental ou à une propriété objective, indépendante des attitudes du sujet pensant ? Sur le lien entre connaissance et psychologie, voir la section 2.2.42 sur la question de la connaissance par introspection. Le fait que l'on attribue des connaissances sans état mental correspondant (certitude ou croyance) semble indiquer que 'savoir' et 'croire' ne se situent pas vraiment sur le même plan ; on parle de 'logique épistémique' et d''attitudes' épistémiques au sens large mais, au sens étroit, 'savoir' n'est donc pas toujours l'expression d'une attitude.

pensée en acte et de pensée pure. La distinction entre logique et *psychologie* est plus ou moins comparable en logique épistémique à la distinction entre logique et *ontologie* en logique classique : les opérations de la logique classique ne déterminent pas la vérité ou fausseté des atomes tels que p, elles déterminent seulement ce qui s'ensuit de la vérité ou fausseté hypothétique de cet atome. Par exemple, la logique dit que *si* 'p' est vrai *alors* 'p ou q' est vraie. De façon similaire, les opérations de la logique épistémique ne déterminent pas la vérité ou fausseté des formules telles que $B_a p$, elles déterminent seulement ce qui s'ensuit de la vérité ou fausseté de $B_a p$. Par exemple, la logique épistémique dit que si '*a* croit que p' est vraie alors '*a* croit que p ou q' est vraie. La logique classique ne détermine donc pas plus la valeur des faits ontologiques que la logique épistémique ne détermine celle des faits psychologiques : toutes deux décrivent uniquement des relations hypothétiques entre plusieurs types de faits.

Présentée comme telle, la logique doxastique n'a effectivement rien à voir avec le psychologisme logique puisqu'elle n'explique absolument pas les conditions dans lesquelles un agent parvient à croire quelque chose et le considère comme vrai. Et cependant, cette logique d'attitude semble établir une passerelle entre le discours logique sur la vérité en tant que telle et le discours psychologique sur l'opinion : elle ne décrit pas les processus physiques qu'engage un acte de croyance, certes, mais elle introduit des relations d'inférence entre des actes psychologiques. Est-ce déjà s'engager dans une forme de *psychologisme modéré* que de parler de croyances, d'attitudes propositionnelles, d'actes de pensée même sans les décrire ? Si le psychologisme logique a trait à la vérité comme description d'ensembles de croyances, on peut considérer la logique épistémique au sens large comme en faisant partie, mais à condition que la caractérisation des modalités épistémiques (par des systèmes d'axiomes ou des modèles sémantiques) constitue une sorte de description ou de justification des relations de vérité entre des formules.

A l'instar de ce genre d'amalgame, certaines présentations peuvent porter à confusion en ce qui concerne la frontière entre logique, logique épistémique et psychologie. Lorsque Frege relègue l'acte de saisie d'une pensée dans le domaine de la psychologie, il associe à cet acte la notion de croyance. [Engel 1996] illustre ce problème au sujet de la distinction entre sens et référence :

> Frege (...) entend se préoccuper seulement de la nature des pensées et de leur contenu, en tenant comme un 'mystère' l'acte de saisie d'une pensée, et en reléguant précisément à la psychologie la tâche d'en élucider la nature. Son attitude consiste à s'occuper exclusivement de *ce qui est connu*, en ignorant le *comment* de la connaissance.[78]

Mais il ne faut pas en conclure pour autant que parler de la croyance fait de la logique épistémique une étude inévitablement psychologique, pas plus que parler de vérité et de fausseté ne fait de la logique classique une étude ontologique. La différence essentielle entre logique et psychologie réside non pas dans leur objet d'étude mais dans le niveau d'analyse de cet objet, et la citation ci-dessus ne fait qu'étayer la présentation des croyances proposée par la logique modale épistémique : la logique étudie la vérité, la psychologie étudie l'opinion.

[78] [Engel 1996], pp. 106-107.

Plus précisément : la logique des croyances étudie les relations déductives entre *ce* qui est cru ou *ce* qui est connu, c'est-à-dire les relations entre ce que Frege appelait des pensées et qui a été interprété en termes de propositions ; or puisque ces propositions ont un contenu déterminé en termes de vérité ou de fausseté, ces relations n'ont rien à voir avec la manière dont le sujet pensant 'saisit' ces propositions. Par contraste, la psychologie étudie *comment* des énoncés sont crus, et cet aspect de la vérité débouche sur les opinions ou croyances en vertu desquelles des énoncés sont considérés comme vrais ou faux. De là l'impression que la logique épistémique puisse servir d'intermédiaire entre les deux disciplines, puisqu'elle étudie les relations entre des croyances vraies ou fausses. Si l'on reprend le passage ci-dessus au sujet de la saisie de la pensée fregéenne, la tâche de la logique épistémique consiste non pas à élucider la nature de cet acte de saisie d'une pensée mais à s'occuper de *ce* qui est saisi, tout en ignorant le *comment* de cette saisie. Si la logique épistémique étend le domaine de discours de la logique sans tomber dans le psychologisme, c'est donc parce qu'elle continue à traiter de relations d'inférence entre des entités *linguistiques*, mais des entités moins abstraites que les propositions de Frege ou les *Sätze an Sich* de Bolzano. La section consacrée à la logique épistémique quantifiée (section 2.3) en fera l'illustration, lorsque le principe d'individuation des objets de discours symbolisera l'acte singulier de saisie de la référence dans un énoncé et servira à rendre compte du cas de l'opacité référentielle. Nous en revenons ainsi à la distinction entre les *propositions* de Frege, abstraites et indépendantes des locuteurs, et les *énoncés* de Hintikka, dont le sens ou contenu propositionnel n'apparaît pas toujours en toute transparence. Dans les deux cas, l'analyse est d'ordre logique parce qu'elle repose sur les mêmes règles de préservation de la vérité : la logique classique s'occupe uniquement de transitions entre des propositions ou pensées vraies, symbolisées par des formules de type ϕ ; la logique épistémique s'occupe également de transitions entre des propositions mais introduit de plus des énoncés ou opinions vraies, symbolisées par des formules de type $K\phi$ ou $B\phi$. Un exemple caractéristique sera l'axiome du système **T** : $K\phi \rightarrow \phi$, qui symbolise la relation logique entre un énoncé épistémique et une proposition. Les lois de l'être vrai sont respectées dans les deux systèmes logiques, seul le *porteur de vérité* est modifié (proposition dans un cas, énoncé dans l'autre).

Pour reprendre l'analogie du télescope de [Frege 1892], où la lune est à la dénotation [*Bedeutung*] ce que l'image de la lune est au sens [*Sinn*] et l'image rétinienne à la représentation [*Vorstellung*], la logique épistémique ne porte pas son regard sur la lune, ce qui est le cas de la logique classique, mais sur l'image de la lune produite dans la lunette. Quant à la psychologie, son étude porte uniquement sur l'image rétinienne et ne se mélange pas aux points de vue objectifs du sens et de la dénotation. En d'autres termes, la logique classique propose une logique de la référence (ou de la dénotation) tandis que la logique épistémique propose une logique du sens.

Sur la base de la distinction sémantique entre sens et représentation, la distinction est faite entre psychologisme logique et logique épistémique. Dummett semble confirmer cette distinction, lorsqu'il marque une séparation entre théorie de la signification et sémantique formelle :

> Une explication sémantique, formulée entièrement à l'aide du concept de référence, montre fort adéquatement la façon dont la valeur de vérité est déterminée à partir des mots qui la constituent

> et de la façon dont ils sont composés. L'explication sémantique pèche cependant, car elle ne revient pas suffisamment en arrière : elle postule une association entre chaque symbole primitif et un référent approprié, sans nous dire *comment* cette association est établie. En logique, cela n'est pas nécessaire ; pour une théorie de la signification, c'est essentiel.[79]

En somme, la seule différence entre logique classique et logique modale épistémique réside dans leur niveau d'analyse des constituants du langage : le langage formel de Hintikka symbolise l'acte psychologique de croyance, introduisant ainsi dans son langage-objet ce que Frege excluait totalement du sien ; mais le fait de symboliser l'acte de saisie de la pensée (par un opérateur modal) n'équivaut pas à son explication.

La différence entre la symbolisation d'un acte psychologique et son explication, essentielle pour dissocier logique épistémique et psychologisme logique, apparaît également dans un passage de [Mohanty 1981] sur les *fonctions individuantes* de Hintikka, c'est-à-dire les fonctions par lesquelles une extension est associée à un symbole dans un monde possible. Ces fonctions correspondent à ce que Russell appelait plus haut des 'fonctions de signification', et celles-ci symbolisent l'acte de dénoter ou d'associer un référent à un signe, ce qui équivaut à un mode de donation de la référence qui caractérise donc le 'sens' de Frege. Cela étant, ces fonctions ne décrivent pas le processus intentionnel de signification (au sens propre du verbe 'signifier' : mettre une pensée en signes) mais se contentent de le conceptualiser, en termes d'objet O et d'états intentionnels W résultant d'un processus noétique N. Et puisque le processus psychologique concerne ce qui se passe en amont de l'assignation de référence, il n'appartient pas au champ d'étude logique de la sémantique relationnelle :

> Le fait est que, dans l'explication de Hintikka, la fonction de signification en question est une fonction donnée en extension, i.e. simplement un ensemble de paires ordonnées $\{\langle W_1, O_1\rangle, \ldots, \langle W_i, O_i\rangle, \ldots\}$. Tout ce qu'elle donne, c'est une liste de ce qui est 'sélectionné' dans chaque monde par N. Nous n'avons pas d'explication de la *raison* pour laquelle O_i est l'image de W_i *sous* N, i.e. *pourquoi* O_i est sélectionné, nous n'avons rien en vertu de quoi O_i est l'objet approprié dans W_i. Tant que notre souci est simplement de fournir une sémantique pour différentes sortes de logiques intensionnelles, l'absence d'un authentique concept de sens n'importe pas. Mais pour les besoins de la phénoménologie, qui est de fournir des structures descriptives des actes de conscience, parmi lesquels lesdites attitudes propositionnelles, la fonction extensionnelle est bien trop inadéquate.[80]

La théorie de la signification souhaitée par Dummett, en vue de compléter une explication sémantique référentielle qui "ne revient pas suffisamment en arrière" à ses yeux, correspond à ce que les phénoménologues recherchent pour leur part par le biais de la *noétique* : une théorie de la constitution des objets par le biais d'actes de visée intentionnelle, dont le résultat est la production d'un domaine d'objets mais qui porte sur les expériences subjectives vécues durant le processus de constitution. La noétique est donc une étude des sources de justification

---

[79] [Dummett 1973a], p. 92 ; nous soulignons. Ailleurs, l'auteur sépare à nouveau conditions de vérité (logique) et conditions de saisie de la vérité (théorie de la signification) : "De même lorsqu'il parle du sens (*Sinn*), il le caractérise comme *ce qui est compris* quand on comprend une expression, mais il n'a rien sur le *comprendre*. Mais si la saisie d'une pensée ou d'un sens est une forme de connaissance ou de compréhension, comme elle semble bien l'être, nous aimerions bien savoir ce qui la distingue de la pensée et du sens eux-mêmes, comme contenus de pensées" (M. Dummett : *Frege, Philosophy of Mathematics*, Duckworth, London (1991), p. 115)

[80] [Mohanty 1981], p. 95.

de la constitution des objets ; elle est constitutive du domaine de référence du sujet et, en cela, n'intéresse pas le sémanticien logique puisque celui-ci présuppose un domaine d'objets initial. L'extension d'un domaine de référence peut croître ou décroître selon les besoins explicatifs, certes, mais la production d'un objet en tant que tel reste du ressort des sciences de l'esprit dont fait partie la psychologie.[81] [Hintikka 1981a] insiste lui-même pour distinguer les deux domaines de compétence. Visiblement, sa description des différents modes d'individuation des objets selon le type de modalité en vue (continuité spatio-temporelle pour les opérateurs temporels, continuité mémorielle pour les opérateurs de perception, etc.) a été comparée à la noétique par Mohanty ; or une telle interprétation *informelle* de la quantification entre des mondes possibles (la 'quantification croisée', ou 'transquantification') n'aboutit pas à de la noétique : les conditions sémantiques pour individuer des objets dans des mondes possibles ne font que symboliser des types de visée dans un monde possible, mais elles ne disent rien sur leur origine :

> il est parfaitement vrai que les fonctions de signification ne décrivent pas en détail la manière dont l'objet d'un acte (ou la référence d'une expression linguistique) est déterminée en différentes circonstances possibles. (...) Contrairement à ce que Mohanty prétend, je n'ai jamais déclaré que les fonctions de signification donnent automatiquement un tel processus opératoire en tant que fonctions, et je ne crois pas non plus qu'elles le font.[82]

Tout au plus pouvons-nous symboliser ces actes de la conscience dont l'assignation de référence est le résultat, au sein d'une logique modale épistémique. Le processus d'*internalisation*, c'est-à-dire l'introduction de notions métalinguistiques au sein du langage-objet, permet une description affinée de la pensée, et la logique épistémique quantifiée en fut l'illustration puisque sa méthode de valuation symbolise le mode de donation de la référence [*Gegebenheitsweise*], c'est-à-dire le sens ou *Sinn* frégéen. A un certain égard, la méthode d'internalisation a pour but de préciser les conditions de vérité d'une formule selon le contexte de discours : en enrichissant le vocabulaire du langage-objet dans son langage formel, le logicien est en mesure de donner plus de détails à son analyse sémantique et d'expliquer *a posteriori* les raisons pour lesquelles un énoncé n'est pas compris par un locuteur, c'est-à-dire ne remplit pas toutes les conditions de satisfaction requises. L'opacité référentielle, présentée comme le résultat d'une multiplicité de la référence à partir de [Hintikka 1957], exprime l'ambiguïté en termes d'individuation non satisfaite et permit notamment à [Hintikka 1969] de considérer le sens et la référence comme les deux aspects référentiels d'une unique théorie générale de la signification comme assignation de référence

---

[81] [Mohanty 1979] décrit la logique transcendantale de Kant et la noétique de Husserl en termes de 'logique épistémique' : "La logique transcendantale serait alors une théorie de notre connaissance *a priori* des objets. Elle est une logique épistémique" (p. 181) ; "Quelle est la relation entre logique formelle et *Noétique* (ou logique épistémique) ? (...) Bien que la logique formelle puisse être 'noétiquement naïve', si elle doit être critique et donc philosophiquement sophistiquée elle doit être en étroite relation avec la logique épistémique" (p. 49). Mais il fait ainsi de manière métaphorique, pour opposer dans les deux cas une logique des objets ou logique 'réelle' à une logique des formes ou logique 'pure'. Sur le lien entre logique épistémique quantifiée et intentionnalité, voir notamment les articles de [Hintikka 1975b] consacrés à l'interprétation de Husserl, en particulier "Les intentions de l'intentionnalité", pp. 143-179. D'autres rapprochements ont été faits entre Husserl et la logique épistémique, notamment par [Beyer 2001] qui compare la signification de Husserl à l'usage attributif de la référence de Donnellan.

[82] [Hintikka 1981a], p. 115.

(dans un monde possible ou dans le monde 'réel', respectivement). Mais il y a d'autres actes épistémiques que l'assignation de référence, parmi lesquels l'acte de raisonnement et les méthodes de preuve qui l'accompagnent. Pour décrire des enchaînements de pensées produites dans le cadre d'une inférence, la question est de savoir si une logique modale épistémique est capable de retranscrire clairement les opérations déductives effectuées par un agent au sein d'un langage formel.

Pour exprimer ces actions épistémiques ou cognitives du sujet lors d'un raisonnement, [van Benthem 1991] oppose deux méthodes d'exposition de la pensée logique. La méthode des logiques explicites ou *extrinsèques*, d'une part, dans laquelle l'acte de pensée est exprimé par un opérateur modal qui gouverne et s'ajoute donc aux pensées frégéennes (les énoncés, vrais ou faux); la logique épistémique de [Hintikka 1962] fait partie de cette tradition, puisque les conditions de vérité des pensées (énoncés) restent classiques (frégéennes) et sont simplement complétées par celles d'opérateurs épistémiques. La méthode des logiques implicites ou *intrinsèques*, d'autre part, dans laquelle l'acte de pensée ne se surajoute pas à la pensée mais est intégré à elle ; le résultat est une révision des connecteurs logiques.[83] La théorie des types constructive de Per Martin-Löf, évoquée ci-dessous, est une logique de type intrinsèque dans la mesure où les formules de base ne symbolisent ni une proposition ni une pensée frégéenne (un énoncé) mais un acte de jugement, et cet acte devient l'unité fondamentale de la signification. Quelle méthode faut-il préférer pour exprimer des actes de pensée, et le choix dépend-il de critères théoriques ou philosophiques ? Une remarque de [Hinzen 2000] semble confirmer ce point de vue, lorsqu'il explique qu'une logique épistémique extrinsèque s'avère incapable d'exprimer le lien *anti-réaliste* entre un acte de raisonnement (la vérification) et son objet (l'énoncé vrai). S'il en est ainsi, c'est parce que le langage-objet d'une logique modale épistémique extrinsèque, même lorsque celle-ci est d'aspiration anti-réaliste, sépare d'emblée action et discours portant sur l'action :

> On a étudié des affirmations anti-réalistes sur la signification et la vérité dans une sémantique des mondes possibles adaptée à une logique modale épistémique. (...) le fait de traiter les attitudes épistémiques comme des opérateurs propositionnels modaux *explicites* nous engage dans l'idée que l'aspect épistémique de la logique, qui constitue une préoccupation intuitionniste centrale, ne sera pas reflété dans notre notion de contenu propositionnel ou notre notion de vérité. Elles ne seront pas *intégrées dans* la logique.[84]

Pour satisfaire cette exigence théorique des intuitionnistes, la théorie des types constructive de Per Martin-Löf a pour but de remplacer la notion objective de proposition par celle de jugement (un énoncé dont la vérité n'est pas seulement stipulée mais prouvée, ou vérifiée), définissant la proposition comme l'ensemble des preuves de son énoncé. Pour tenir compte de

[83] La logique modale épistémique est décrite par [van Benthem 1991] comme "une superstructure modale au sommet d'une sémantique classique ordinaire", peu soucieuse de "réanalyser cette dernière en profondeur. Pour introduire une distinction, le paradigme Hintikka-Halpern est un paradigme de l''épistémique extrinsèque', lequel n'affecte pas la sémantique classique standard, en témoigne se définition clef de la vérité (...) Par opposition, une 'épistémique intrinsèque' plus radicale ne considérait pas la sémantique classique comme allant de soi et réanalyserait la notion de vérité dans son ensemble, à la lumière de considérations épistémiques" (p. 6).
[84] [Hinzen 2000], p. 283. L'auteur fait allusion ici à l'approche extrinsèque de [Williamson 1990,1992], plaidant à l'opposé en faveur d'une formulation intrinsèque ou constructive de l'anti-réalisme sémantique.

cette distinction, le langage-objet de la théorie des types constructive prévoit d'enrichir en conséquence les règles de formation d'une proposition (pensée) : celle-ci n'est plus une entité abstraite préfabriquée mais le résultat d'un acte, d'une construction de preuves (vérification) en termes de jugements.[85] Les langages-objets des logiques extrinsèques et intrinsèques répondent quoi qu'il en soit à un même souci initial : exprimer de manière plus détaillée la constitution du sens des signes propositionnels dans un langage, sans prétendre pour autant que les moyens symboliques apportent une explication de type causal à cette constitution du sens.[86] Ils l'expriment, et c'est ce rôle que le logicien confie à son langage formel. Rien n'est donc dit sur les conditions de constitution du sens, encore une fois. Or si la psychologie traite de ces conditions préalables, alors la logique épistémique ne révèle en aucun cas une forme de psychologisme.

En résumé, la logique épistémique n'est pas un psychologisme tel que nous l'entendons en un sens fort, c'est-à-dire une réduction des opérations logiques à des événements psychiques ou des faits de la conscience. En revanche, la logique épistémique se rapproche en quelque sorte de la psychologie lorsqu'elle introduit dans son langage-objet des notions psychologiques, sans prétendre toutefois à leur explication causale. Pour rendre justice à ce rapprochement du logique vers le psychologique, citons un passage de [Engel 1996], dans lequel l'auteur évoque un psychologisme faible susceptible de définir le statut particulier de la logique épistémique et des logiques d'attitudes propositionnelles en général :

> il me semble qu'il y a, entre le psychologisme fort ou réductionniste et l'antipsychologisme radical, place pour un *psychologisme non réductionniste, faible ou minimal.* La tâche de ce psychologisme consiste à caractériser et décrire précisément les interactions entre les propriétés normatives du monde 3 et les propriétés naturelles du monde 2. Pour indiquer le type de région à laquelle appartient cette enquête, nous pouvons dire que le psychologisme non réductionniste s'intéresse à ce qui se produit dans ce que j'appellerai le *monde 2 ½*, la région intermédiaire où propriétés normatives et propriétés naturelles entrent en contact. Il a comme objectif non pas de décrire *de facto* les types de processus et de représentations psychologiques qui conduisent à la formation des connaissances, mais de spécifier les types de processus qui devraient être ceux qui gouvernent la vie psychologique des agents humains.[87]

Entre le monde 3 des pensées pures ou entités platoniciennes de Frege (les *Sinne*), et le monde 2 des pensées en acte ou événements de la conscience qu'étudie le psychologue, une place est ainsi trouvée pour situer la logique modale épistémique dans un monde 2 ½, un

---

[85] L'expression 'a : A' signifie que l'ensemble des preuves ou jugements 'a' donne la proposition A. Sur cette base de données, la règle de formation d'un énoncé conjonctif devient : $\frac{a:A \quad b:B}{(a,b):A \wedge B}$
Pour les motivations philosophiques de la théorie des types constructive, voir [Martin-Löf 1996]. On aperçoit une affinité entre cette théorie et la logique modale illocutoire de la section 3.3.3 : toute construction propositionnelle a:A revêt la forme d'un jugement ϕ de type 'ϕ est vrai', soit $A\phi$.

[86] La logique épistémique exprime le résultat de l'opacité référentielle (un échec de l'individuation) mais ne dit rien sur les raisons de cet échec, encore une fois. Elle n'est pas en mesure d'expliquer par exemple les *idiolectes*, c'est-à-dire la raison du choix d'un mode de donation d'une référence parmi d'autres : "Comment ces sens 'communs' ou 'publics' se distinguent-ils des, ou éventuellement se rattachent-ils aux, sens privés que les locuteurs peuvent entretenir dans leurs idiolectes ? Cette question peut paraître absurde si l'on est convaincu, comme on peut l'être à la suite des arguments de Wittgenstein contre la notion de 'langage privé', qu'il ne peut y avoir de sens autres que publics, mais là aussi l'élucidation de la différence entre ce que connaît un locuteur individuel et les sens publics dans le langage devient pressante, à partir du moment où l'on admet que cette différence existe. Quelle est la relation entre les idiolectes et le langage commun ? ([Engel 1996], p. 117).

monde de prescriptions dans lequel sont présentées les interactions entre les propriétés normatives du monde 3 (les lois de l'être vrai) et les propriétés naturelles du monde 2 (les actes de pensée) ; le résultat de cette combinaison donne une étude sémantique des conditions de vérité des énoncés de croyance, c'est-à-dire des opinions ou jugements. Pour reprendre une remarque de [Hintikka 1997c] sur l'identité entre 'sens' et 'référence dans un modèle', notons toutefois que les mondes 3 et 2 ½ ne sont pas deux mondes distincts mais un seul et même monde observé de deux points de vue différents : le point de vue des objets, chez Frege et les partisans de la logique comme langage universel ; le point de vue du sujet, chez Hintikka et les partisans de la logique comme calcul. Ces deux perspectives et conceptions rivales de la logique seront détaillées plus bas (section 2.1.4.2).

Sur la base de ce psychologisme faible, nous allons examiner la sémantique de [Ellis 1976] et son projet de fonder la logique sur des bases épistémologiques. Pourquoi décrit-il son projet comme un projet psychologiste, et dans quel sens faut-il l'entendre ?

2.1.3.2 Une confusion sur la notion de vérité

Dans le compte-rendu d'un ouvrage de Brian Ellis consacré à la théorie des croyances rationnelles, [Gardenförs 1980] associe lois de la pensée et psychologie :

> Une thèse majeure du livre est que les lois de la logique sont des *lois de la pensée*. Ces lois 'gouvernent la structure de systèmes de croyance idéalement rationnels pour des langages idéalisés, qui modèlent (...) des systèmes de croyance humaine ordinaires pour les langues naturelles' (p. v)'. Ainsi, la logique devient *une branche de la psychologie*. Ellis insiste beaucoup à dire que cette forme de *psychologisme* ne doit pas être confondue avec le fondationnalisme, rejeté par Popper, qui affirme que les affirmations de base de la science sont justifiées par l'expérience.[88]

Or on sait avec [Frege 1892] que l'expression de lois de la pensée est ambiguë, étant donné le sens ambigu de la notion de 'pensée' :

> On comprend peut-être l'expression 'lois de la pensée' comme 'lois de la nature', entendant par là l'aspect général du procès psychique de la pensée. Une loi de la pensée prise en ce sens serait une loi psychologique. A partir de quoi l'on pourrait croire que la logique traite du déroulement psychique de la pensée et des lois psychologiques auxquelles il est soumis. Mais on méconnaîtrait la tâche de la logique en ne donnant pas à la vérité la place qui lui est due. (...) L'opinion fausse et l'opinion vraie adviennent l'une comme l'autre selon des lois psychologiques. On peut bien vouloir en tirer une analyse du processus psychique qui accompagne une opinion, cela ne peut pas remplacer une preuve de l'objet de cette opinion.[89]

Pour distinguer la pensée logique de la pensée psychologique, Frege recourt à la notion d'être vrai (et non de tenir pour vrai, qui concerne la psychologie) :

---

[87] [Engel 1996], p. 123.
[88] [Gardenförs 1980], p. 161; nous soulignons. L'ouvrage d'Ellis en question est *Rational Belief Systems*, in American Philosophical Quarterly, Library of Philosophy 3, Basil Blackwell, Oxford (1979).
[89] [Frege 1919a], pp. 170-1.

j'appelle pensée [*Gedanke*] ce dont on peut demander s'il est vrai ou faux.[90]

C'est précisément ce rôle de la notion de vérité que Ellis associe à la logique frégéenne pour mieux s'en démarquer : selon lui, les critères de rationalité d'un ensemble de croyances doivent se trouver *dans* la psychologie. Une telle position va à rebours de la position de Frege sur la logique comme ensemble de lois prescriptives de la pensée pure, par opposition aux pensées réelles que décrit la psychologie :

> Tout comme les lois de la nature, les lois de la vérité sont des lois descriptives, quoique elles puissent être considérées comme une base en vue de prescriptions, et ainsi comme une base en vue de lois normatives : en vue des lois de la pensée *rationnellement correcte*. (...) Toujours en lutte contre le psychologisme, Frege insiste fortement sur le fait que la logique ne concerne pas les lois *psychologiques* de la pensée effective [*actual thinking*] ; la tâche de la logique est de découvrir les lois de la vérité, non les lois descriptives empiriques des croyances ou pensées qui se produisent réellement.[91]

Le problème concerne la relation explicative entre la notion épistémologique de croyance et la notion logique de vérité. Selon Ellis, une description de la vérité en termes d'ensembles de croyances donne une justification épistémique à la vérité d'un énoncé et, ce faisant, renverse le projet de logique épistémique en une *épistémologie de la logique*. Comment se traduit le passage d'une logique des croyances vraies (ou fausses) à une logique de la vérité (ou fausseté) des croyances, et dans quelle mesure ce renversement manifeste-t-il une forme de psychologisme ?

Le psychologisme revendiqué par Ellis se manifeste par son rejet du rôle fondamental de la vérité dans la description des opérations rationnelles, dans lesquelles on assiste à l'enchaînement d'une croyance à une autre. Il explique son psychologisme en ces termes : le remplacement de la sémantique standard vériconditionnelle, basée sur les termes *primitifs* de vérité et de fausseté, par une sémantique des croyances où la cohérence prime sur la vérité :

> Par sémantique standard, j'entends le type de sémantique que l'on peut trouver dans pratiquement tous les manuels de logique, dans lesquels les connecteurs et opérateurs logiques d'une structure de langage formellement (syntaxiquement) définie sont définis en termes de conditions de vérité et de fausseté. J'appellerai *sémantique de vérité* tous les fondements d'un système logique de ce genre. Le besoin d'une sémantique de vérité dérive du concept standard de validité d'un argument, selon lequel un argument est valide si et seulement si il n'y a pas d'*interprétation* de ses termes non logiques (ou autrement dit, pas de monde possible) dans laquelle ses prémisses sont *vraies* et sa conclusion est *fausse*. En effet, il résulte de cette définition que pour comprendre suffisamment une phrase pour toutes les fins de la logique, il nous faut savoir sous quelles interprétations (ou dans quels mondes possibles) elle serait vraie ou fausse. Je procède à partir d'un concept de validité différent, qui conduit à un programme d'analyse différent, et donc à un type de fondement pour les systèmes logiques différent. Je considère qu'un argument d'un langage donné est valide ssi il n'y a pas de *croyance rationnelle* dans cette langue pour laquelle ses prémisses sont *acceptées* et sa conclusion est *rejetée*. Pour

---

[90] [Frege 1919a], p. 173. La formule de Frege pose le problème du porteur de vérité, *ce* qui est vrai ou faux. Par la suite, notre rejet de l'entité objective 'proposition' transfèrera ce rôle de porteur aux énoncés indexicaux.

[91] U. Meixner : "Is logic the science of the laws of truth ?", in *Building on Frege* (*News Essays on Sense, Content, and Concept*, A. Newen & U. Nortmann & R. Stuhlmann-Laeisz (eds.), CSLI Publications, Standord (2001), p. 329.

moi, la *validité* est ainsi une notion *épistémique*, et mon problème consiste à définir un système de croyance rationnelle dans une langue, ainsi qu'à spécifier les conditions d'acceptation et de rejet pour les phrases de cette langue.[92]

Ce renversement conceptuel entre vérité et croyance se traduit par une inversion des termes dans le métalangage et le langage-objet : vérité et fausseté deviennent des notions du langage-objet, la croyance devenant la notion sémantique centrale du métalangage. Ce n'est pas la vérité ou fausseté des énoncés qui définit la signification de leur combinaison, comme ce peut être le cas dans une sémantique vériconditionnelle, mais leur acceptabilité par un individu au sein d'un système de croyances.[93] Un exemple de ce renversement de perspective est donné par les conditions de rationalité (et non plus de vérité) d'une croyance chez Ellis : pour déterminer quelles croyances peuvent entrer dans un ensemble de croyances rationnel, ou admissible, il s'agit de spécifier les énoncés qui sont *considérés comme vrais*, les énoncés qui sont *considérés comme faux*, et les énoncés qui sont *considérés comme pas encore décidés*. Tout en refusant la sémantique relationnelle et son discours des mondes possibles, Ellis décrit ces trois qualités des énoncés non par des opérateurs mais des ensembles d'évaluations T, F et X sur l'ensemble des énoncés du langage donné, où $T\alpha$ signifie que α est accepté comme vrai, $F\alpha$ que α est accepté comme faux, et $X\alpha$ que l'intégration de α dans l'ensemble des croyances est encore indécidé. Pour montrer que les conditions d'acceptabilité d'Ellis reposent tout autant sur les contenus propositionnels et leurs relations logiques que chez Hintikka, une traduction mutuelle de leurs langages-objets est envisageable. Puisque la distinction entre connaissance et croyance n'apparaît pas chez Ellis, dans la mesure où le critère d'acceptabilité exprime une croyance justifiée, on peut concevoir ses expressions $T\alpha$, $F\alpha$ et $X\phi$ comme les contreparties des modalités doxastiques $B\phi$, $B\neg\phi$ et $(\neg B\phi \wedge \neg B\neg\phi)$ de la logique épistémique. En d'autres termes, croire une formule ϕ c'est considérer ϕ comme vrai, croire que $\neg\phi$ c'est considérer ϕ comme fausse, tandis que $X\alpha$ désigne l'absence de jugement au sujet de ϕ. Voici deux exemples de condition d'acceptabilité d'un énoncé, où T et F sont assimilables à des déclarations de croyance justifiée portant sur des énoncés considérés comme connus. Soit un ensemble de croyances *B* au sein d'un langage L, les ensembles de croyances renvoyant à des types d'énoncés distincts ; alors $T(\alpha \wedge \beta)$ figure dans *B* si et seulement si $T\alpha$ et $T\beta$ figurent dans *B*, $F(\alpha \rightarrow \beta)$ figure dans *B* si et seulement si ni $F\alpha$ ni $T\beta$ ne figurent dans *B*. L_0 correspond au langage des énoncés classiques et peut être complété par d'autres langages $L_1, \ldots L_n$, chacun portant sur des classes de formules dont les conditions d'acceptabilité sont spécifiques (énoncés nécessaires, contrefactuels, causaux, etc.). Or il se trouve que les autres conditions d'acceptabilité de L_0 proposées par Ellis confirment notre objection à son soi-disant psychologisme, et cette objection est la suivante.

[92] [Ellis 1976], p. 187.

[93] "Pour Ellis, par opposition, la rationalité apparaît avant la vérité. La compréhension par une personne d'un langage, dit Ellis, 'dépend du vaste schéma d'acceptations et de rejets qu'affiche son système de croyances' (p. 25). Ainsi, il considère l'*acceptabilité* comme une notion plus fondamentale que la vérité. En spécifiant les conditions d'acceptabilité pour un système de croyance rationnel, on peut déterminer à quel moment une déclaration de vérité ou de fausseté concernant un énoncé quelconque d'un langage peut être incorporée dans le système de croyance. Les lois de la pensée sont, pour Ellis, des conditions d'acceptabilité de ce genre" ([Gardenförs 1980], p. 162).

Quelle différence avec la logique vériconditionnelle ? Aucune, selon nous, sinon dans l'utilisation de la notion de 'vérité'.

Premièrement, Ellis semble confondre épistémologie et logique, justification et hypothèse, vérité matérielle et vérité formelle lorsqu'il oppose les notions de vérité et de croyance, et son holisme épistémologique le conduit pour cette raison à préférer parler en termes de conditions d'acceptabilité plutôt que de conditions de vérité. En d'autres termes, Ellis veut remplacer la sémantique de vérité parce qu'il confond *détermination* de la vérité et *assignation* de vérité. En effet, la vérité d'un énoncé est pour lui la conséquence de sa capacité à être intégrée au sein d'un ensemble d'énoncés considérés comme vrais, et l'on aperçoit ici une conception de la vérité-cohérence qu'il maintient jusqu'à l'intérieur de sa sémantique formelle. Or les critères de cohérence qu'il propose sont les mêmes que ceux d'un logicien classique tel que Quine, ce qui n'empêche pas ce dernier d'adopter une sémantique *véri*conditionnelle en termes de conditions de vérité : le logicien postule des assignations de vérité sans avoir besoin de les justifier, dans la mesure où la vérité logique d'une formule complexe se caractérise par sa vérité pour toute assignation de valeurs à ses atomes. La différence entre les deux auteurs se situe dans leur utilisation de la notion de 'vérité' en logique : bien que Quine défende tout comme Ellis un holisme épistémologique, en vertu duquel la vérité d'un énoncé particulier dépend de sa relation aux autres énoncés au sein d'un système de connaissance, cela ne l'empêche pas de concevoir la vérité du point de vue logique comme une simple valeur postulée : en logique, on suppose que tel énoncé est vrai ou faux puis l'on en tire des conséquences hypothétiques. La seule différence avec Ellis est d'ordre terminologique : celui-ci ne parle pas de conditions de 'vérité' mais d''acceptabilité', qui expriment des croyances mais obéissent aux même règles conditionnelles d'inférence que les premières.

Deuxièmement, et c'est une conséquence de la dernière phrase ci-dessus, les notions du métalangage et du langage-objet d'Ellis ne sont certes pas les mêmes que dans une sémantique vériconditionnelle. S'il est vrai que l'auteur a substitué aux valeurs sémantiques de vérité ou de fausseté les valeurs sémantiques d'acceptation et de rejet, les relations d'inférence entre les énoncés conservent toutefois la même *forme* ou *structure* qu'en logique classique. L'acte de croyance est donc premier dans la constitution d'une sémantique formelle, et la vérité telle qu'Ellis en parle devient le résultat d'une application des règles d'acceptabilité d'énoncés au sein d'un ensemble de croyances. Un énoncé est 'vrai' s'il respecte les critères de cohérence. Or il apparaît clairement que sémantique de vérité (vériconditionnelle) et sémantique de croyance, bien qu'elles inversent chacune la relation d'ordre entre vérité et croyance, se conforment toutes les deux aux mêmes critères de cohérence. Pour résumer la position d'Ellis et notre objection à ses prétentions psychologistes, nous dirons que sa sémantique d'acceptabilité n'est rien d'autre qu'un ensemble d'opérations booléennes semblables à celles de la logique classique, la seule différence résidant dans les termes utilisés pour décrire ces opérations. Dans les exemples donnés plus haut, on peut reformuler les deux conditions d'acceptabilité en termes vériconditionnels : $v(\phi \wedge \psi)$ = vrai ssi $v(\phi) = v(\psi)$ = vrai ; $v(\phi \rightarrow \psi)$ = faux ssi $v(\phi)$ = vrai et $v(\psi)$ = faux. Nous voyons ainsi que les lois d'acceptabilité rationnelle d'Ellis ne sont autres que les lois de l'être vrai de Frege, quoique présentées sous une appellation différente (l'être acceptable) et admettant une valeur d'indécision X. Celle-ci permet à l'auteur de décrire les ensembles de croyances comme des

ensembles d'énoncés ouverts ou en attente d'être complétés par d'autres informations (condition de clôture). La conséquence notoire des conditions d'acceptabilité dans un ensemble ouvert est le rejet du tiers exclu : $X(\phi \vee \neg\phi)$ est une formule acceptable lorsque ϕ n'est pas encore justifiable pour le sujet croyant, et il n'est pas exigé que $T(\phi \vee \neg\phi)$ figure dans tout ensemble *B* de L ; en revanche, tout énoncé tend à être considéré comme vrai ou faux, dans l'attente d'une preuve.[94] Ce principe d'extensibilité des ensembles de croyances amène Ellis à compliquer la formulation de certaines règles sémantiques, à l'image de son adaptation des critères de consistance : $T\phi$ apparaît dans L si $F\phi$ n'apparaît dans aucune extension achevée de *B* d'une extension de L quelconque, et $F\phi$ apparaît dans *B* si $T\phi$ n'apparaît dans aucune extension achevée de *B* d'une extension de L quelconque. Lorsqu'un langage L satisfait ces deux formes de consistance, le système de croyances est dit strictement rationnel.

Un intérêt de la sémantique de croyance d'Ellis tient dans son traitement de la clôture déductive des croyances, qui évite soigneusement de tomber dans le piège à venir de l'omniscience logique (voir section 2.2.4.1) :

> les systèmes de croyance strictement rationnels ne sont vraiment conçus que pour des dieux, et ceux-ci n'ont de toute façon pas besoin d'un système de logique. Par conséquent, je préfère le concept de rationalité plus faible, et donc je définis une tautologie de L comme une phrase de L qui ne reçoit pas une évaluation F dans un système de croyance rationnel de L.[95]

En termes de sémantique relationnelle, cela signifie que, dans l'ensemble des croyances de l'agent *a*, $B_a\phi$ et $\phi \Rightarrow \psi$ n'obligent pas toujours à attribuer $B_a\psi$ au sujet croyant et n'interdisent pas le cas d'indécision $\neg B_a\psi$; seule $B_a\neg\psi$ est exclue des conditions d'acceptabilité d'un tel ensemble de croyances. La raison de cette différence est que, chez Ellis, les croyances ne représentent pas des ensembles d'énoncés *clos*, car chaque conséquence logique ψ d'une croyance ϕ n'est pas forcément incluse dans l'ensemble de croyances *B* ; autrement dit, tout monde doxastiquement possible n'est pas 'achevé' et la condition de clôture du système modal **K** n'apparaît pas ici.

En dehors de cet avantage, nous pouvons conclure à partir des remarques précédentes sur ce point : la sémantique de croyance d'Ellis n'a de différence que terminologique avec la sémantique de vérité ; les conditions de vérité sont rebaptisées du nom de conditions d'acceptabilité mais, dans les deux cas, les règles sémantiques suivent une même structure *algébrique* : un même ensemble d'opérations générales les sous-tend, dans lesquelles la notion de vérité est philosophiquement *neutre* et ne renvoie à aucun type de fait en particulier.[96] La neutralité de la notion de vérité confirme notre idée que [Ellis 1976] n'est pas

---

[94] "L'idéal de rationalité est satisfait principalement du fait de l'exigence que les systèmes de croyance puissent être complétés [*completable*] par le biais de toute extension de L_0. Pour les T-systèmes, cela veut dire que, pour tout langage qui inclut syntaxiquement L_0, il y a un certain moyen de remplacer les évaluations en X qui figurent dans le système de croyance par des évaluations en T ou en F, sans violer aucune des autres lois" ([Gardenförs 1980], p. 163). Le cas $X\alpha$ sera revu plus loin dans [Quine 1974], en termes de fonction verdictive d'abstention. Mais nous ne considérerons cependant pas ce comportement comme une raison suffisante pour abandonner la bivalence classique. Voir section 3.2.3.

[95] [Ellis 1976], p. 192.

[96] "Nous pouvons donc accepter la conception sémantique de la vérité sans abandonner la moindre attitude épistémologique que nous ayons pu avoir ; nous pouvons demeurer des réalistes naïfs, des réalistes critiques ou

un véritable psychologisme, au sens où il ne définit pas les propriétés de la vérité en termes de lois psychiques causales.

2.1.3.3 La neutralité du concept sémantique de vérité

Tout au plus le psychologisme revendiqué par Ellis illustre-t-il un psychologisme faible, où les notions de vérité et de fausseté sont simplement reformulées en termes psychologiques d'acceptation et de rejet mais conservent leurs propriétés sémantiques, formelles.

Des psychologistes du 19e siècle ont utilisé les mêmes termes d'acceptation et de rejet pour *expliquer* la loi du tiers exclu, parmi lesquels Alexander Bain :

> La loi du Tiers Exclu est simplement une généralisation de l'expérience universelle que certains états mentaux sont destructifs d'autres états. Elle formule une certaine loi absolument constante, que l'apparition d'un mode positif quelconque de conscience ne peut pas se produire sans exclure un mode négatif corrélatif (...) D'où il s'ensuit que si la conscience n'est pas dans l'un des deux modes elle doit l'être dans l'autre.[97]

Par mode de conscience, Bain désigne une attitude du sujet que l'on pourrait symboliser aujourd'hui par l'opérateur modal B ; si tel est le cas, alors la seconde phrase présente le tiers exclu comme une relation conditionnelle dans laquelle un mode positif de conscience, c'est-à-dire la reconnaissance de vérité symbolisée $B\phi$, entraîne le rejet de son opposé qu'est le mode négatif de conscience, c'est-à-dire la reconnaissance de fausseté $\neg B\neg\phi$. On obtient ainsi la contrepartie modale du tiers exclu psychologiste : $B\phi \rightarrow \neg B\neg\phi$, qui correspond également au schéma caractéristique du système modal **D**. Cela dit, une telle symbolisation suppose que l'acte de 'destruction de la conscience' et le mode négatif soient tous deux exprimables par le connecteur de négation $\neg$, et que celui-ci ne porte pas uniquement sur des formules ou contenus de conscience ($B\neg$), mais également sur la conscience elle-même lorsqu'elle est 'détruite' ($\neg B$). Si l'on interprète le tiers exclu comme l'idée selon laquelle de deux jugements (modes de conscience) contradictoires au moins l'un d'eux est *vrai*, l'interprétation psychologiste équivaut apparemment à l'idée selon laquelle tout jugement ou opinion implique le rejet de son opposé ; or puisqu'une opinion consiste à croire que quelque chose est vrai ($B\phi$), l'opposé de cette opinion consiste à croire que la même chose est fausse ($B\neg\phi$) plutôt qu'à ne pas la croire comme vraie ($\neg B\phi$), étant donné qu'une non-croyance n'exprime pas forcément un mode de conscience et peut signifier au contraire l'absence pure et simple de l'acte de conscience. Le tiers exclu du psychologiste met donc aux prises deux actes de conscience effectifs, dont l'un est un 'mode positif' et l'autre un 'mode négatif' : $B\phi$ et $B\neg\phi$, de sorte que le tiers exclu de Bain n'oppose pas $B\phi$ et $\neg B\phi$. La comparaison entre la version psychologiste de Bain et la version moderne doxastique pose problème, si l'on admet la définition du conditionnel en termes de disjonction ; on obtient ainsi $(B\phi \rightarrow \neg B\neg\phi) \equiv (\neg B\phi \vee$

---

des idéalistes, des empiristes ou des métaphysiciens –quoi que nous ayons été auparavant. La conception sémantique est complètement neutre par rapport à toutes ces questions" ([Tarski 1944], p. 362). La distinction est clairement marquée également dans [Quine 1993] entre logique et épistémologie, ou vérité et croyance justifiée, mais elle ne l'est pas chez [Ellis 1976] lorsque celui-ci qualifie la validité de notion épistémique.

[97] A. Bain : *Logic* (1870), trad. française *Logique déductive et inductive*, Paris, Baillère (1875), vol. 1, p. 323.

$\neg B\neg\phi) \equiv (B\phi \vee B\neg\phi)$, après substitution de $\neg B$ par B.[98] Le problème réside dans le fait que la version doxastique du tiers exclu n'est pas nécessairement $(B\phi \vee B\neg\phi)$ et peut être sinon $B(\phi \vee \neg\phi)$, qui porte sur un seul et même acte de conscience composé de deux modes opposés. Mais malgré la différence logique admise de nos jours entre ces deux formules (elles ne sont pas équivalentes), Bain ne pouvait être sensible à ce problème puisque, dans son approche psychologiste, actes de conscience et contenus de conscience ne faisaient qu'un : $B\phi$ et ϕ, $B\neg\phi$ et $\neg\phi$ ne font qu'un pour lui, de sorte que la version du tiers exclu qu'il propose est sans doute plus proche de la formule admise par les intuitionnistes : $(\phi \rightarrow \neg\neg\phi)$, la double négation signifiant la destruction d'un mode de conscience négatif.[99]

Quoi qu'il en soit, la façon dont Bain et Ellis basent la logique sur des actes psychologiques d'acceptation et de rejet n'est visiblement pas identique : le premier admettait le principe du tiers exclu sous plusieurs formulations comprenant $B(\phi \vee \neg\phi)$, et le second a rejeté ce principe sous sa formulation $T(\phi \vee \neg\phi)$. La citation de Bain semble s'adresser davantage au principe de double négation qu'au tiers exclu, mais l'on retrouvera cela dit une formulation similaire en termes d'assertion et de rejet dans [Russell 1973].[100] De plus, et surtout, les actes dont parle Ellis n'*expliquent* pas les lois logiques mais les accompagnent. Le problème des lois logiques auxquelles des actes psychologiques se superposent est comparable au débat psychophysique concernant la relation d'identité entre états mentaux et états physiques : certains affirment une réduction du psychique au physique, prétendant que le psychique survient sur le physique ; d'autres (parmi lesquels Herbert Feigl) plaidaient en faveur d'un parallélisme psycho-physique neutre et non réductionniste. Bain plaide pour une solution du premier type, concernant la relation entre logique et psychologique : il n'y aurait pas seulement un parallélisme logico-psychologique, mais une sorte de survenance du logique sur le psychologique. Pour quelle position plaide Ellis ? Une explication psychologique des lois logiques ne tient pas à leur simple description en termes d'actes de croyance.

En effet, le type de psychologisme que propose Ellis n'a rien d'un psychologisme fort ou réductionniste, dans lequel les lois logiques de préservation de la vérité seraient causées par des lois psychiques d'acceptation et de rejet. On peut admettre en un sens que la sémantique des croyances est d'ordre psychologique, dans la mesure où elle se base sur ce que l'agent est censé considérer comme vrai ou comme faux. Les lois qu'il suit dans ses raisonnements logiques ne sont peut-être pas des lois issues d'un monde d'entités platoniques, le monde des

---

[98] La substitution est effectuée pour rétablir les actes de conscience que décrit le psychologiste : le tiers exclu oppose deux actes de conscience $B\phi/B\neg\phi$ et non deux absences d'actes $\neg B\phi/\neg B\neg\phi$, ce qui serait une présentation absurde du tiers exclu psychologiste.

[99] Bain considère visiblement le tiers exclu et la double négation comme des principes interchangeables : la double négation entraîne l'affirmation (et vive versa), parce que l'auteur présuppose la bivalence stricte (chaque proposition reçoit une et seule valeur de vérité, soit le vrai soit le faux) et parce que la négation est involutive (elle inverse la valeur de l'argument). Par conséquent, la double négation signifie la fausseté de la fausseté et implique la vérité. Il en irait autrement dans une logique déviante à plus de deux valeurs de vérité, mais Bain n'en aurait pas accepté l'usage.

[100] Russell est anti-psychologiste, toutefois, et sa version psychologique en termes de rejet et d'assertion est clairement distinguée de la loi *logique* du tiers exclu. Voir la section 3.3.2, dans laquelle la dénégation de Russell est examinée sur un mode non plus psychologique mais pragmatique (conséquence de l'analyse performative de la thèse KK). La version de Russell n'est pas celle de Bain, en tout cas, puisqu'elle dit que la dénégation entraîne l'assertion opposée ; si l'on interprète la dénégation comme la 'destruction d'un mode de conscience', on obtient la formule inversée : $(\neg B\phi \rightarrow B\neg\phi)$. Mais nous verrons aussi que dénégation et destruction d'un mode de conscience ne sont pas simplement réductibles à la forme $\neg B$.

lois de l'être vrai de Frege, mais ce détail métaphysique ne fait aucune différence pour la pratique logique actuelle : peu nombreux sont les logiciens admettant encore l'hypothèse d'entités de ce genre afin de justifier les opérations logiques. Or si les règles qui gouvernent l'usage de la négation ou de la conjonction ne sont pas objectives, elles ne sont pas réductibles pour autant à des faits de conscience. Comment justifier l'identité des règles de raisonnement entre les locuteurs, si ce n'est ni tout à fait par des lois de l'être vrai, ni tout à fait par des lois du considérer comme vrai ? L'esprit et le langage, les croyances et les énoncés semblent suivre communément de mêmes structures de transition entre des formes d'énoncés.

L'identité de structure entre les conditions sémantiques de [Hintikka 1962] et les conditions d'acceptabilité des ensembles *B* de [Ellis 1976] est mise en évidence au sein de la logique algébrique : à ce niveau d'abstraction du discours, aucune notion telle que la croyance ou la vérité n'apparaît et seuls des signes vides décrivent des groupes d'opérations, valables aussi bien pour des relations entre des nombres (arithmétique) qu'entre des énoncés des langues naturelles (logique philosophique) ou d'autres objets quelconques. Dans ce registre des structures, la logique classique peut être interprétée comme un ensemble muni d'une loi de composition interne (une application d'un ensemble d'éléments vers des éléments de cet ensemble ; en logique formelle, les éléments peuvent être des valeurs sémantiques) et dont les connecteurs obéissent à des propriétés précises : distributivité, associativité, idempotence, etc. A partir de ce jargon algébriste, la logique classique est assimilée à une structure d'anneau et décrite également comme un treillis booléen. Sans entrer dans les détails, cette mise en abstraction des logiques interprétées a pour but de montrer que les notions philosophiques de proposition, d'énoncé ou de vérité ne sont pas essentielles pour la caractérisation des lois logiques, ce qui amène par exemple à remplacer vérité et fausseté par de simples marques abstraites telles que V et F, 1 et 0, ou + et –. Dans la sémantique de croyance d'Ellis, les exemples de conditions d'acceptabilité vues plus haut n'avaient pas besoin d'être philosophiquement interprétées pour être identifiables à de simples opérations booléennes : être vrai ou être considéré comme vrai peuvent être librement remplacées par 1 ou +, leurs structures n'en sont pas moins identiques et n'obéissent pas moins aux mêmes propriétés algébriques, neutres et sans incidence sur le débat réductionniste entre logique et psychologie. En logique modale, la sémantique relationnelle de Hintikka suggère elle-même la mise entre parenthèses de la notion de vérité lorsqu'il redéfinit le fait d'être vrai comme le simple fait d'appartenir à un monde possible. Une structure peut être définie sans avoir besoin d'interpréter les objets sur lesquels elle porte, et c'est pour cette raison que la réinterprétation des valeurs de vérité en d'autres valeurs n'a aucun effet sur les fondements de la logique ; le psychologisme impliquerait plutôt la remise en cause de ces lois structurelles, ce qui n'est le cas ni en logique modale épistémique ni dans la théorie des ensembles de croyance de Ellis.

Quelle que soit donc la nature de la notion de vérité, les états psychologiques peuvent être introduits en logique comme tout argument au sein d'une chaîne d'inférence formelle ; cela n'implique en rien une réduction psychologiste de la vérité à l'opinion. Logique classique et logique épistémique manifestent tout autant des structures algébriques qui organisent leurs règles sémantiques et définissent leurs conditions de consistance. Mais l'affirmation d'Ellis selon laquelle ses règles de cohérence dériveraient de la psychologie sont clairement réfutées par [Gardenförs 1980], car rien n'est en mesure de montrer que ce sont des lois

psychologiques qui déterminent les règles structurelles au sein d'une inférence. Le critère métalogique de consistance est par exemple impossible à *justifier* en termes psychologiques :

> Ellis affirme, sans preuve, que les critères d'acceptabilité pour le langage L_0 ont pour conséquence qu'un énoncé α est un théorème du calcul propositionnel si et seulement si $F\alpha$ ne figure pas dans un quelconque T-système rationnel pour L_0.[101]

Pour le dire autrement, la phrase ci-dessus signifie qu'un énoncé est valide ou toujours vrai si et seulement s'il n'a pas de contre-modèle. Or la remarque de Gardenförs préfigure le cas de la logique paraconsistante et l'utilisation de logiques déviantes pour déterminer d'autres conditions d'acceptabilité, d'autres normes de rationalité des croyances. Pourquoi les ensembles de mondes issus de [Hintikka 1962] ou les ensembles de croyances *B* de [Ellis 1976] correspondent-ils à une seule et même structure algébrique de type booléen, du moins en ce qui concerne les cas de croyances T et F, et pourquoi ne pas choisir d'autres sortes de structures pour décrire les modèles d'acceptabilité des ensembles de croyances ? L'attitude d'Ellis est comparable à cet égard à celle de Hintikka, puisque son traitement des modalités (aléthiques) consiste à enrichir ou inclure dans L d'autres ensembles de croyances en plus des croyances déclaratives (factuelles) de *B*. C'est d'ailleurs dans le but de proposer des modèles pour ces types d'énoncés de croyance plus complexes (contrefactuels, aléthiques) que Ellis avait justifié son rejet de la sémantique de vérité.[102]

Au total, la véritable différence de la logique modale épistémique vis-à-vis psychologisme réside dans la *neutralité* des concepts sémantiques, que ce soit la vérité ou un autre nom de valeur ; mais la logique modale a nécessité une modification du cadre sémantique classique, passant d'un modèle unique qu'est l'univers de discours à un ensemble de modèles relationnels. C'est cette modification qui, selon nous, a constitué le véritable moyen d'introduire le concept de croyance dans l'analyse logique et de le dissocier de la psychologie. Nous allons développer dans ce qui suit les deux présupposés implicites qui empêchaient la constitution d'une logique des croyances, manifestant deux sortes de dogmes logiques.

2.1.4 Deux dogmes de la logique

Dans sa littérature consacrée à l'histoire de la logique, Hintikka a tenté de montrer à plusieurs reprises que les obstacles rencontrés par la logique modale s'expriment principalement à travers deux présupposés majeurs, qu'il considère comme deux dogmes

---

[101] [Gardenförs 1980], p. 163.

[102] "Mon rejet de la sémantique de vérité et du concept traditionnel de vérité résulte de ma conviction qu'il n'y a pas de sémantique de vérité adéquate pour les déclarations de nécessité et de possibilité, les conditionnels contrefactuels, les jugements moraux, les déclarations causales, les affirmations de forme 'A explique B' et les déclarations de probabilité inductive. La recherche des conditions de vérité pour ce genre d'affirmations a été une préoccupation des philosophes depuis plusieurs générations, et je suis convaincu qu'aucune solution adéquate à tous ces problèmes n'est en vue. Il se peut qu'il existe des solutions à ces problèmes que nous n'ayons pas encore découvertes. Mais nul doute que nos fondements de la logique ne devraient pas être fournis sur la base du fait que de telles solutions existent. La question de savoir si l'on peut dire d'une phrase d'une langue qu'elle est vraie ou fausse en un sens absolu n'est pas une question dont les fondements de la logique devraient dépendre". ([Ellis 1976], p. 191). Ellis entend par modalité les énoncés de nécessité ('il est nécessaire que –') exprimant des lois physiques de la nature, considérant que nos opérateurs extrinsèques 'B' ou 'K' ne figurent pas dans son langage objet.

prégnants de la logique du début du 20e siècle : le premier est le *réalisme* ou *rationalisme logique*, le second est la thèse de la logique comme *langage universel* et résulte de la première.[103] Bien que ces deux positions soient plus ou moins délaissées dans la pratique logique actuelle, la première nous intéressera dans le cadre de la logique philosophique puisqu'elle pose le problème de la légitimité des constantes logiques.

2.1.4.1 Contre le dogme du rationalisme logique : Tolérance et relativisme logique

On oppose plus ou moins le Principe de Tolérance de Carnap à la position des rationalistes logiques, en vertu de laquelle les règles d'inférence logiques sont des lois universelles et valables pour toute formule ϕ quelle qu'elle soit. Mais on peut interpréter l'universalisme logique de deux façons distinctes : en termes d'*invariance* des connecteurs, ou en termes de *généralité* du domaine de discours.

($\forall$I) Universalisme comme invariance

Le critère d'*invariance* signifie que les règles d'inférence logiques sont valables pour tout langage et dans tout type de raisonnement ; la thèse de l'universalisme comme invariance s'applique aux constantes logiques (les opérateurs) et implique que leur signification (les opérations qu'elles effectuent) ne change pas d'un langage formel à un autre.

Un partisan de l'universalisme comme invariance des connecteurs est [Quine 1975], lorsqu'il prétend notamment que 'changer la signification des connecteurs, c'est changer le sujet'. Hintikka est un partisan de la logique modale mais ne remet pas en cause le sens classique des opérateurs ; en cela, il ne se distingue pas de l'adversaires des logiques modales qu'est Quine. Leur opposition porte donc sur un autre critère que ($\forall$I).

($\forall$G) Universalisme comme généralité

Le critère de *généralité* signifie que les constantes logiques s'appliquent dans n'importe quel modèle, quel que soit le contenu matériel de l'énoncé.

Le critère de l'universalisme comme généralité d'application dans un modèle permet notamment de distinguer la logique des autres disciplines, le langage pur des théories. Contrairement aux énoncés de la physique, de la biologie ou de la géographie, qui appartiennent à des théories parce qu'elles portent sur une classe d'objets restreinte, la logique n'est censée porter sur aucune théorie en particulier : la double négation, le tiers exclu ou la commutativité de la conjonction s'appliquent à n'importe type d'énoncé et garantissent ainsi la portée universelle de la logique.

---

[103] [Largeault 1993] décrit ainsi (p. 8) le rationalisme logique : "Les thèses rationalistes s'appuient sur la conviction que les vérités logiques expriment une nécessité *a priori*, indépendante de l'expérience, que l'expérience n'explique point. Puisque la logique possède un caractère de certitude absolue, ses lois ne peuvent pas, estime-t-on, résulter seulement des propriétés contingentes de la langue et de la grammaire. Il faut qu'elles représentent des exigences immuables de la pensée et de la *raison*. En réciproque, si le formel est apriorique, il doit avoir une valeur absolue et conférer mécaniquement le caractère du vrai à tout contenu de connaissance qui peut y rentrer. Ainsi s'explique, comme indiqué par Brouwer, la croyance à l'infaillibilité et à l'unicité de la logique classique". Le caractère obsolète de la notion apriorique de 'raison' dans l'épistémologie actuelle, remplacée par la notion empiriste de rationalité, explique en large partie l'abandon progressif de cette forme de monisme logique : "La logique de Frege-Russell fut dans le temps la dernière tentative d'asseoir la discipline sur une base ontologique. Rétrospectivement une construction aussi ambitieuse que celle d'une logique sous-jacente à toute la réalité, ou du moins d'une théorie des traits généraux d'une description linguistique totale de la réalité, paraît implausible. Elle s'est aussitôt achoppée à des contradictions (paradoxes de l'autoréférence)" ([Largeault 1993], p. 12).

Nous voyons donc que l'universalisme logique contient deux idées distinctes : l'invariance des constantes logiques est un premier critère d'universalisme, et l'application à tous les domaines de discours (tous les modèles) en est un second. Le Principe de Tolérance exprime un rejet du premier critère mais ne remet pas en cause le second. Quant à la position de Hintikka, nous allons voir que ce dernier admet à la fois (∀I) et (∀G) ; celui-ci maintient en effet l'universalité d'application des principes logiques quel que soit le contenu des énoncés. Cela était déjà vrai avec la théorie des types ramifiés de Russell, dans laquelle plusieurs langages étaient distingués sans aboutir pour autant à une partition des domaines de discours. Logique modale et théorie des types ne s'opposent donc pas au critère de généralité de l'universalisme logique, dans la mesure où

> au sein d'un type il n'y a pas de restriction à un domaine spécifique, et en ce sens l'universalité est préservée. Nous avons un univers stratifié, mais ici encore c'est *l'*univers, non un univers de discours changeable à volonté. Les questions portant sur le système sont aussi absentes des *Principia Mathematica* que dans l'œuvre de Frege.[104]

Le pluralisme de la logique modale correspond à la réinterprétation possible des expressions dans un modèle $\mathcal{M}$: les opérations logiques demeurent constantes et s'appliquent uniformément à toutes les expressions d'un modèle, mais ce modèle n'est pas unique et se compose de plusieurs mondes. Dans la sémantique relationnelle de la logique modale, la relation R entre des mondes possibles exprime cette idée de duplication de plusieurs interprétations au sein d'un 'ensemble modèle' W. Si l'on veut rendre justice à la distinction entre universalistes et pluralistes, un troisième critère devrait donc être ajouté à (∀G).

(!∀G) Universalisme comme *unicité*

Le critère d'unicité signifie que les règles de logique s'appliquent à un domaine d'interprétation unique, c'est-à-dire une classe d'unité {W} sans sous-ensembles w pour réinterpréter les objets de discours. Autrement dit, le domaine de discours ne varie jamais et contient la totalité des objets dans un langage unique. Par contraste, l'univers de discours n'est pas unique en logique modale : celle-ci introduit la totalité des objets dans son domaine de discours, mais l'interprétation de ces objets peut être modifiée. Cette modification entraîne la distinction entre plusieurs domaines de discours : soit une valeur individuelle appartient à un domaine particulier mais n'appartient pas à un autre ; soit cette valeur appartient à deux domaines distincts mais ne satisfait pas les mêmes propriétés dans ces domaines. C'est cette opposition à l'idée d'un univers de discours constant, exprimée par (!∀G), qui fait de Hintikka un opposant à l'universalisme logique (considéré comme la conjonction des critères (∀I), (∀G) et (!∀G).

D'après nos définitions ci-dessus, le rationaliste logique est celui qui admet à la fois (∀I), (∀G) et (!∀G). La section présente s'intéressera d'abord au critère d'invariance des constantes logiques ; la prochaine section portera sur le critère et la distinction majeure entre logiciens classiques et logiciens modaux, qui concerne non pas (∀G) mais (!∀G).

Le premier présupposé des universalistes, le critère (∀I), est lié à leur conception de la syntaxe logique : quelles règles de transformation sont admissibles pour des constantes

[104] [van Heijenoort 1967], p. 326.

logiques dans un système formel, et lesquelles ne le sont pas ? Par opposition aux partisans universalistes de la logique classique, selon lesquels les significations des constantes logiques sont données une fois pour toutes et s'appliquent à n'importe quel énoncé, l'idée de tolérance suppose au contraire que rien n'oblige le logicien à maintenir les propriétés classiques de la conjonction, de la disjonction, etc. :

> En logique, il n'y a pas de morale. Chacun est libre de concevoir sa propre logique, c'est-à-dire sa propre forme de langage, comme il le souhaite. Tout ce qui est exigé de lui, c'est que s'il souhaite l'examiner, il doit exposer ses méthodes avec clarté, et donner des règles syntaxiques à la place d'arguments philosophiques.[105]

Selon l'Autrichien, le problème posé par le critère (∀I) de l'universalisme logique est une question purement *interne* aux langages formels et ne trouvera donc aucune réponse empirique en dehors des langages formels eux-mêmes, c'est-à-dire au sein des langues naturelles. Cela revient à dire que ces dernières, les langues historiquement constituées, sont incapable d'apporter une preuve décisive en faveur d'une traduction 'adéquate' des constantes logiques dans les langues naturelles : en l'absence de critères transcendants pour justifier le sens des constantes logiques, Carnap tolère donc une révision de principe des règles d'inférence logique, et la signification de $\wedge$, $\vee$ ou $\neg$ peut varier d'un langage formel $\mathcal{L}$ à un autre. Le critère (∀I) est donc rejeté.

La signification *intuitive* des constantes logiques n'est pas considérée ainsi comme un critère décisif en vue de la construction d'un système formel, et une révision des connecteurs classique est autorisée en logique. Le rejet de l'intuition n'est toutefois pas une attitude propre aux pluralistes logiques : bien avant la tolérance de Carnap, l'idéographie de Frege avait déjà eu pour but de rejeter le rôle de l'intuition dans la justification des opérations logiques et d'insister sur la procédure de démonstration. Ce qui distingue toutefois l'idée de système formel chez Frege et chez Carnap (ainsi que chez Hilbert), c'est que le premier considérait comme inconcevable toute révision des procédures de calcul au sein du système formel qui sous-tendait son idéographie : chacune des opérations logiques figurant dans une formule reçoit une et seule définition, sans possibilité de les modifier au sein de notre pratique du langage. Pour résumer grossièrement, le *logicisme* se distinguait ainsi du formalisme par le fait que les opérations effectuées portaient sur des objets purs, *a priori* mais ne réduisent pas à de simples symboles sur du papier. Les opérations de la logique conservent un contenu et symbolisent quelque chose, en un sens du terme 'symbole' qui n'est pas celui des formalistes.

La conséquence du pluralisme de Carnap, dans la lignée du formalisme de Hilbert, est qu'il n'y a pas un mais plusieurs langages logiques concevables : chacun de ces langages se distingue en fonction des définitions de ses constantes et des opérations qui les définissent. On sait que, en logique moderne, un système formel peut se composer à la fois d'une syntaxe et d'une sémantique, et les règles de déduction peuvent être exprimées en termes de conditions de vérité des formules lorsque le système est complet : pour toute formule α et β dans un système $\mathcal{L}$, β est déductible de α si et seulement si β est une conséquence logique de α, c'est-

[105] [Carnap 1934], p. 52.

à-dire si la vérité de β s'ensuit toujours de la vérité de α. Commençons par examiner les différences sémantiques possibles entre deux systèmes formels, afin de comprendre l'effet du relativisme logique.

On peut distinguer des langages formels de deux manières : par les opérations effectuées sur leur domaine, en vertu du critère (∀I) de l'universalisme logique ; par le contenu de leur domaine de discours, en vertu du critère (∀G). Cette distinction porte sur les éléments qui constituent un modèle $\mathcal{M}$: le domaine $\mathcal{D}$ dans lequel s'effectue l'interprétation des expressions; la fonction de valuation *v*, qui correspond en logique propositionnelle à la fonction d'interprétation (en termes de valeurs de vérité) des formules non analysées ; la fonction d'interprétation *f*, enfin, qui correspond en logique du premier ordre à la fonction d'interprétation des formules analysés (individus, prédicats).

Le domaine de discours $\mathcal{D}$ comprend un ensemble d'individus satisfaisant des prédications dans un langage. Ce domaine peut être organisé en catégories d'objets : le domaine de l'arithmétique contient un domaine de référence plus restreint que celui du langage ordinaire, puisqu'il contient uniquement des nombres comme valeurs d'individus et quelques prédicats en guise d'opérations sur ces nombres. Quant à la logique, le critère (∀G) de l'universalisme suppose que ce domaine n'a pas de limites et doit comprendre toutes les valeurs de toutes les théories possibles, dont les domaines constituent des sous-ensembles (des sous-modèles) de l'univers de discours. L'universalisme logique semble confirmer l'idée de vérité logique comme vérité dans *tout* modèle, puisque ce qui est logiquement vrai est satisfait par n'importe quelle valeur de l'univers de discours $\mathcal{D}$. Or si l'on prétend qu'une vérité logique est indépendante du modèle au sein duquel elle s'applique, en quoi un changement dans $\mathcal{D}$ peut-il impliquer un changement dans la logique ? La réponse est que les opposants au critère (∀G) de l'universalisme logique, en particulier, n'exigent pas un changement du domaine $\mathcal{D}$ mais un changement dans la signification des constantes logiques, c'est-à-dire dans les propriétés de *v* : une opposition entre plusieurs philosophies de la logique peut être illustrée par une différence dans la caractérisation des constantes logiques.

La fonction *f* correspond à l'interprétation des expressions de $\mathcal{L}$ et assigne des valeurs de $\mathcal{D}$ à chacune des expressions de $\mathcal{L}$: des objets pour les termes d'individus (a, b, c, ...), des propriétés pour les termes de prédicats (P, Q, R, ...). Lorsque des valeurs ont été assignés aux termes de prédicats et d'individus de l'énoncé ϕ, une fonction d'interprétation particulière, la fonction de valuation *v*, assigne une valeur sémantique à l'énoncé ϕ qui en résulte : le vrai ou le faux. Pour les formules complexes de type ⊕ϕ et ϕ⊕ψ –où ⊕ désigne un opérateur unaire ou binaire quelconque– *v* décrit la signification des constantes ⊕ en termes de conditions de vérité pour les formules ϕ et ψ que ces constantes gouvernent ou connectent.

Les conditions de vérité des énoncés complexes peuvent s'exprimer également en termes syntaxiques par différentes méthodes de preuve : une axiomatique ou un calcul des séquents de type Gentzen, par exemple, permet de définir ces opérateurs par des critères de déduction logique : ψ est déductible de ϕ dans $\mathcal{L}$ si et seulement si ψ peut être obtenue à partir de ϕ suite à une série finie d'étapes de démonstration transformant ϕ en ψ dans $\mathcal{L}$. Dans la partie sémantique, ces relations déductives s'expriment en termes de condition de préservation de la vérité des énoncés : ψ est une conséquence logique de ϕ dans $\mathcal{L}$ si et seulement si ψ est vrai dans chaque modèle $\mathcal{M}$ de $\mathcal{L}$ dans lequel ϕ est vrai. Reprenons l'exemple de l'opération de

conjonction, symbolisée par l'opérateur $\wedge$; dans ce cas, le calcul des séquents donne une définition déductive ou syntaxique de cet opérateur à partir de deux modes de transformation de la prémisse (par introduction ou élimination de l'opérateur) dans la conclusion. Quant à la sémantique de $\mathcal{L}$, elle explique que la valeur de vérité (la référence de ϕ dans $\mathcal{D}$) de la conclusion ψ demeure inchangée après que ces deux règles de transformation aient été appliquées sur ϕ.[106] Quel est l'impact de ces procédures sur l'idée d'universalisme logique ?

Si l'on considère l'universalisme logique comme la conjonction des critères ($\forall$I), ($\forall$G), alors les logiciens modaux tels que Hintikka sont des universalistes ; l'universalisme logique que ce dernier rejette ne provient donc pas des critères ($\forall$I)-($\forall$G), mais du critère (!$\forall$G) d'unicité du modèle $\mathcal{M}$. Or le rejet de (!$\forall$G) ne fait pas de Hintikka un pluraliste logique tel que nous l'entendons ici, à travers le critère ($\forall$G). Par opposition à la logique modale épistémique, que Hintikka conçoit comme une extension de la logique classique, le *pluralisme logique* se définit comme la possibilité de réviser les propriétés des constantes logiques et équivaut au rejet de ($\forall$G) : ce n'est pas l'extension du domaine de $\mathcal{D}$ (universelle aussi bien pour les logiciens classiques que modaux) qui change mais le sens des connecteurs définis par v. Autrement dit, pour le pluraliste la signification des constantes logiques peut varier d'un système à un autre, ce qui signifie que la conjonction, la disjonction, le conditionnel etc. ne s'interprètent pas de la même manière dans tous les systèmes ; pour un logicien modal, la signification de ces constantes peut rester invariable mais v s'applique dans des domaines (mondes possibles) variables.

En résumé : si l'on prend pour référence la fonction de valuation v des formules de la logique classique, à partir de laquelle chaque constante logique reçoit une caractérisation précise, on peut distinguer plusieurs types de logiques selon les règles associées à v. L'application invariable des constantes logiques dans un domaine unique, leur application variable dans un domaine unique, leur application invariable dans plusieurs domaines et leur application variable dans plusieurs domaines constituent quatre éventualités distinctes : le réalisme logique ne conçoit qu'un seul type d'interprétation des constante logique et un seul domaine de référence, tandis que le pluralisme admet en plusieurs. En ce qui concerne la logique modale épistémique héritée de [Hintikka 1962], le fait qu'elle soit une extension 'classique' signifie que les connecteurs logiques s'appliquent de manière invariable dans des domaines variables : la signification des connecteurs classiques est inchangée, mais la valuation des formules modales de type $\Box\phi$ et $\Diamond\phi$ s'effectue dans plusieurs mondes possibles. Le maintien des constantes classiques et du critère d'invariance ($\forall$I) peut se justifier pour différentes raisons : Quine et Hintikka sont des partisans de la logique classique au sens où ils maintiennent l'interprétation booléenne des opérateurs, leur différence résidant sur la structure du modèle $\mathcal{M}$; mais leur conservatisme n'implique pas un réalisme logique

---

[106] La définition de la conjonction en termes de séquents donne ceci :

Introduction de la conjonction (I-$\wedge$) Elimination de la conjonction (E-$\wedge$)

$$\frac{\phi \quad \psi}{\phi \wedge \psi} \qquad\qquad \frac{\phi \wedge \psi}{\phi \quad \psi}$$

En théorie des modèles (partie sémantique), ces deux règles sont validées par la définition de l'opérateur de conjonction $\wedge$: $v(\phi \wedge \psi) = 1$ si et seulement si $v(\phi) = 1$ et $v(\psi) = 1$. Concernant la Tolérance logique, celle-ci suppose que règles d'inférence et conditions de vérité des connecteurs peuvent être modifiées à l'intérieur d'un langage formel clos.

selon lequel il serait *inconcevable* de changer l'interprétation des constantes logiques. Le maintien de (∀I) peut se justifier par des arguments divers qui vont du rationalisme logique à la maxime de mutilation minimum de Quine, ou au caractère intuitif des constantes logiques classiques. La section sur le principe de charité présentera un motif de ce genre, dans lequel la défense des connecteurs classiques ne sera pas établie sur des arguments aprioristes mais sur une reconstitution *a posteriori* du discours (section 3.2).

Le principe de tolérance de Carnap présente la logique comme un langage pur, dont les règles de formation et de transformation des énoncés sont établies indépendamment de toute théorie d'arrière-plan et de tout domaine de discours. L'indépendance des opérations logiques vis-à-vis du domaine de référence satisfait la neutralité [*topic-neutral*] de la vérité logique, mais le droit à la *révision* des règles de transformation fait de la vérité logique une propriété interne à un langage ; il en résulte un rejet du réalisme logique et un renforcement du pluralisme en logique : il n'y a pas un mais plusieurs systèmes de logiques concevables. Visiblement, le pluralisme signifie la possibilité *en droit* de changer les caractéristiques de la conséquence logique dans un langage artificiel ; mais il ne dit rien au sujet de la *correction* ou *adéquation informelle* des connecteurs logiques tels que nous les utilisons : faut-il corriger *de fait* ces opérations pour avoir une meilleure compréhension des raisonnements dans les langues naturelles ? Un moyen de couper court à cette question de l'adéquation serait d'adopter une position plus radicale encore que le pluralisme : l'*instrumentalisme* logique, selon lequel l'idée même de correspondance entre les deux catégories des langages formels et naturels n'a pas de sens. Pour quelle raison devrait-on dire de la définition du connecteur ∧ qu'elle saisit 'adéquatement' le sens de la conjonction de coordination 'et' du langage ordinaire ? La relation d'adéquation entre ces deux langages distincts, le langage formel (artificiel) et le langage ordinaire (appliqué), est pourtant au cœur de notre essai parce que de lui dépend la question de la légitimité d'une logique modale épistémique.

L'intérêt principal de la Tolérance de Carnap, c'est qu'elle émancipe le langage logique des idées reçues du langage ordinaire et remplace le critère flou de l'intuition par celui, clair et précis, de l'explication des concepts par définition. Rien n'échappe au contrôle de cette définition, toute procédure de calcul est expliquée dans un langage formel et, en cela, la pureté du langage formalisé satisfait ce que Leibniz attendait d'une procédure de calcul : selon sa métaphore du fil d'Ariane, le calcul nous permet de nous diriger avec sécurité dans le 'labyrinthe' de nos raisonnements et sans créer de confusion, puisque chaque étape du raisonnement est soigneusement définie dans ses relations avec l'ensemble des termes du langage formel. Cela étant, le souci de *clarté* peut paraître insuffisant si la logique philosophique a pour but déclaré de résoudre des problèmes philosophiques : la Tolérance de Carnap ne conduit-elle pas inévitablement à un relativisme sceptique, c'est-à-dire à une séparation définitive entre les règles d'usage du raisonnement ordinaire et les règles d'inférence du langage formel ? S'il n'y a pas de moyen de traduction fiable pour relier les deux types de langage, alors c'est la notion même de 'correction' ou 'adéquation informelle' qui risque de devenir un non-sens et, en conséquence, de condamner la logique philosophique au même titre. Dans le cadre de notre problématique, centrée sur la question de l'adéquation entre les langages artificiels et naturels, un problème fondamental posé par la Tolérance carnapienne concerne les conditions en vertu desquelles des parties du langage ordinaire

pourraient être légitimement reconstruites dans les termes d'un langage formel : comment justifier une telle reconstruction (interprétation dans un langage formel), et celle-ci ne nécessite-t-elle pas quelques règles de traduction minimales et invariables entre les termes du langage artificiel et du langage ordinaire ? Dans le cas précis de la logique modale épistémique, le problème consistera à se demander si les attitudes épistémiques exigent une révision des opérations logiques et un recours à des logiques déviantes (non-classiques).

Nous verrons que la question des logiques déviantes ne concerne pas tant que cela la logique modale épistémique, tout au moins en ce qui concerne celle produites dans la lignée de [Hintikka 1962] ; mais nous constaterons toutefois que des analyses révisionnistes ont été proposées pour modifier certaines règles d'inférence indésirables des opérateurs épistémiques. A l'origine, la logique épistémique se présente au contraire comme une *extension* de la logique classique : les connecteurs vérifonctionnels sont maintenus, comme nous avons pu le constater dans la sémantique relationnelle des espaces épistémiques. Pour cette raison, la véritable différence entre la logique modale épistémique et la logique classique ne concerne pas l'interprétation v mais le domaine d'interprétation $\mathcal{D}$: un changement s'effectue entre la conception classique du domaine de discours (domaine unique) et la théorie des modèles (domaine relatif). C'est dans cette distinction mineure que Hintikka aperçoit une amélioration majeure pour l'histoire de la logique du 20e siècle : le passage de l'idée de logique comme langage à l'idée de logique comme calcul, dont le rejet de l'universalisme logique porte sur le critère (!∀G).

2.1.4.2 Logique comme langage universel, logique comme calcul

L'aversion des universalistes pour la logique modale serait liée à leur conception du modèle, c'est-à-dire du domaine de discours sur lequel porte un système formel. La critique de Hintikka contre l'universalisme logique ne porte pas sur leur refus de faire varier l'interprétation des constantes logiques, puisque lui-même maintient les fonctions de valuation classiques en l'état dans sa logique modale épistémique. Sa critique porte uniquement sur l'argument de l'*unicité* du domaine de discours en logique classique.

Dans le but d'affaiblir l'argument majeur des opposants à la logique modale –la non-vérifonctionnalité des modalités– van Heijenoort puis Hintikka proposèrent d'élargir le cadre de la réflexion en insistant sur deux conceptions opposées de la logique : la logique comme *langage universel*, la logique comme *calcul*.[107] Les opposants à la logique modale étudiés plus haut appartenaient à la première tradition, et leur rejet des modalités tient en partie à l'idée de

[107] Pour une critique de cette distinction systématique entre deux approches de la logique, voir Volker Peckaus : "Calculus ratiocinator versus characteristica universalis ? The two traditions in logic, revisited", *History and Philosophy of Logic* **25**(2004), pp. 3-14. Cela dit, la distinction entre langage et calcul permet à [Hintikka 1997a, 1997b, 1999, 2000] de réexaminer plusieurs arguments célèbres en philosophie de la logique et de les considérer comme des symptômes de l'universalisme en logique. Ce dernier aurait trois conséquences majeures : l'hypothèse d'un modèle unique, qui représente la totalité du monde ; l'ineffabilité de la sémantique ; l'atomisme logique. L'hypothèse du monde unique aurait entraîné ainsi la réification par Frege des *Sinne* comme entités platoniques ; l'ineffabilité de la sémantique aurait entraîné la distinction de Wittgenstein entre dire et montrer ; la thèse de l'atomisme logique, enfin, aurait encouragé l'approche comportementale puis stimulatoire de la signification chez Quine (dans un seul monde). Selon Hintikka, un abandon de ces présupposés permettrait de surmonter notamment la thèse quinienne de l'indétermination de la traduction. La solution se trouve-t-elle dans une approche non-atomiste de la signification des énoncés ? Nous reviendrons sur la thèse de l'indétermination dans la suite (section 3.2.).

langage comme médium universel entre les sujets pensants et ce qu'ils pensent. Les sémanticiens des mondes possibles appartiennent à la seconde tradition et s'estiment en mesure de surmonter les obstacles techniques de la non-vérifonctionnalité. Revenons tout d'abord sur les deux notions citées de langage universel et de calcul, avant d'examiner leur impact sur la logique modale.

La distinction entre langage universel et calcul remonte au projet élaboré par Leibniz d'une *lingua characterica*, c'est-à-dire une langue caractéristique capable d'exprimer toutes les pensées communes à l'espèce humaine. Un tel langage devait dépasser les particularités idiomatiques de chaque langue naturelle et n'en retenir que le fond essentiel, d'où la nécessité d'une réécriture de ces langues au sein d'un unique langage artificiel qui ne retiendrait des phrases que leurs *formes* essentielles. La logique traditionnelle d'Aristote constituait déjà un langage formel, dans la mesure où les termes syllogistiques étaient remplacés par des lettres schématiques abstraites. Mais il manquait à cette syllogistique une procédure rationnelle capable à la fois d'invalider les modes non concluants et de déduire au sein des modes concluants les modes imparfaits à partir des modes parfaits. Alors qu'Aristote établissait ces distinctions sur la base d'évidences intuitives ou d'observations empiriques de l'usage linguistique, Leibniz attendait de sa langue caractéristique un critère de certitude pour conduire aussi bien la pensée en logique qu'en mathématiques.[108] Pour ce faire, les relations entre ces formules sont définies au sein d'un *calculus ratiocinator*, c'est-à-dire une combinatoire de symboles telle que [Blanché 1970] en retient la définition :

> Un calcul est une suite d'expressions formées de symboles, construites et enchaînées conformément à des règles explicitement énoncées et choisies de façon que leur application conduise toujours, au terme d'un nombre fini d'étapes, à une expression considérée comme le 'résultat' du calcul.[109]

Etendu à n'importe quel type de symbole au-delà des nombres, le calcul logique s'applique ainsi à un symbolisme pour constituer un système formel effectif. En ce sens, la formalisation permet de passer du langage symbolique au calcul en distinguant parmi les termes du langage ceux pour lesquels un calcul est permis ou non : les règles de formation et de transformation modernes obéissent à ce schéma général.

Si van Heijenoort et Hintikka opposent langage et calcul, c'est parce que les deux auteurs conçoivent le langage comme un ensemble de formules portant sur un modèle universel et unique, conformément aux critères ($\forall$G) et (!$\forall$G) ci-dessus de l'universalisme logique ; par calcul, ils entendent l'idée selon laquelle une formule de $\mathcal{L}$ peut être réinterprétée selon différentes fonctions d'interprétations (rejet de ($\forall$G)) ou dans plusieurs domaines distincts (rejet de (!$\forall$G)). Or la possibilité de réinterpréter dans plusieurs domaines distincts du modèle $\mathcal{M}$, si elle implique la négation de (!$\forall$G), n'implique pas en revanche la négation de ($\forall$G) : une vérité logique peut être illustrée par n'importe quelle formule dans $\mathcal{M}$ mais peut porter sur plusieurs modèles distincts de $\mathcal{M}$, comme nous l'avons constaté précédemment avec la

[108] Pour consolider les règles syllogistiques d'Aristote, Łukasiewicz en a proposé une formalisation et axiomatisation complètes. Voir son *Aristotle's syllogistic from the standpoint of modern formal logic*, Clarendon Press, Oxford (1951).
[109] [Blanché 1970], p. 209.

sémantique relationnelle de la logique modale épistémique. La distinction de van Heijenoort et Hintikka consiste à insister sur l'aspect artificiel des langages formels car, depuis l'époque de Leibniz, l'idée de 'caractéristique réelle' a été abandonnée entre-temps et délaissée au seul profit de l'idée de 'caractéristique logique'. On ne prétend plus de nos jours à un langage *symbolique* au sens propre, capable d'être en rapport direct de représentation avec les choses, bien que cet idéal se retrouve encore dans l'idéographie de Frege (dépouiller la pensée pure de ses habits vernaculaires) ; il ne reste que l'idée d'une syntaxe pure, indépendante des contingences des grammaires empiriques. Hintikka exprime ainsi l'effet central de la notion de calcul :

> Ce sur quoi je mets l'accent dans mon utilisation du *terminus technicus* 'langage comme calcul', c'est uniquement et simplement le sens d'une *réinterprétabilité*.[110]

Mais une fois que les idéaux de langue philosophique ou de caractéristique réelle ont été abandonnés, comment justifier la configuration choisie pour un langage symbolique et une algèbre universelle, dans la mesure où ces règles de formation et de transformation des expressions d'un langage artificiel ne renvoient plus aux choses elles-mêmes et risquent d'être réinterprétables sans limite ?

Quel que soit l'argument retenu pour justifier la configuration particulière d'un langage symbolique, le Principe de Tolérance de Carnap réduit le symbolique au formel et prétend que rien d'autre que la rigueur démonstrative ne justifie le choix d'un langage par rapport à un autre : il n'y a pas de morale en logique, tout est permis sous condition de clarté et de rigueur. Le langage formel n'a donc rien d'universel, et la relation symbolique entre le langage et les choses est un leurre ; mais le calcul conserve toutes ses vertus méthodologiques à l'intérieur d'un langage artificiel délimité. D'où la priorité de la structure sur le langage et des opérations sur les termes, chez les relativistes.[111] C'est cette délimitation ou relativisation des règles déductives logiques que les universalistes ne pouvaient concevoir, d'après Hintikka. La raison de cette impossibilité constitue une des trois conséquences que Hintikka associe à la thèse de la logique comme langage (ou médium) universel :

> D'après la conception universaliste, le langage (*le* langage, 'le seul langage que je comprenne' selon les termes de Wittgenstein) est un intermédiaire inévitable entre moi et le monde, un médium dont je ne peux me dispenser. Je ne peux pas marcher pour ainsi dire hors des plates-bandes de mon langage (et du système conceptuel qu'il incarne) et le considérer depuis l'extérieur.[112]

---

[110] [Hintikka 1997a], p. 5. L'auteur compare son travail d'historien de la logique à celui d'une *déconstruction* des présuppositions majeures de la logique moderne. Les deux dogmes de l'universalisme sont des cas du genre, mais Hintikka prétend de plus reconstruire ces idées en vue de résoudre des problèmes de philosophie. Dans le cas qui nous occupe à présent, comment reconstruire un langage formel, et sur la base de quelle méthode d'interprétation ?

[111] "Les logiciens de la tendance algébriste, qui continuent l'œuvre de Boole [1847], n'ont pas de *langage* formel : ils écrivent des équations ; pour exprimer une proposition, ils égalent une équation à 0 ou à 1 (resp. le faux et le vrai). Ils mettent les structures *avant* le langage qui les décrit, disons, avant les axiomes : cela représente un facteur de *pluralité* ou une pluralité potentielle. (L'école algébriste, par L. Löwenheim et Th. Skolem, est à l'origine de la sémantique ou théorie des modèles)." [Largeault 1993], p. 11 ; nous soulignons.

[112] [Hintikka 1997a], p. 8.

Nous reviendrons plus en détail sur sa relecture des présupposés universalistes de Quine contre la logique modale, mais notons dès à présent le lien établi entre universalisme et *ineffabilité de la sémantique*. Cette caractéristique, intimement liée au discours du *Tractatus* de Wittgenstein, est une conséquence naturelle de la thèse universaliste :

> D'après la conception universaliste, je ne peux donc pas examiner dans mon langage les relations qui le relient au monde. Ces relations constituent les significations des mots et autres expressions de mon langage. Leur totalité est ce qui connu sous l'appellation de *sémantique* de ce langage. Par conséquent, un des corollaires les plus importants de la position universaliste est l'*ineffabilité de la sémantique*. Puisque les significations de nos mots et expressions, au sens de la signification en tant que distincte de la référence, sont nos concepts, un universaliste est tenu de croire au caractère inexprimable de toutes les vérités conceptuelles.[113]

Par opposition à la logique comme médium universel, il oppose la logique comme calcul ou théorie *des modèles* (au pluriel):

> La conception opposée, la vision du langage comme calcul, peut être caractérisée de la façon la plus brève en disant que d'après elle toutes les choses qu'un universaliste conçoit comme impossibles sont possibles.[114]

La possibilité d'interpréter autrement les expressions dans un domaine, par exemple, représente une façon de procéder que le logicien universaliste ne saurait concevoir : le langage formel est le même que *le* langage commun à tous les locuteurs, la seule différence résidant dans la redéfinition plus précise de ses termes ; il n'y a pas de sens à parler de plusieurs langages distincts, et l'habitude prise depuis Tarski, qui consiste à parler de logique en termes de vérité ou de démonstration *relativement* à un modèle, est un symptôme de la logique comme calcul. La logique des modalités repose sur cette conception des domaines d'interprétation multiples.

L'histoire de la logique symbolique moderne est partagée entre ces deux courants doctrinaires : la logique comme langage universel trouve son expression notamment dans les écrits de Frege, (le premier) Wittgenstein, Russell, Tarski et Quine ; la logique comme calcul est introduite par le courant algébriste de Schröder et Boole, puis par Peirce, Skolem, Löwenheim, Carnap, ainsi que Hintikka. Le courant relativiste (qui insiste sur le calcul) existait avant l'avènement de la logique universelle selon Frege et Russell ; mais la prise de pouvoir historique de la logique mathématique ou logistique au tournant du 20e siècle expliqua alors le déclin provisoire de l'approche algébrique et la victoire provisoire de l'universalisme en logique. Cela étant, le problème des *antinomies* avait contribué en un sens à renouveler la conception de la logique comme calcul : la résolution de cette antinomie par la théorie des types ramifiés de Russell exigeait la distinction entre plusieurs types de termes et plusieurs niveaux de langage correspondants, conduisant apparemment à séparer plusieurs types de modèles pour chaque niveau de langage. En réalité, le problème des antinomies, s'il a conduit à une distinction entre deux *niveaux* de discours, n'entraîne pas nécessairement une

---

[113] *Ibid.* La distinction fregéenne, admise par Quine, entre le 'sens' d'une entité linguistique et sa 'référence' est un autre symptôme de son universalisme, d'après Hintikka. Voir section 2.3.1.
[114] *Ibid.*

distinction entre plusieurs modèles ou *domaines* de discours : le domaine de Russell restait unique et contenait tous les objets du monde, bien que la scission entre langage-objet et métalangage ait motivé par ailleurs la théorie des modèles et son idée centrale d'une multiplication des domaines de discours possibles.[115] Plusieurs logiciens acquis à la distinction tarskienne entre langage-objet et métalangage demeureront ainsi de fervents universalistes, à l'image de Quine ; rappelons ainsi qu'un des arguments de l'Américain contre la logique modale consistait précisément à distinguer le conditionnel du langage-objet et l'implication du métalangage. Cela veut dire que la thèse de l'universalisme ne repose pas uniquement sur l'absence de métalangage mais sur l'idée générale qu'une certaine relation entre le langage et le monde s'impose au sein de notre langage ordinaire. Mais de quel langage s'agit-il ici ? Chez Wittgenstein comme chez Frege puis Quine, la sémantique formelle est soit impossible, soit redondante : soit les liens entre les objets ou faits du monde et les termes ou énoncés du langage sont inexprimables (Wittgenstein) ; soit ils sont exprimables, mais le résultat de cette expression est redondant car il ne fait que reformuler dans un langage artificiel ce qui est déjà présupposé dans le métalangage, c'est-à-dire le langage ordinaire. Nous avons en effet cette impression de déjà-dit inutile, lorsque les conditions de vérité des connecteurs logiques sont exprimées au sein d'une sémantique formelle : l'énoncé 'A $\wedge$ B' est vrai si et seulement si 'A' est vrai et 'B' est vrai. Quelle différence entre le *definiens* et le *definiendum*, ici, mis à part le symbolisme utilisé ? En d'autres termes, ce que l'on appelle la convention T de Tarski ne représente pour les universalistes qu'une retranscription formelle des règles d'interprétation appliquées dans *le* langage, par le biais d'une fonction d'interprétation *f* des expressions prédicatives. Puisque le domaine sur lequel *f* s'applique est unique, le choix d'interprétation est unique. De même pour la définition des constantes logiques : la fonction de valuation *v* des expressions propositionnelles ne représente pour les universalistes que la retranscription formelle de *la* signification des connecteurs logiques dans *le* langage, et ces règles sont supposées invariables : '$\wedge$' est 'et', '$\vee$' et 'ou', sans qu'il y ait de discussion possible sur le choix des règles d'assignation de valeur de vérité. Ce seul et unique langage-objet détermine le champ d'application de la logique classique des Quine, Frege et Russell ; mais d'autres types de modèles pourraient être imaginés– si l'on considère que la logique classique n'est pas la meilleure en vue d'une compréhension optimale, c'est-à-dire d'une théorie de la signification exhaustive des énoncés ordinaires.[116]

---

[115] Deux remarques à la décharge des universalistes. Premièrement : la préférence pour le calcul n'implique pas nécessairement le relativisme logique, à l'image de George Boole pour qui la signification des connecteurs (c'est-à-dire les règles de valuation de *v*) s'imposait d'elle-même et demeurait indiscutable. Deuxièmement : bien que nous ayons montré ci-dessus que le problème des antinomies ne constitue pas un réel danger pour l'universalisme logique, on peut ajouter que sa résolution n'exigeait pas selon Wittgenstein le recours à une théorie des types ramifiés parce que l'antinomie venait selon lui d'une confusion symbolique ; voir le *Tractatus*, §§ 3.331-3.333.

[116] L'universalisme logique peut trouver aussi son expression dans d'autres systèmes non-classiques : dans la logique paraconsistante, pour Ross Brady ; dans la logique intuitionniste, pour Dummett ; dans la logique relevante, pour Belnap et Anderson. Contrairement à Hintikka, ces logiciens cherchent à modifier puis imposer d'autres règles d'interprétation des connecteurs logiques. L'universalisme peut donc s'exprimer par différentes fonctions de valuation *v* ; cette pluralité implique des rivalités sur le choix des connecteurs déviants et, en ce sens, la logique modale épistémique de Hintikka est d'inspiration classique.

Qu'est-ce qui distingue nettement les deux conceptions de la logique ? Les universalistes interdisent toute réinterprétation des termes de leur langage formel, et ce langage est selon eux unique. Les relativistes (partisans de la logique comme calcul) admettent au contraire la libre réinterprétation des éléments du langage formel et acceptent par conséquent qu'il y en ait plusieurs. Mais que faut-il entendre par le fait de réinterpréter les termes et de rejeter le critère (2') ? Si l'on parcourt les articles de Hintikka relatifs à la logique comme calcul, la réinterprétation porte sur la fonction d'interprétation *f* des termes *non*-logiques et non sur les constantes logiques : la révision en question n'est donc pas une révision des termes logiques, et la question de la logique déviante n'est pas concernée par ce genre de révision. Autrement dit, Hintikka s'intéresse davantage à la relation multiple entre une expression et sa référence dans un domaine $\mathcal{D}$ qu'à la définition multiple des connecteurs logiques. D'où son maintien des constantes logiques classiques, bien que ces constantes puissent être relativisées au sein d'un calcul et d'une structure algébrique.[117] En modifiant les propriétés des objets (les variables de prédicats que ces objets satisfont) dans un domaine, c'est évidemment la notion de *monde possible* que Hintikka cherche à justifier et opposer à la thèse d'ineffabilité de la sémantique : il est possible d'exprimer dans un métalangage la relation entre un terme et l'objet auquel il réfère. Mais la thèse de l'ineffabilité est ambiguë : veut-elle dire qu'il est impossible d'exprimer cette relation d'assignation de référence ou qu'il est possible d'en parler, mais que le résultat de ce discours est soit redondant soit trivial ? Tandis que Frege rejetait la sémantique formelle pour des raisons d'impossibilité, c'est pour des raisons de trivialité qu'il faut comprendre la position de Quine. D'un côté, ces deux logiciens classiques admettaient l'idée de système formel, contrairement à la conception ineffable que Wittgenstein se faisait des règles de logique dans le *Tractatus*. D'un autre côté, l'idéographie de Frege avait pour but non pas de modifier mais de clarifier les relations conceptuelles au sein d'une syntaxe. Ces relations mathématiques étaient supposées dériver d'un fonds commun du langage universel et dont les lois de l'être vrai seraient l'expression formelle ; la raison de cette ineffabilité et l'absence consécutive de toute façon de voir sémantique chez Frege est que, pour un universaliste,

> les résultats de cette 'façon de voir' doivent être exprimables dans notre langage. Or ce langage présuppose dans tous ses usages certaines relations sémantiques (relations de représentation) entre le langage et la réalité. (Sans cela, nous ne pourrions pas utiliser le langage dans nos transactions avec la réalité). Mais puisque ces relations sémantiques sont présupposées dans chaque *usage* du langage, elles ne peuvent pas être exprimées dans le langage.[118]

---

[117] Une structure booléenne, du nom de George Boole, est une algèbre $\langle \wedge, \vee, \neg, \rightarrow, 1\ 0 \rangle$ définie comme un treillis borné ($1 \vee a = 1$, $0 \wedge a = 0$), distributif ($[(a \vee b) \wedge c = (a \wedge c) \vee (b \wedge c)]$), $[(a \wedge b) \vee c = (a \vee c) \wedge (b \vee c)]$) et complémenté ($a \vee \neg a = 1$, $a \wedge \neg a = 0$). Il existe bien d'autres types de structures algébriques, parmi lesquelles la structure de Heyting (non complémentée, entre autres). Dans un calcul des propositions, l'opposition entre ces deux structures correspond à la distinction entre le calcul classique et le calcul intuitionniste. Hintikka maintient sa préférence pour le premier et ne propose pas de réinterprétation de ces structures : ce n'est pas la fonction de valuation *v* qui serait à changer en priorité.

[118] [Frege 1971], p. 718. [van Heijenoort 1967] confirme l'absence des questions métalogiques chez Frege, en conséquence de sa thèse d'ineffabilité de la sémantique : "Une autre conséquence importante de l'universalité de la logique est que rien ne peut être dit ou ne doit être dit situé en dehors du système. Et en effet, Frege ne soulève jamais de question métasystématique (consistance, indépendance des axiomes, complétude). Frege est parfaitement conscient en fait que tout système formel exige des règles qui ne sont pas exprimées dans le système". (p. 326)

Vouloir exprimer ces relations de représentation, c'est vouloir attraper sa propre ombre, pour ainsi dire.

Mais que fait un métalangage lorsqu'il définit la fonction d'interprétation f par une application des expressions d'un langage dans un domaine $\mathcal{D}$, sinon exprimer dans un 'mode formel de discours' cette relation de représentation supposée inexprimable ? Encore une fois, la critique universaliste de la sémantique formelle semble porter davantage sur la prétention à *expliquer* les relations sémantiques dans un métalangage plutôt que sur le prétention à simplement les exprimer.[119] Quine ne contestait nullement la sémantique formelle de Tarski, après tout, laquelle consiste pourtant à sortir d'un langage pour parler de ses relations sémantiques. Mais paradoxalement, Hintikka classe Tarski lui-même, promoteur de la théorie des modèles, parmi les partisans de l'universalisme logique, pour la raison que le Polonais ne croyait pas à la possibilité de définir le concept de vérité dans le langage ordinaire, qu'il considérait lui aussi comme universel.[120] Or si les relations sémantiques ne sont pas exprimables dans ce fond commun de la pensée qu'est supposé être le langage ordinaire, comment justifier ces relations et faire mieux que de les stipuler au sein de langages formels, arbitraires ? Nous retombons dans le relativisme impliqué par la Tolérance de Carnap, et que l'on puisse construire en pratique des systèmes formels distincts n'est donc pas un argument suffisant pour rejeter la thèse de l'ineffabilité de la sémantique. Il s'agit plutôt de savoir si notre pratique sémantique du langage ordinaire (notre *Ursprache* naturelle) est, elle, explicable au sein d'une sémantique formelle. Pour les universalistes, elle ne l'est pas. Pour certains relativistes, elle peut l'être (dans sa totalité ou par fragments), et la logique philosophique en exprime la prétention. Le reproche que Hintikka adresse à Quine porte en particulier sur son traitement de la théorie des modèles, que le dernier ne considère pas comme impossible mais superflu dans le cadre d'une théorie logique :

> Ce n'est pas un hasard si, parmi les dizaines d'articles que Quine a publiées, il n'y en a pas un seul consacré à la théorie des modèles. (...) Les modèles sont pour Quine un simple procédé technique qui peut être utilisé par exemple pour étudier l'indépendance mutuelle des axiomes déductifs, rien de plus.[121]

Quel avantage explicatif la théorie des modèles apporte-t-elle en sémantique, d'après Hintikka, et quel changement décisif apporte-t-elle par rapport à la sémantique classique ?

Hintikka présente le problème de la sémantique comme un problème de *représentation* des relations de la réalité au sein du langage, mais il reste à déterminer le mode d'expression de

---

[119] [Hintikka 1979] rappelle que "d'après Frege, 'on ne peut pas dire : une représentation est vraie lorsqu'elle s'accorde avec la réalité' (Voir *Logik*, 1897)" (p. 720). Or cette thèse de Frege est considérée comme signifiant que la relation langage-monde est injustifiable ou indémontrable hors du langage dans lequel cette relation est énoncée. Le cœur de l'argument d'ineffabilité, nous insistons, semble porter ainsi sur l'impossibilité de *justifier* ces relations ou de garantir leur *sens*.

[120] Contre la thèse de l'inexprimabilité d'un concept consistant de vérité à l'intérieur de la langue naturelle, Hintikka prétend que sa logique IF permet au contraire d'obtenir une définition de la vérité au sein d'un seul et même langage et, par conséquent, de surmonter la thèse de Tarski. Peut-on *justifier* et, donc, surmonter la thèse de l'ineffabilité de la sémantique par la logique IF? Ce problème sera effleuré plus tard, dans le cadre de la thèse quinienne de l'indétermination de la traduction et du principe de charité.

[121] [Hintikka 1997b], p. 468.

ces relations au sein d'une théorie des modèles. Répétons-le : si l'on regarde ses travaux en logique modale épistémique, on constate que le changement apporté par rapport à la logique classique concerne uniquement la fonction d'interprétation des termes *non*-logiques : la signification des formules complexes est toujours définie en termes de conditions de vérité des connecteurs classiques, mais la valuation s'effectue désormais dans plusieurs modèles *distincts* et justifie le recours à des opérateurs modaux. C'est cette idée simple de multiplier les points de vue sur le monde que les partisans de l'universalisme logique ne pouvaient concevoir, selon Hintikka.[122] On peut expliquer cette absence par la fonction que la majorité des universalistes associaient à leur théorie logique : une théorie mathématique, pour Frege et Russell, une théorie ontologique pour Wittgenstein, une théorie du langage scientifique pour Quine ; à l'inverse, la majorité des logiciens modaux issus de la tradition de Montague ou Kripke sont nourris de préoccupations portées sur la linguistique et la description du langage ordinaire ; la distinction précédente de Hare entre phrastique et neustique, pertinente dans une théorie des actes de discours, devient concept superflu aux yeux d'un théoricien porté sur l'étude des nombres ou la conception scientifique du monde. Mais n'en déplaise à Quine, Hintikka prétend que cette autre conséquence de son universalisme logique, *l'hypothèse du monde unique*, ne s'impose pas non plus dans la construction des théories scientifiques :

> Par exemple, un théoricien des probabilités considère un état de choses ou cours des événements réel, considéré peut-être comme un des 'petits mondes' de Jimmy Savage [*Foundations of Statistics*, 1954], comme un simple point dans l'espace d'étude. (...) Ce que vise un scientifique qui emploie des concepts probabilistes ou modèles-théorétiques peut très bien être en dernière analyse la description et, sinon, la maîtrise intellectuelle de *notre seul et unique monde réel.* Mais les moyens de procéder du scientifique comportent très souvent des conceptualisations qui transcendent les frontières de ce monde. Cela se produit par exemple dès que des concepts probabilistes sont employés, comme c'est le cas dans une bonne partie de la science physique actuelle.[123]

En conclusion de cette étude des présuppositions en philosophie de la logique, nous remarquerons que la contribution de Hintikka au débat du monisme-pluralisme ne plaide ni pour un pluralisme global, ni pour un pluralisme local : sa sémantique relationnelle ne modifie pas la fonction de *valuation v* pour les énoncés complexes, c'est-à-dire la signification des connecteurs (mots logiques) reliant des énoncés atomiques en termes de conditions de vérité ;

---

[122] On retrouve ce présupposé du modèle unique dans la logique de Frege : "Frege, qui ne détestait rien tant que la pluralité d'interprétations, où il blâmait l'ambiguïté, estime que la conception algébrique renverse l'ordre naturel des choses ; il met le langage et les axiomes avant la structure". ([Largeault 1993], p. 11). [van Heijenoort 1967] confirme que "l'universalité de la logique exprime en soi une caractéristique importante du système de Frege. Dans ce système, les quantifications liant les variables d'individus parcourent *tous* les objets. (p. 325). [Largeault 1993] prétend de même que "Frege (*Begriffschrift*, 1879) concevait la logique comme une théorie 'absolue' : l'univers du discours doit être *unique* et les variables d'individu prennent pour valeurs n'importe quels objets, tandis que les variables de prédicat parcourent toutes les propriétés. Cette logique n'a qu'un modèle, au lieu, comme celle de l'école algébriste, de comporter une pluralité d'interprétations avec des domaines et des relations différents. L'adhésion à un principe de monovalence est surprenante, étant donné qu'on connaissait, en géométrie et en algèbre, des exemples de théories dotées de plusieurs modèles. Le système de Frege a *une seule* interprétation, qu'on pourrait appeler la structure logique du monde, aussi le premier ordre n'y est-il pas disjoint de ce qui le dépasse". (p. 12)

[123] [Hintikka 1997b], p. 459. Considérer le monde réel comme un domaine de référence(s) parmi d'autres possibles, en théorie des probabilités, c'est considérer le monde de référence $w_0 \in W$ comme un élément dans l'ensemble W des mondes possibles w.

elle porte seulement sur la fonction d'*interprétation f*, c'est-à-dire sur la relation entre les expressions non-logiques d'un langage et leurs référents assignés dans un modèle. A cet égard, la logique modale épistémique n'implique pas un changement des connecteurs logiques mais du domaine de référence, et seule cette relation terme à objet compte dans sa conception de la logique comme calcul. La multiplication des mondes dans un modèle ne signifie pas une division entre plusieurs catégories d'expressions dans des sous-modèles, ce qui remettrait en cause l'idée de vérité logique comme vérité dans tous les modèles : la modification de $\mathcal{D}$ consiste à passer d'un monde unique à une pluralité de mondes, et chacun d'eux ne représentent pas des parties d'un domaine unique mais plusieurs domaines entiers. Le domaine de discours de la logique modale est tout aussi universel que celui de Quine ou Frege, donc, mais il n'est plus unique et peut être varié selon le contexte (aspect neustique) du discours.[124] Contrairement aux pluralistes tels que Restall ou Beall, conçus dans le sens d'un relativisme des constantes logiques, la logique des attitudes propositionnelles de Hintikka reste basée sur une *infrastructure classique*, pour reprendre les termes de [van Benthem 1991].[125] Dans le cadre de sa sémantique des jeux GTS, consécutive à sa sémantique des modèles, le Finlandais insistera de nouveau sur l'idée qu'il n'est pas question de modifier la signification des connecteurs logiques pour comprendre le sens de certains énoncés non-vérifonctionnels, comme c'est le cas dans la logique intuitionniste de Heyting ; on trouve dans le passage qui suit une issue au relativisme induit par le Principe de Tolérance de Carnap :

> A la différence des changements dans notre logique que Quine déclare possibles, ces variations ne sont pas de simples possibilités théoriques abstraites. En effet, j'ai déclaré que les logiques

---

[124] Les passages de [van Heijenoort 1967] qui suivent montrent la filiation entre la sémantique relationnelle et la logique comme calcul héritée de Boole et Schröder : "Pour Frege, l'ameublement ontologique de l'univers se divise en objets et en fonctions. Boole a sa classe d'univers, et De Morgan son univers de discours, dénotés par '1'. Mais ceux-ci n'ont guère de teneur ontologique. Ils peuvent être *changés à volonté*. L'univers de discours comprend seulement ce que nous nous accordons à considérer à un certain moment, dans un certain contexte. Pour Frege, il ne peut pas être question de *changement d'univers*. On ne pourrait même pas dire qu'il se limite à *un* univers. Son univers est *l'*univers. Pas nécessairement l'univers physique, bien sûr, parce que pour Frege certains objets ne sont pas physiques. L'univers de Frege se compose de tout ce qu'il y a, et il est fixe." (p. 325) ; "Alors que l'approche Frege-Russell des fondations de la logique pourrait être qualifiée d'approche *axiomatique*, celle de Löwenheim pourrait être qualifiée d'approche *ensembliste*. Si nous suivons cette approche, les questions de validité des formules bien formées dans des domaines différents sont mis en exergue. Le titre même de l'article, *Über Möglichkeiten im Relativkalkül*, fait référence à l'idée suivante : si une formule est valide dans un domaine, elle peut être ou ne pas être valide dans un *autre* domaine. (p. 327) ; "Avec l'article de Löwenheim, nous avons une profonde rupture avec l'approche Frege-Russell des fondations de la logique et un retour à, du moins une connexion avec la logique pré-frégéenne ou non-frégéenne. Löwenheim utilise la notation logique de Schröder, mais ce qui est plus important, il prend avec Schröder également la liberté de *changer l'univers de discours à volonté* et de baser des réflexions sur ce genre de changements. Et de même que Frege fut ignoré pendant un certain temps du fait de cette rupture avec la tradition en vigueur, de même Löwenheim fut ignoré lui aussi pendant un certain moment du fait de cette rupture avec la nouvelle tradition en vigueur. Derrière la tendance Frege-Russell en logique, Löwenheim renoue contact avec Boole et Schröder, tout en apportant d'importantes contributions à part entière à la logique". (p. 328)

[125] Sur le pluralisme logique, voir notamment J.C. Beall J.C. & G. Restall : "Logical Pluralism", *Australasian Journal of Philosophy* **78**(4) (2000), pp. 475-93. Quant à la logique modale épistémique de Hintikka, [van Benthem 1991] la décrit comme "une superstructure modale au sommet d'une sémantique classique ordinaire", peu soucieuse de "réanalyser cette dernière en profondeur. Pour introduire une distinction, le paradigme Hintikka-Halpern est un paradigme de l''épistémique extrinsèque', lequel n'affecte pas la sémantique classique standard, en témoigne se définition clef de la vérité (...) Par opposition, une 'épistémique intrinsèque' plus radicale ne considérait pas la sémantique classique comme allant de soi et réanalyserait la notion de vérité dans son ensemble, à la lumière de considérations épistémiques" (p. 6).

constructivistes devraient être considérées comme obtenues de cette façon même. Même le type de changement particulier dans nos jeux sémantiques est intéressant, dans ce cas. Au lieu de changer les règles de signification des jeux sémantiques, nous devrions restreindre les stratégies accessibles pour les joueurs aux stratégies calculables (récursives).[126]

Nous remarquerons dans ce qui suit que l'une des principales objections adressées dans la littérature contre la logique épistémique porte sur la validité relative de ses théorèmes : une vérité logique ne doit-elle pas être vraie dans tout l'univers de discours ? Armé de cette distinction historique entre une logique comme médium universel et comme calcul, la distinction entre vérité logique et vérité dans un *modèle* devrait permettre de clarifier la situation et de répondre aux objecteurs. La majorité d'entre eux partagent selon nous l'hypothèse du monde unique, conséquence de leur universalisme en logique. Arrêtons-nous donc sur ces objections récurrentes, d'abord convaincantes mais qui perdent de leur impact à l'aune de la déconstruction historique proposée ci-dessus par Hintikka et van Heijenoort.

2.1.5 Vérité logique et validité dans une structure

La distinction entre plusieurs domaines de discours a produit une distinction entre deux notions jusque là synonymes : la *vérité logique* (dans tous les modèles) et la *validité* (dans une classes de modèles).Tant que le modèle $\mathcal{M}$ représente un domaine $\mathcal{D}$ unique, ce qui est le cas en logique classique (non modale), toute vérité logique est une formule valide ; mais une fois que l'on admet la réinterprétation des expressions dans un domaine, le modèle prend plusieurs expressions possibles et implique qu'une formule peut être valide sans être logiquement vraie.

Prenons un exemple : soit $\mathcal{M}$ un modèle composé de deux 'mondes' $w = \{p, q, r\}$ et $w^* = \{p,r\}$. La formule '$p \wedge q$' sera vraie dans $\mathcal{M}$ parce qu'elle est vraie pour toute interprétation dans les mondes de ce modèle : $v(w,p \wedge q) = 1$ et $v(w^*,p \wedge q) = 1$, donc $v(W,p \wedge q) = 1$. Mais on ne dira pas de '$p \wedge q$' qu'elle en devient une tautologie pour autant. Par conséquent, une formule vraie dans un modèle n'est pas forcément une formule logiquement vraie. On peut élargir le champ de validité d'une formule : une formule est valide, c'est-à-dire vraie dans tous les modèles d'une structure, si elle est vraie pour toute interprétation dans un modèle quelconque de $\mathcal{M}$. La formule '$p \wedge q$' ci-dessus n'est pas un exemple de formule valide dans une structure, parce que l'on peut facilement trouver un contre-modèle à cette formule : un troisième domaine dans lequel on trouve soit $\neg p$ soit $\neg q$, ou dans lequel on ne trouve ni p ni q. La distinction entre vérité dans un modèle, validité dans une structure et validité modale (dans toute structure de Kripke) va être abordée dans ce qui suit : elle a servi d'argument à certains opposants à la logique modale épistémique qui n'admettaient pas (sciemment ou inconsciemment) la distinction entre vérité logique et validité. Peut-on caractériser la vérité logique comme ce qui est valide dans *n'importe quelle* structure (ou modèle), c'est-à-dire indépendamment des propriétés de R et du domaine d'interprétation ?

L'allusion faite ci-dessus au Principe de Tolérance de Carnap avait pour but de mettre en évidence la difficulté d'une logique philosophique : comment saisir la signification des

[126] [Hintikka 1999], p. 20. Voir également [Hintikka 2001], où l'auteur prétend que la notion intuitionniste de 'construction' peut être redéfinie dans sa logique IF (dont l'infrastructure est classique) par l'axiome du choix et en termes de fonctions de Skolem. Mais un logicien intuitionniste pourrait objecter à cette méthode qu'elle présuppose des moyens rejetés par la philosophie intuitionniste (la formulation d'une preuve ne signifie pas la construction effective de cette preuve).

énoncés du langage ordinaire au sein d'un langage formel ? La 'caractéristique réelle' que Leibniz espérait concevoir est depuis longtemps réduite dans ses prétentions. Mais la réduction du langage formel à une combinatoire de signes révisable à l'envi risque de conduire à un relativisme, sinon à un scepticisme sur l'application de la notion de vérité, avons-nous dit : supposons d'une formule quelconque ϕ qu'elle soit vraie dans un langage $\mathcal{L}_1$, fausse dans un langage $\mathcal{L}_2$ et dépourvue de sens (ni vraie ni fausse) dans un langage $\mathcal{L}_3$; ϕ peut-elle être logiquement vraie, ou la classe des vérités logiques constitue-t-elle une classe particulière de formules, la classe des formules vraies dans *tout* langage $\mathcal{L}$? Plutôt que d'aborder la question du pluralisme logique en général, nous nous limiterons au problème des vérités logiques en logique modale épistémique. A première vue, il semble que la relativité des vérités logiques soit incompatible avec l'exercice de la logique philosophique. En effet, pourquoi se fier à un langage formel dès lors que ses règles de fonctionnement sont révisables donc improbables ? Il s'agit de trouver un moyen par lequel concilier la pluralité des systèmes logiques et la pertinence de leurs enseignements.

Ce qui suit va donc se concentrer sur la question de l'*adéquation informelle* d'un langage formel : étant donné la relativité des systèmes formels, dans quelle mesure la logique modale épistémique issue de [Hintikka 1962] a-t-elle une réelle valeur philosophique, c'est-à-dire une capacité à saisir la signification 'caractéristique' (s'il y a) des concepts philosophiques de savoir et de croyance ? Avant d'aborder la réponse méthodologique de Hintikka sur ce point, nous reviendrons sur les notions de vérité et de validité, de modèle et de structure. Si la vérité logique est relative au choix d'une structure particulière, existe-t-il un critère pour justifier ce choix en *logique philosophique* ?

2.1.5.1 Tolérance formelle, adéquation informelle

Comment justifier le contenu d'une classe de vérités logiques, et dans quelle mesure un langage logique n'est-il pas un langage formel quelconque? Comment s'effectue la construction d'un système de logique philosophique, et pour quelle raison formalise-t-on des concepts ordinaires ? Pour comprendre cette méthode, commençons par revenir sur ce qui constitue la vertu première de la méthode d'analyse logique, et voyons si celle-ci est préservée dans les logiques philosophiques.

L'intérêt accordé à la démonstration logique réside dans la fiabilité de ses résultats : la conclusion est nécessairement tirée des prémisses, en vertu de règles d'inférence claires et distinctes. Maintenant, cette nécessité est à entendre en deux sens : un sens absolu, qui signifie que la conclusion est vraie en tant que telle ('en soi') ; un sens relatif, qui signifie que la conclusion est vraie relativement à ou sur la base de prémisses.[127] Qu'en est-il en particulier pour les théorèmes de la logique modale épistémique ? Ce sont manifestement des vérités relatives : que ce soit en syntaxe ou en sémantique, les axiomes ou tautologies obtenus sont

[127] La nécessité absolue correspond à la vérité *apodictique* de la logique traditionnelle, la nécessité relative à la vérité *hypothético-déductive* des logiques modernes. On ne parlerait pas de 'logiques philosophiques' si la logique se résumait à quelques principes fondamentaux et apodictiques, telle que la logique était conçue il y a encore peu (composée des trois piliers que seraient tiers exclu, non-contradiction et identité). Mais ce qui justifie l'appellation de logique philosophique risque dans le même temps de rendre ses résultats contestables : comment garantir les théorèmes d'un système de logique modale dont les règles sont toutes hypothétiques, donc révisables ?

révisables, bien que leur caractère de vérités générales soit supposé au départ ; il devient recommandable de les réviser lorsqu'un paradoxe (une conclusion inadmissible obtenue sur la base de prémisses admises) apparaît dans le système formel. Le rôle du paradoxe dans les logiques philosophiques (épistémique, temporelle, déontique, etc.) montre bien le caractère révisable des axiomatiques et des modèles sémantiques, construits parallèlement afin de satisfaire le critère métalogique d'adéquation.[128] Mais ce n'est pas cette sorte d'adéquation, l'adéquation formelle appelée aussi complétude, qu'une logique philosophique a le devoir d'apporter en priorité : pour reprendre les termes de Hintikka, la sémantique formelle a pour but d'apporter un *éclairage* sur certains termes fondamentaux de la théorie de la connaissance : croyance, connaissance, justification. L'utilisation de la logique formelle pour traiter ces notions a l'avantage d'apporter ce que Leibniz exprimait par sa métaphore du 'fil d'Ariane' : le caractère explicite et rigoureusement défini des règles d'inférence (de transformation des énoncés) permet de montrer *clairement* par quelles étapes la vérité d'un énoncé est obtenue à partir d'autres énoncés admis. Maintenant, les attitudes propositionnelles peuvent-elles être définies par des propriétés axiomatiques aussi générales que pour l'espace en géométrie ou les nombres en arithmétique ? Non seulement une logique épistémique suppose que les verbes 'croire' et 'savoir' peuvent être caractérisés de la sorte, mais le langage logique suppose également que les théorèmes obtenus devront être vrais dans *n'importe quel* modèle. Ce projet est-il réalisable, ou faut-il préférer une mission plus modeste consistant à construire plusieurs logiques épistémiques possibles suivant les différentes significations des verbes d'attitude ?

Dans l'idéal, le but de la logique épistémique est de construire un langage dans lequel les règles d'inférence reproduiraient les vérités intuitives relatives aux énoncés de savoir et de croyance. Cet idéal est formulé dans [Haack 1978] en termes de corrélation entre la validité logique d'un énoncé, au sein d'un langage formel, et sa validité extra-logique au sein du langage ordinaire. L'idée de corrélation initiale en logique philosophique est proche de la distinction médiévale entre deux types de 'logique', entre une logique de sens commun et une logique plus rigoureuse :

> A la suite de Peirce (lequel a emprunté à son tour la terminologie aux logiciens médiévaux), on peut appeler notre jugement irréfléchi de la validité des arguments informels le *logica utens*, et les jugements plus précis et rigoureux, tels que conçus dans les systèmes formels à partir d'une réflexion sur ces jugements, le *logica docens*.[129]

La notion de validité extra-logique pose problème, puisqu'elle suppose que le langage ordinaire (non artificiel) contient des règles par lesquelles on puisse démontrer qu'un énoncé ordinaire est toujours vrai. Or la construction de langages formels ayant précisément pour but d'apporter des règles de démonstration dont le langage ordinaire ne dispose pas ou qu'il ne montre pas explicitement, comment comparer la validité d'un énoncé entre deux types de langage dont l'un ne contient même pas de méthodes de démonstration ? L'idée de corrélation

[128] Dans [Meyer 2000], qui donne des exemples de logique épistémique où plusieurs axiomes combinés donnent des théorèmes contre-intuitifs, tous les auteurs ne choisissent pas le même ensemble d'axiomes pour résoudre le problème. La situation est comparable au trilemme de l'Argument du Dominateur.
[129] [Haack 1978], p. 16.

ci-dessus s'oppose par ailleurs à la fonction initiale de la logique : à savoir, éliminer les ambiguïtés du langage ordinaire et remplacer la règle directrice de l'intuition, douteuse et parfois contradictoire, par la méthode de justification déductive. En l'absence de corrélation possible entre les énoncés des deux types de langages, c'est notre notion même d'adéquation informelle qui risque de perdre son sens ; par défaut, il faut entendre par cette adéquation l'idée d'un langage formel dont les théorèmes interprétés rendraient compte de vérités 'intuitivement admises' dans le langage ordinaire, mais sans que la corrélation établie entre les deux types d'énoncés soit elle-même justifiable. Aucune garantie ne semble être donnée pour conforter ces intuitions, sinon un accord global ou convenu sur les résultats obtenus par un système logique. En ce sens, l'intérêt d'une logique philosophique particulière repose sur la confiance que nous accordons à ses règles d'inférence ainsi que sur les conclusions inattendues qu'elle est capable de faire apparaître par déduction.

Dans la pratique, la logique philosophique fonctionne ainsi selon une procédure 'empirique' de méthode et essai, dans laquelle le choix des axiomes ou des tautologies est révisable en fait.[130] Nous constaterons que c'est également le cas en logique épistémique, lorsqu'une conclusion aussi paradoxale que l'omniscience logique se produit dans les structures normales de Kripke. Pour cette raison, on peut considérer la logique philosophique comme une construction *a posteriori* de structures *a priori* : en amont, des règles d'inférence de la logique pure permettent des opérations purement abstraites qui, en aval, peuvent être appliquées à des formules interprétées ; le choix de modélisation peut être modifié si l'interprétation donne des conclusions indésirables.[131] L'accord obtenu entre la validité formelle et l'intuition informelle fait de la logique philosophique une discipline mobilisant les compétences à la fois du logicien, pour ce qui est de la construction du système de logique, et du philosophe, pour ce qui est de l'appréciation des résultats de son application sur des énoncés du langage ordinaire.

---

[130] [da Costa & Bueno 1996] s'appuient sur cette révisabilité de fait pour justifier l'utilisation de leur logique doxastique paraconsistante : "Une logique exhaustive, appropriée à tous les domaines, est difficile à trouver. Nous en restons en général avec (plusieurs) logiques alternatives qui décrivent seulement certains de leurs aspects, et il y a un grand nombre de raisons heuristiques et pragmatiques de choisir entre de telles logiques, ce qui dépend en particulier des traits spécifiques que l'on trouve dans ce genre de domaines, bien entendu" (p. 54). Une conséquence de cette révision admise est que la nécessité des démonstrations est toujours locale et jamais absolue; aucun système de logique ne peut prétendre à l'universalité d'application, qu'il soit classique ou non-classique : "Au lieu de déclarer, avec les propositions 'radicales' plus anciennes, que la logique classique est un instrument pour appréhender la structure la plus générale du monde (elle est censée être vraie, après tout !), leur nouvelle version 'radicale' déclare la même chose pour autant que cela concerne certaines logiques non-classiques. On s'interroge, dans un tel cas, sur l'importance de l'enseignement issu de l'expérience –c'est-à-dire, de l'histoire récente de la logique. En effet, quelle est la teneur, la pertinence de cette histoire, avec les changements et déplacements qu'elle a produits, pour notre compréhension philosophique de la logique ? Compte tenu de ces circonstances, comment ne pas être *faillibiliste* malgré tout ?" (*loc. cit.*, pp. 53-54)

[131] [Gardiès 1979] résume ainsi l'approche par tâtonnement qui sous-tend les logiques philosophiques : "Les résultats obtenus dans l'analyse des concepts apparemment les plus familiers, comme ceux du possible, de permis ou d'obligatoire, tendraient à montrer que tous les systèmes et les modèles auxquels on parvient sont *paradoxaux*. Après l'exposé de chaque modèle ou de chaque système on serait tenté d'objecter que ce que nous appelons possibilité, permission ou obligation, ce n'est jamais *exactement* cela. Devant quoi la seule réplique valable qui puisse rester à l'auteur du modèle serait, nous l'avions dit, de sommer l'objecteur de lui fournir le modèle correspondant à ce qu'on entend par là : si l'on n'est pas d'accord sur l'*adéquation* du modèle à la chose, qu'on propose un *autre modèle*, à la condition de ne pas se réfugier dans le simple témoignage négatif d'une intuition dont tout laisse à penser qu'elle est *fluctuante*" (p. 232 ; nous soulignons).

C'est ainsi que semble procéder la logique philosophique. Cela dit, la révisabilité des termes d'une inférence en logique ne signifie pas que tout soit constamment révocable au sein d'une démonstration : en quoi une vérité logique demeure-t-elle la plus générale qui soit, à défaut d'être 'universelle', et ne doit-il pas se trouver dans l'inférence quelque chose dont on ne puisse pas discuter, sur quoi repose la correction de la démonstration ? Après tout, la présence d'un paradoxe nécessite une révision des prémisses sur la base d'une règle d'inférence : la réduction par l'absurde, en vertu de laquelle une conséquence contradictoire implique la négation de la prémisse.[132] Cette inférence intermédiaire entre la prémisse et la conclusion n'est pas révisée elle-même en règle générale, car elle permet d'établir le paradoxe et de justifier un changement à l'intérieur des règles du système. Néanmoins, cette même réduction à l'absurde utilise les mots logiques de la conjonction et la négation qui, si l'on suit le pluralisme logique, peuvent être révisées à leur tour. Comment juger l'adéquation (informelle) d'un système de logique philosophique si même les termes de ce jugement sont révisables ? Faut-il faire une distinction entre des notions logiques révisables, les opérateurs, et des procédures logiques non révisables, les règles de substitution uniforme et de détachement ? L'éclosion récente des logiques substructurelles semble montrer au contraire qu'aucune procédure logique n'est à l'abri d'une révision, pas même les règles d'inférence.[133]

L'idée de révision des lois logiques n'est pas incompatible avec leur généralité : une fois qu'une opération déductive est admise (comme par exemple l'élimination de la conjonction ou le modus ponens), elle s'applique nécessairement à n'importe quel objet de discours. Une nécessité toute locale, conformément à l'approche hypothétique des logiques formelles. En d'autres termes, la recherche de l'adéquation informelle d'une logique philosophique, bien qu'elle inclue une procédure de révision du formalisme, conserve en revanche une caractéristique essentielle des langages logiques et qui les distingue d'autres formalismes : celui de s'appliquer à n'importe quel objet. Pour distinguer ainsi la vérité logique, propre à un langage pur, et la vérité non-logique propre à une théorie particulière (géométrie, biologie, physique, etc.), [Granger 1979] exprime la vérité logique en termes de vérité *dans une structure* :

> Le formalisme logique de la modalité, du temps, et a fortiori ceux de la causalité, du commandement et de l'évaluation morale, ont en fin de compte un statut épistémologique ambigu qui les rend plus proches des langues que la logique ordinaire, sans pourtant les rendre aptes à exprimer totalement ce qu'expriment les premières. Ce sont des logiques *sur* des structures, et elles constituent par conséquent déjà des théories d'*objets* qui ont cessé d'être absolument quelconques. En ce sens, elles s'écartent de la pure symbolisation logique, se rapprochant du syncrétisme des langues naturelles. Mais leur construction demeure cependant

---

[132] La forme logique de la réduction à l'absurde est : $[(\phi \rightarrow \psi) \wedge (\phi \rightarrow \neg\psi)] \rightarrow \neg\phi$. Lorsque ϕ est une conjonction de prémisses $A_1, A_2, \ldots A_n$, on obtient $\neg\phi = (A_1 \vee A_2 \vee \ldots \vee A_n)$. Quelle(s) prémisse(s) A faut-il alors rejeter ? Ce problème du choix de la prémisse à sacrifier est caractéristique des dilemmes de la philosophie antique (Argument Dominateur, sorites) ; il représente l'aspect empirique qui entre en ligne de compte lors de la révision d'une logique philosophique.

[133] Les logiques substructurelles sont des logiques non-classiques plus faibles que la logique classique (le nombre des théorèmes est restreint), caractérisées par l'absence d'une ou plusieurs règles *structurelles* de la logique classique (contraction, affaiblissement, etc.). Cela montre que même les règles d'inférence pures, définies sans aucun opérateur, peuvent être modifiées. Nous reviendrons brièvement sur cette famille de logiques affaiblies, dans le cadre de la solution au paradoxe de l'omniscience logique proposée par Dubucs et Marion.

résolument fidèle au paradigme des logiques pures en se limitant à la mise en œuvre des aspects syntaxiques et sémantiques des symboles.[134]

Suite à cette remarque, il faut donc distinguer deux sortes de relativité des énoncés vrais : à une théorie d'objets particuliers, d'une part, et à un langage d'objets quelconques, d'autre part. La première sorte de vérité s'applique seulement à un ensemble partiel d'objets, tandis que la seconde vérité (logique) s'applique à n'importe quel objet. La relativité des logiques s'expliquerait ainsi par une modification des règles d'inférence et une variation dans la liste des mots logiques, mais sans que cette relativité ne remette en cause l'indifférence de la vérité logique par rapport au contenu. Elle reste en ce sens un critère grammatical, et le rôle de l'expérience dans la construction des logiques philosophiques ne concerne pas le choix des objets de discours, mais le choix des structures abstraites susceptibles de décrire adéquatement un concept philosophique.

Malgré tout, Quine n'admettait pas la logique épistémique et maintenait que, quel que soit le degré d'abstraction de ses opérations déductives, les opérateurs de croyance et de savoir restent trop 'colorés' pour faire figure de mots logiques. Sa critique signifie ceci : que le caractère abstrait d'une opération n'est pas une condition suffisante pour faire des opérateurs des mots proprement *logiques* : c'est l'usage ordinaire d'un terme qui justifie sa capacité à être traité comme un mot logique ou non. La logique épistémique ne serait autrement dit qu'un langage formalisé parmi d'autres, mais un langage dont les opérations ne sont valables qu'en fonction d'une certaine interprétation restrictive des opérateurs. D'après Quine, les mots logiques ont un caractère 'obvie' que les modalités ne peuvent avoir au même titre que la négation, le conditionnel, etc. Cet argument est-il recevable ? Il repose apparemment sur le rejet catégorique de la moindre relativité *dans l'usage* des vérités logiques : si tel est le cas, l'idée de vérité *relative à* une structure serait déjà trop circonstancielle pour satisfaire la notion de vérité logique, qui n'est relative à aucune structure en particulier. C'est ce genre d'affirmation que nous allons évoquer dans ce qui suit, puisqu'il semble avoir incité plusieurs logiciens à rejeter le projet de logique épistémique : une vérité proprement logique ne doit pas se définir en termes de structures, donc toute formule pour la validité de laquelle les structures jouent un rôle essentiel n'est pas une formule logiquement vraie.

2.1.5.2 Vérité et structure

Que des formules modales puissent être validées ou non selon l'interprétation donnée à l'opérateur peut paraître surprenant : s'il est besoin de préciser ce que signifie $\Box$ dans $\Box\phi \rightarrow \phi$ pour admettre son application aux langues naturelles, cela ne revient-il pas à traiter la modalité comme un élément tantôt logique, tantôt *extra*-logique ? Pas si l'on considère la vérité logique comme une vérité basée uniquement sur la forme du schéma d'énoncé, quel que soit le sujet du discours. La distinction entre ces deux conditions de vérité générales s'exprime en termes de vérité dans un monde ou dans un modèle : une formule peut être vraie dans un monde (ou état de choses) particulier, fausse dans un autre. C'est le cas des vérités empiriques

[134] [Granger 1979], p. 93. Les logiques modales ne sont pas des logiques au sens de Gonseth ou Granger, dans la mesure où elles portent sur des objets certes quelconques mais dont les propriétés ne sont satisfaites que dans certaines structures. Nous l'avons constaté avec les structures de Lewis et, en ce sens, les théorèmes modaux qui s'ajoutent à la structure minimale (commune) K ne décrivent déjà plus un objet *absolument* quelconque.

ou accidentelles. Quant aux vérités logiques, elles sont supposées vraies en toute situation, donc pour toute classe de mondes que l'on désigne par des modèles. Etant donné qu'un modèle $\mathcal{M}$ de logique propositionnelle se définit comme une structure $\mathcal{F}$ munie d'une fonction d'interprétation v, on peut dire qu'une vérité logique est une vérité relative à une structure de langage, uniquement. Nous avons vu que Quine ne distinguait pas entre plusieurs structures, puisque son universalisme l'amenait à considérer son modèle de discours comme *le* modèle, c'est-à-dire l'ensemble des objets et faits constitutifs du monde. Mais lorsque l'on adopte l'approche de la logique comme calcul, la définition de la validité ou vérité logique se complique :

> Fitting distingue trois notions de conséquence logique : la conséquence *locale* définie en termes de vérité relative à un monde, la conséquence *globale* définie en termes de validité dans un *modèle*, et la conséquence *globale* définie en termes de validité dans une *structure*.[135]

La critique d'une logique modale peut être double. D'un point de vue formel, tout d'abord, si l'on considère la logique comme l'étude des vérités logiques alors la notion de validité, relative à des classes de structures, est déjà trop restrictive pour elle ; on trouve une semblable position chez des auteurs tels que Granger et Gonseth, selon lesquels la logique s'apparente à une physique de l'objet quelconque. Dans ce cas, une formule valide dans toute structure de Kripke n'est pas une vérité logique et concerne une théorie formelle de type algébrique, mais pas la logique. D'un point de vue philosophique, d'autre part, on peut estimer qu'aucune structure formelle ne serait en mesure de donner une définition satisfaisante d'un concept ordinaire tel que le passé, l'obligation, la croyance, etc. La première critique, conforme à l'idée de logique comme physique de l'objet *absolument* quelconque, se retrouve dans [Hocutt 1972] :

> Il va de soi que, pour un logicien modal, ce qui fait d'une logique modale qu'elle est *logique* ce sont ses caractéristiques purement formelles ; ce qui fait qu'elle est *modale* (i.e. aléthique, épistémique, doxastique, déontique, ou autre), c'est l'interprétation particulière assignée à l'opérateur modal 'non interprété'. Le résultat est que les logiques modales sont des *formalismes interprétés*, notion que je trouve paradoxale. Ce paradoxe n'est pas signalé en règle générale. Ni le fait que la capacité de variation d'interprétation (en termes de notions diverses telles que nécessité, savoir, croyance, obligation ou autre) montre très bien que l'opérateur modal n'est pas une constante logique de *n'importe quel* type.[136]

La seconde critique, majoritaire dans la littérature sur la logique épistémique durant les années 1960-70, porte sur l'applicabilité des logiques modales sur les langues naturelles. Même un logicien modal tel que E. J. Lemmon ne croit pas pour cette raison précise au projet de logique épistémique :

---

[135] [Gochet & Gribomont & Thayse 2000], pp. 133-4.

[136] [Hocutt 1972], p. 438, note 1. L'objection repose sur une absence d'*évidence* dans l'application des lois logiques modales, par opposition aux lois logiques classiques : "dès le début, la logique modale a montré une *incertitude* concernant la validité de ses principes logiques qui est totalement étrangère à la logique classique frégéenne", in L.T.F. Gamut : *Intensional Logic and Logical Grammar*, The University of Chicago Press (Chicago and London) (1991), p. 20 ; nous soulignons. Nous reviendrons sur cet argument contestable de l'évidence, contesté plus loin dans le 'modèle explicatif' de Hintikka.

> une logique *réaliste* de la connaissance ne contient pas de théorème distinct en dehors de [Kϕ → ϕ] et de ses conséquences logiques.[137]

Pour résumer les deux objections précédentes, la distinction établie précédemment entre langage universel et calcul se répercute sur la notion de *validité* en logique : la vérité logique des universalistes est la vérité dans un modèle unique, donc sans structure, et qui comprend l'ensemble des tautologies ; la vérité logique des relativistes est l'ensemble des tautologies plus l'ensemble des formules valides dans une ou plusieurs structures. Les partisans de la logique comme langage universel ont un seul et même critère pour la vérité logique et la validité : être vrai dans tout modèle, précisant bien que leur modèle est unique et propose un seul et unique univers de discours pour l'interprétation des formules.[138] Par opposition, les partisans du relativisme (opposés au critère (!∀G) de l'universalisme logique) distinguent la vérité logique de la validité, ainsi que plusieurs ensembles de formules valides : valides en logique classique propositionnelle, puis valides en logique modale (dans **K**, **T**, **S4**, **S5**, etc.). Pour ce qui nous intéresse ici, la logique modale conçoit la vérité logique relativement à des structures de langage : les systèmes de Lewis, définis par des propriétés distinctives de R dans un ensemble W de mondes w. La validité modale est donc indifférente aux objets du domaine de discours (elle respecte à cet égard le critère (∀G) des universalistes), mais sensible aux relations R entre les mondes d'un modèle : les formules **T**-valides sont celles pour lesquelles R satisfait la relation de réflexivité, les formules **S4**-valides celles pour lesquelles R satisfait la relation de transitivité, etc. Un symptôme de l'universalisme logique semble se manifester dans le rejet par certains logiciens de l'idée de validité *relative* à un modèle et sa classe de structures. De là leur distinction entre langage formel et langage logique, dont les vérités caractéristiques ne reposent pas seulement sur leur abstraction ou indifférence par rapport au contenu de discours (sur le critère (∀G), encore une fois). Peut-on soutenir que certaines vérités logiques s'imposent indifféremment à toute structure de langage particulière ? Etant donné que les structures de Kripke sont décrites comme des 'superstructures modales' surajoutées à l'infrastructure de la logique classique, celle-ci constitue-t-elle une structure ultime, la seule structure du langage logique, ou n'est-elle qu'une structure parmi d'autres ? Le commentaire qui suit de [Kielkopf 1971] semble supposer que les tautologies classiques sont des formes de raisonnement *absolument* nécessaires –valides dans toute structure– et méritent elles seules le titre de vérité *logique* (et pas simplement *formelle*) :

> les thèses ne sont pas nécessairement vraies parce qu'elles permettent de valider une inférence ou de donner des formes de prémisses qui peuvent être utilisées dans n'importe quel argument. Un présupposé fondamental de cet essai est que les tautologies de la logique classique propositionnelle sont des paradigmes de ce que devraient être les thèses d'une logique. Il est alors clair que (...) le système des thèses doxastiques de Hintikka ne peut pas être une logique de la croyance.[139]

---

[137] [Lemmon 1967], p. 78.

[138] L'expression 'dans *tout* modèle' peut induire en erreur, parce qu'elle laisse penser qu'il y a plusieurs modèles distincts en logique classique ; en effet, on ne parle pas de totalité en règle générale sans présupposer l'existence de parties et, donc, de relations structurées entre ces parties. Or pour le logicien classique, il n'y a pas de structure dans son modèle qui ne contient aucune partie et comprend une totalité d'objets indistincte, un seul et unique domaine de référence. Pour cette raison, 'être vrai dans tout modèle' signifie 'pour tout type d'objets' et ne signifie pas la même chose que la validité modale : 'être vrai dans toute structure de Kripke'.

[139] [Kielkopf 1971], pp. 3-4.

Le critère hypothético-déductif apparaît donc comme insuffisant à l'auteur cité pour fonder une logique. Peut-on admettre que les vérités logiques soient des vérités formelles singulières parce que vraies 'par nature', en vertu de règles de signification que l'on admettrait comme inhérentes aux mots logiques ? Même Quine n'admettrait pas une distinction *catégorielle* de ce genre entre les constantes classiques et enrichies, parmi lesquelles les opérateurs modaux.

Pour justifier cette position radicale contre la logique modale et sa notion relative de vérité dans une structure, le critère *formel* est insuffisant : nous avons signalé plus tôt que, d'après le relativisme radical de [Tarski 1986], n'importe quel terme du langage naturel est susceptible d'être adapté dans un langage formel de façon à se comporter comme un mot logique. En fait, l'argument de Kielkopf ci-dessus repose plutôt sur un critère *informel* de logicité : contrairement aux tautologies classiques, les théorèmes modaux ne seraient pas de nature logique parce qu'ils reposeraient sur des opérateurs dont l'interprétation n'est pas conforme à l'*usage*. La 'vérité dans l'usage' des énoncés servirait de contrepartie à la 'validité' des formules : puisque aucune règle de sémantique formelle ne suffit pour justifier à elle seule la vérité d'un énoncé des langues naturelles, le rôle du philosophe serait de justifier le choix d'une certaine structure modale en montrant que les formules valides qu'elle contient symbolise un ensemble d'énoncés vérifiés dans l'usage. Pour donner un exemple, le philosophe considérerait le système **T** comme caractéristique du concept de connaissance dans la mesure où le verbe 'savoir' présente une *grammaire de succès* : toute connaissance est considérée comme vraie, donc le schéma ($K\phi \rightarrow \phi$) est une formalisation adéquate de ce verbe.[140] Peut-on juger la pertinence d'une logique philosophique à partir de cette méthode, laquelle suppose une invariance d'usage pour les concepts des langues naturelles susceptibles de servir de mots logiques dans un langage formel ?[141] Faut-il entendre par là que les mots naturels tels que 'et', 'ou', 'si ... alors', 'ne ... pas' seraient correctement retranscrits dans la logique classique, contrairement à d'autres syncatégorèmes ou termes formateurs tels que 'nécessairement', 'peut-être', 'obligatoirement', 'savoir', etc. ? Le caractère purement arbitraire de la formalisation logique selon Tarski n'est pas maintenu ici, puisque l'on suppose qu'à chaque terme du langage ordinaire reviendrait une fonction déterminée (lexique ou mot

[140] Nous présentons le critère de l'usage comme un moyen de répondre à la question posée dans [Lemmon 1959] : "Mais avec quels *faits* devons-nous faire coïncider les affirmations que nous déduisons des formules logiques –les *faits logiques* ? Peut-être est-il sans risque de dire que c'est un fait logique que soit il est train de pleuvoir maintenant soit il n'est pas en train de pleuvoir maintenant. Mais il reste la question : Comment savons-nous que cela est vrai ? Il ne peut y avoir d'appel en retour vers le système logique, à ce moment ; nous ne pouvons pas dire que nous le savons parce que cela illustre un théorème du calcul propositionnel ; car c'est ce calcul lui-même que nous mettons ici à l'épreuve de la correction." (p. 28 ; nous soulignons). Un tel critère reste précaire, puisque la multiplicité des usages complique notre souci d'adéquation informel pour un système de logique philosophique.

[141] Pour juger l'adéquation d'un système de logique interprété, une logique philosophique, [Rescher 1960] insiste sur "la nécessité de développer, préalablement à l'énonciation de quelque théorie que ce soit, un *critère d'acceptabilité* adéquat ou *critère d'adéquation* pour les règles d'une théorie des affirmations de croyance, critère qui puisse servir à la fois de test en vue de la sélection des règles proposées et de guide en vue de la construction de la théorie. En l'absence d'un tel critère *directeur* ... une confiance aveugle au concept d'*acceptabilité intuitive* –lequel constitue notre seul guide en l'absence d'un critère explicite – nous fourvoiera dans la construction d'une théorie des affirmations de croyance". (pp. 91-92). L'absence de tout critère d'interprétation de ce genre, affirmée par [Fisher 1964] et confirmée plus tard par [Hintikka 1968], renforce le rôle de l'intuition pré-théorique dans la construction des systèmes modaux appliqués.

logique, fonction ou argument). Nous arrivons par cet argument de la *signification dans l'usage* à l'objection principale portée contre la logique épistémique : une telle logique philosophique serait vouée à l'échec parce que les verbes d'attitude propositionnelles ont un usage variable et ne peuvent fonctionner comme des mots logiques.

2.2 La logique des concepts et leur usage

Etre valide dans un langage formel, ce n'est pas forcément être valable dans la langue naturelle : on peut démontrer n'importe quoi à l'intérieur d'un langage formel dès lors que des règles de transformation valide des formules y sont stipulées. Si par exemple, dans ce langage artificiel, une formule conjonctive $A \wedge B$ est définie sémantiquement comme ce dont la vérité exige que A et B soient fausses toutes les deux, nous ne dirions pas du signe $\wedge$ qu'il désigne une conjonction mais que, au contraire, il représente dans *ce* système le connecteur du rejet.[142] Malgré cet argument quelque peu forcé, on suppose que chaque signe d'un langage formel revêt un sens particulier ($\wedge$ signifie la négation, par exemple) et que, contrairement au choix du signe, le sens des mots logiques est plus ou moins donné sur la base d'une compétence linguistique préalable ($A \wedge B$ signifie que A et B sont vrais tous les deux). Autrement dit : à chaque opérateur logique serait associable une expression particulière de la langue naturelle et dont le sens serait plus ou moins clairement saisi. Peut-on en conclure qu'il existe une relation d'adéquation de ce genre entre linguistique empirique et logique formelle, c'est-à-dire entre les termes formateurs de la grammaire naturelle et les opérateurs formels ? C'est en tout cas sur l'inadéquation prétendue entre 'savoir' et 'K' puis entre 'croire' et 'B' que repose une grande part des critiques au système de [Hintikka 1962], pour la raison que ses théorèmes épistémiques ne seraient admissibles que par le biais d'une interprétation 'irréaliste' des opérateurs de croyance et de savoir. Pour retracer quelques arguments majeurs contre cette logique modale épistémique, nous retiendrons trois objections récurrentes : la logique épistémique est irréaliste ; ce n'est pas une logique, mais une éthique ; elle est une logique classique déguisée. Derrière ces trois objections, c'est le statut épistémologique de la logique en général qui est en cause.

2.2.1 Une logique "irréaliste"

Suite à l'hypothèse précédente d'une corrélation entre la vérité logique d'une formule et l'invariance d'usage de ses termes logiques, les théorèmes de [Hintikka 1962] sont accusés d'attribuer aux opérateurs de croyance ou de connaissance des usages qui ne leur conviennent pas : les verbes psychologiques auraient une signification trop fluctuante pour prétendre au titre de mots logiques. C'est l'argument que l'on retrouve en substance chez de nombreux commentateurs de [Hintikka 1962].

---

[142] D'après Quine, le sens des foncteurs de vérité est invariable. Changer le sens des connecteurs, c'est parler d'une autre référence, en vertu de sa doctrine du changement de sujet ; voir [Quine 1975], Chapitre 6 : d'après lui, changer les conditions de vérité d'un connecteur "est une hypothèse absurde. Il ne reste pas d'essence résiduelle [du connecteur] quand on a ôté les sons, les notations, et les lois en conformité desquelles on emploie ces sons et ces notations". (p. 120). Contre cette théorie de la 'désignation rigide' des connecteurs logiques, voir toutefois [Paoli 2003], selon lequel l'argument ci-dessus de Quine contre le pluralisme global reposerait sur une caractérisation incomplète des connecteurs logiques.

Le problème apparaît déjà dans la structure minimale **K**, commune à tous les modèles de Kripke. En vertu des règles d'inférence du système axiomatique **K**, toutes les thèses α du calcul des propositions classique CPC sont 'nécessitées', ce qui signifie sémantiquement que les tautologiques classiques sont préservées dans *tous* les mondes possibles : $\models\phi \Rightarrow \models_K\Box\phi$; et puisque les tautologies sont également valides dans n'importe quelle structure, la loi de distribution $\Box(\phi \rightarrow \psi) \Rightarrow \Box\phi \rightarrow \Box\psi$ est valide également. De telles règles sont considérées comme 'irréalistes' parce qu'elles font de l'agent épistémique un être *logiquement omniscient* : il connaît chacune des conséquences logiques de ce qu'il sait, et de même pour l'opérateur de croyance.

La règle de distribution est défendable dans un certain sens : si *a dit* qu'il sait que β résulte de α, alors on ne peut guère concevoir qu'il *dise* ensuite qu'il sait que α est le cas sans être disposé à accepter β par la même occasion ;[143] d'autant plus que, dans [Hintikka 1962], une des conditions préalables à l'analyse des attitudes épistémiques est que chaque *déclaration* de croyance ou de connaissance soit produite à une seule et même occasion, c'est-à-dire un intervalle au cours duquel l'agent ne change pas d'avis.[144] Pour cette raison, il faut supposer que, dans $K(\phi \rightarrow \psi)$, la conséquence $K\psi$ est tirée au même moment que l'énonciation de $K\phi$. L'unité de temps justifie ainsi cette interprétation charitable de la règle de distribution.

En revanche, la règle de nécessitation est bien plus contestable puisqu'elle suppose que l'agent sait ou croit toutes les tautologies du calcul propositionnel, y compris celles qu'il n'a pas affirmées explicitement : l'espace épistémique doit être conforme à l'espace logique, en vertu du cadre sémantique des structures de Kripke, et cette particularité formelle fait de l'agent un être capable de connaître le nombre infini des tautologies classiques. Une telle exigence peut également être défendue charitablement, si l'on définit la connaissance comme une simple *disposition* à prendre connaissance de quelque chose. Cette disposition présuppose toutefois une compétence logique minimale pour chaque agent, ce que les objecteurs ont considéré comme une 'idéalisation' outrancière des capacités cognitives de l'agent ; aucun agent 'réel' de la vie courante n'a un comportement logique impeccable, et l'existence d'actions incohérentes ou irrationnelles semble montrer qu'une caractérisation dispositionnelle des opérateurs épistémiques est déjà trop exigeante. En réponse à l'accusation courante lancée contre la logique modale épistémique :

> Hintikka a formalisé des sens très forts de 'savoir' et 'croire', pour lesquels il n'y a probablement pas d'instances humaines[145]

---

[143] Le lien entre l'affirmation explicite d'une relation logique et la connaissance de cette relation est formulée par Hintikka en termes de *conscience* : "En général, notre examen n'est applicable à ce que les gens savent réellement (ou activement, comme nous le dirons) que s'ils ont conscience [*aware*] des conséquences de ce qu'ils savent" (p. 34). Le savoir affirmé est explicite, par rapport à un savoir implicite de type dispositionnel. L'opposition entre les deux approches crée une ambiguïté quant à l'interprétation des théorèmes épistémiques.

[144] "Les déclarations [*statements*] en question doivent être faites à *une seule et même occasion*. L'importance de cette notion d'occasion n'a pas toujours été appréciée comme il se doit. Néanmoins, il est clair qu'un critère de consistance logique n'est applicable à un certain nombre d'affirmations que lorsqu'elles sont totalement comparables ; et cela présuppose, entre autres, que la distance temporelle entre elles (y compris ce qui se passe dans l'intervalle) est hors de propos. (Elles doivent être faites au sein d'un 'présent logique spécieux')" ([Hintikka 1962], p. 7).

[145] [Castañeda 1964], p. 133.

l'accusé a précisé que les théorèmes modaux ne seraient vérifiables que par des agents rationnels parfaits, c'est-à-dire tels qu'ils "suivent suffisamment loin les conséquences de ce qu'ils savent ou croient dans chacun des mondes connaissables". Dans [Hintikka 1962], une redéfinition est proposée en réponse pour les notions métalogiques de validité, de contradiction et de simple vérité dans un modèle (satisfiabilité) : lorsqu'une formule épistémique est valide pour un agent parfait, c'est-à-dire vraie dans tous les mondes que cet agent considère comme épistémiquement possibles, cette formule n'est pas 'activement' (réellement) mais 'virtuellement valide' et l'auteur la qualifie d''auto-confirmante' [*self-sustaining*] ; lorsqu'une formule épistémique est contradictoire pour un agent idéal, c'est-à-dire toujours fausse, elle est 'virtuellement contradictoire' et l'auteur la qualifie d''indéfendable'. Quant à l'ensemble des formules susceptibles d'être tantôt vraies tantôt fausses pour cet agent parfait, elles sont dits simplement 'défendables'. Ce néologisme a pour but de distinguer la vérité logique d'une formule pour un agent normal et sa garantie épistémique pour des agents idéaux : une formule logiquement vraie est vraie en vertu des significations des connecteurs logiques, tandis qu'une formule est valable ('vraie dans l'usage') en vertu de la signification des concepts correspondants. L'objection centrale des commentateurs de la logique modale épistémique consiste à maintenir une distinction entre ce qui est logiquement vrai et ce qui est démontrable : être logiquement vrai en vertu de la signification des connecteurs logiques est une chose ; être toujours vrai dans un modèle en vertu d'un choix arbitraire du domaine de référence ou des opérations des connecteurs en est une autre. Or puisque les théorèmes de [Hintikka 1962] ne sont vérifiés que par des agents idéaux, le système est accusé d'être par conséquent 'irréaliste' parce que validé dans un modèle arbitrairement construit. L'accusation porte ainsi sur le fossé qui sépare le langage formel des logiques philosophiques et les langues 'historiquement données', auxquelles elles sont pourtant censées s'appliquer.

[Linsky 1968] relève un autre défaut impliqué par la sémantique relationnelle : non seulement l'agent est logiquement omniscient (et omnicroyant), mais il est tenu d'attribuer la même omniscience (omnicroyance) à tous les autres agents.[146] La raison est d'ordre technique, puisque la condition de non-contradiction imposée aux mondes épistémiques signifie que toute formule dans la portée d'un opérateur épistémique doit être consistante ; et puisque tout ce qui se situe dans la portée de K doit être consistant, la valeur sémantique de $K_aK_b\phi$ dépend de la valeur sémantique de ϕ. Hintikka appelle 'transmissibilité du savoir' cette propriété de K : toute attribution de savoir à un autre implique que l'on sait soi-même, même si l'on ne dispose d'aucune preuve en faveur de l'objet de connaissance en question.[147]

[146] [Linsky 1968] s'intéresse en particulier à la logique doxastique, mais son exemple s'applique à n'importe quel opérateur fort ($\Box$ = B ou $\Box$ = K) puisqu'il ne dépend pas d'une propriété caractéristique de l'un ou l'autre : "La *fausseté* manifeste 'Personne ne croit que quelqu'un croit une contradiction logique' est prouvable également par l'utilisation du système doxastique de Hintikka". ([Linsky 1968], p. 502). Soit ϕ une formule contradictoire: le système modèle de [Hintikka 1962] est tel que $\Box_a\Box_b\phi$, bien que possible dans les faits, est indéfendable pour tout agent épistémique rationnel *a* et *b*.

[147] "Il se peut que cet exemple [$K_aK_bp \rightarrow K_ap$] ne soit pas totalement dénué d'intérêt philosophique. Ce qu'il montre, en effet, c'est que le savoir est transmissible. Si je sais que vous savez que *p* est vrai, je sais moi-même virtuellement que *p* est vrai. En conséquence, si vous me dites que vous savez que *p* est vrai, je ne peux pas nier (de manière défendable) savoir moi-même si *p* est vrai sans indiquer le fait que je ne vous fais pas totalement confiance (à votre jugement ou votre sincérité). Ceci sert à expliquer pourquoi, et dans quel sens, 'là où quelqu'un m'a dit "je sais", je suis habilité à dire que *je* sais aussi, en seconde main'" ([Hintikka 1962], p. 61).

Pourquoi cette transmission, sinon parce que les mondes épistémiques sont définis au sein d'une logique modale *normale* ? Cette conséquence contre-intuitive résulte de la définition duale des opérateurs modaux ($K\phi \rightarrow \neg P\neg\phi$), qui revient à attribuer des propriétés formelles normales des opérateurs modaux aux propriétés naturelles irrégulières (donc non-normales) des attitudes épistémiques. A l'image de la distinction métalogique entre validité et auto-confirmation, ou entre contradiction et indéfendabilité, les résultats de [Hintikka 1962] produisent des conclusions apparemment *ad hoc* sur le comportement cognitif des agents humains ; or

> ce que l'on veut n'est pas une *élaboration formelle* mais une *explication informelle*. (...) A titre de défense de la logique épistémique celle-ci est circulaire, bien sûr. Elle garantit simplement les théorèmes de la logique épistémique par un *fiat*: ce sont les vérités concernant le savoir au sens de 'savoir', quel qu'il soit, pour lequel ce sont des vérités au sujet du savoir.[148]

C'est pourquoi la logique modale épistémique a été très tôt discréditée ou considérée comme une raffinerie technique sans intérêt, parce que sans rapport avec ce que les agent savent ou croient *réellement*.[149]

Cela étant, les notions de 'réalité' des agents ou de 'vérité' des formules restent vagues. Nous avons rappelé plus tôt que la valeur sémantique des énoncés des langues naturelles n'est déterminable qu'après les avoir été traduits dans un langage formel et évalués par le biais de règles sémantiques. Faut-il en conclure qu'un énoncé ne peut pas être bien compris sans recourir à un langage formel qui établit ses règles de signification ? Si tel est le cas, alors le seul moyen de construire un langage formel qui ne soit pas arbitraire consiste à faire appel au critère de signification dans l'usage ou d'intuition langagière.[150] Pour marquer la différence entre ce qui est vrai pour des agents réels et ce qui est démontrable pour des agents idéaux (déductivement parfaits), [Kielkopf 1971] propose de distinguer entre les notions de *logique* et de *système* :

---

[148] [Hocutt 1972], p. 438.

[149] Hintikka est conscient que ses règles sémantiques ne sont satisfaites que par un prototype idéal d'agent rationnel, comparable à l'*homo oeconomicus* des théories économétriques : "Nos résultats ne sont pas applicables directement à ce qui est vrai ou faux dans le monde réel à nous. Ils ne nous disent quelque chose de déterminé au sujet de la vérité et de la fausseté des affirmations que dans un monde dans lequel tout le monde suit les conséquences de ce qu'il sait aussi loin qu'elles le conduisent. Un énoncé est auto-confirmant [*self-sustaining*] s'il est vrai dans tous les mondes de ce genre, il est défendable s'il est vrai dans un monde de ce genre au moins, et ainsi de suite. Ils ne sont applicables à des déclarations réelles que dans la mesure où notre monde réel se rapproche de l'un des 'mondes possibles les plus connaissables', ainsi que nous pouvons les appeler, ou que l'on peut le faire rapprocher de l'un d'eux en attirant l'attention des gens sur les conséquences de ce qu'ils savent" ([Hintikka 1962], p. 36). Si l'on poursuit l'analogie avec l'économétrie, la préférence actuelle pour les modèles micro-économiques réalistes devrait nous inciter à suivre la philosophie du langage ordinaire et à abandonner la philosophie formelle, trop schématique. C'est pourtant cette voie qu'il s'agira de maintenir et justifier.

[150] [Eberle 1974] insiste sur le rôle de l'intuition extra-logique dans la construction d'un système logique adéquat : "mis à part les doutes quant à savoir si cette réinterprétation correspond réellement à de claires intuitions (...) l'objection demeure que les systèmes ne parviennent pas à traiter les notions coutumières de croyance et de savoir." (p. 356 ; nous soulignons).

> On sait bien, et Hintikka s'en est lui-même rendu compte, que des formules *falsifiables* sont *prouvables*. Il serait donc erroné de qualifier son système de *logique* doxastique. Dans ce qui suit, je le qualifierai simplement de *système* doxastique.[151]

Compte tenu de cette distinction, un langage formel est *logique* si et seulement si les interprétations de ses vérités ne sont pas falsifiées dans l'usage : si ($K\phi \rightarrow \phi$) est intuitivement vraie, c'est-à-dire toujours vérifiée dans l'usage en plus d'être démontrable dans **T**, c'est parce qu'elle exprime une implication conforme à la signification dans l'usage des deux termes de vérité et de savoir.[152] La distinction utilisée ici reprend l'hypothèse de travail mentionnée plus haut, à savoir l'idée de corrélation entre la valeur sémantique d'une formule et la signification dans l'usage de la phrase correspondante. Dans une logique philosophique, le cadre sémantique utilisé ne doit pas être un simple artifice formel : les formules valides dans un système formel *et logique* doivent être confirmées dans leurs applications aux énoncés des langues naturelles, faute de quoi les systèmes de logique philosophique ne seraient que des langages sans usage, donc sans portée philosophique. Telle est en tout cas l'objection qui revient en permanence contre [Hintikka 1962] ; elle suppose une notion de vérité 'dans les faits' ou 'dans l'usage', à l'appui des constructions formelles qui parlent de vérité ou de fausseté à l'intérieur d'un langage simplement formel.

La distinction entre un agent idéal et un agent réel peut s'exprimer autrement : elle correspond à ce que les sujets pensent réellement ou *devraient* penser dans des conditions idéales. Une telle distinction apparaît déjà dans l'opposition entre l'approche dite *descriptive* et *normative* de la logique : un système de logique exprime-t-il ce que le sujet pense réellement, ou ce qu'il devrait penser pour raisonner correctement ? Les verbes 'croire' et 'savoir' expriment des attitudes du sujet face aux relations logiques entre des ensembles d'énoncés : pour toute une variété d'attitudes propositionnelles exprimées par une forme de pensée (volonté, crainte, croyance, etc.), le sujet pense que ϕ est valide si ϕ est valide ; s'il pense que ϕ et si ψ est une conséquence de ϕ, alors il pense que ψ, etc. Or de même que Quine reprochait aux partisans de l'implication stricte d'introduire l'idée d'inférence dans le langage-objet, ne peut-on pas reprocher également aux logiciens épistémiques de confondre pensée pure et pensée en acte, lorsqu'ils font *mention* de certaines règles d'inférence logique (clôture déductive, distribution) pour décrire leur *usage* par un sujet (sa capacité à appliquer ces inférences) ? Les deux objections qui suivent s'appuient sur cette critique d'ordre méthodologique : les inférences épistémiques ne feraient rien d'autre que mentionner des inférences classiques et, de ce fait, elles ne nous apprendraient rien de nouveau sur les attitudes épistémiques en tant que telles.

---

[151] [Kielkopf 1971], p. 135. [Hintikka 1962] lui-même reconnaît que "les règles ($A.\neg K$) et ($A.\neg P$) ne sont absolument pas concernées par la *vérité* des déclarations ; elles nous disent simplement que certaines adjonctions préservent toujours la défendabilité d'ensembles d'énoncés" (p. 32).

[152] La relation implicative entre savoir et vérité est admise presque unanimement par les logiciens philosophes, dont on aperçoit une version contraposée ($\neg\phi \rightarrow \neg K\phi$) dans le *Théétète* de Platon : "Socrate – Est-il possible de parvenir à la vérité lorsque l'on ne peut parvenir à ce qui est ? Théétète – C'est impossible. Socrate – Mais si un homme ne peut pas parvenir à la vérité d'une chose, se peut-il qu'il connaisse cette chose ? Théétète – Non, Socrate, comme le pourrait-il ? (186c-d)". Pour quelques arguments critiques contre cette 'clause de succès' : $K\phi \rightarrow \phi$, voir toutefois [Lenzen 1978], pp. 7-11, ainsi que la section 2.2.4.3.

2.2.2 Logique épistémique = une éthique de la logique classique ?

La distinction entre les deux types de pensée ci-dessus renvoie également à deux conceptions de la logique en général : la logique comme *physique* (descriptive), la logique comme *éthique* (normative).[153] [Hintikka 1962] précise que sa caractérisation des opérateurs modaux illustre une approche *prescriptive* des compétences logiques de l'agent :

> Le fait que lesdites lois de la logique ne soient pas des 'lois de la pensée' au sens de lois naturelles semble être admis de nos jours, de façon générale. Mais les lois de la logique ne sont pas non plus des lois de la pensée au sens de commandements, sauf peut-être des lois de la pensée la plus pointue possible. Compte tenu d'un certain nombre de prémisses, la logique ne nous dit pas quelles sont les conclusions que nous *devons* en tirer ; elle nous dit simplement quelles conclusions nous *pouvons* en tirer –si nous le souhaitons et si nous sommes suffisamment intelligents.[154]

L'agent qui sait que ϕ ne doit pas mais est susceptible de savoir que ψ sur la base de l'inférence ($\phi \Rightarrow \psi$), et cette capacité toute relative implique que les théorèmes épistémiques présupposent des conditions de rationalité minimales chez *a*.

Pour résumer les objections portées contre la logique modale épistémique, les auteurs ci-dessous prétendent que le système est confronté à un dilemme insoluble : soit les inférences sont logiquement vraies, auquel cas elles ne caractérisent pas les propriétés de la croyance ou du savoir ; soit elles caractérisent ces propriétés, auquel cas elles ne sont pas logiquement vraies.[155] Commençons par l'argument selon lequel les inférences épistémiques ne sont pas logiquement *vraies*, avant de poursuivre par celui selon lequel les inférences logiques ne sont pas *épistémiques*.

Pour montrer que la logique modale épistémique n'est pas une logique proprement dite, [Kielkopf 1971] propose de considérer comme logique un théorème prouvable dans *n'importe quel* système formel, indépendamment de sa structure sémantique. Peut-on trouver des formules susceptibles d'être vérifiées dans n'importe quel système formel ? Sans parler de la distinction entre vérité logique et validité dans une structure, on pourrait objecter à Kielkopf qu'aucune formule n'est logiquement vraie sans présupposer un certain nombre de croyances fondamentales, aussi peu nombreuses soient-elles : le choix de la bivalence, par exemple, n'est pas neutre et entraîne la validation ou non du tiers exclu, de la réduction par l'absurde, de la double négation, etc. Par ailleurs, on peut considérer que le choix d'un ensemble de vérités logiques dépend de l'utilité d'un système formel en fonction d'un besoin prédéfini ; en dehors de cette motivation instrumentale des systèmes de logique, il n'y aurait point de théorème fondamental. Peut-on supposer au contraire l'existence d'un socle minimal de

---

[153] Le sens prescriptif de la logique comme 'éthique de la pensée', décrivant la façon dont nous *devrions* penser pour penser correctement, est attribuée à Herbart. Voir [Engel 1996], p. 131.

[154] [Hintikka 1962], p. 37 ; nous soulignons.

[155] Lorsqu'il s'agit de montrer l'adéquation informelle des théorèmes épistémiques, le dilemme vient de ce que "ces tentatives soit (1) nient les contre-exemples en vidant effectivement le terme de 'savoir' de sa substance, soit (2) nient que les théorèmes épistémiques doivent être *vrais*. Ils nous laissent donc avec l'une ou l'autre de ces deux questions très pressantes: (1) Dans quel sens la logique épistémique est-elle *épistémique* ? ou (2) Dans quel sens est-elle *logique* ?" ([Hocutt 1972], p. 436). Toute adéquation informelle entre la signification d'usage des verbes 'savoir' et 'croire' et leurs théorèmes est donc impossible, d'après l'auteur cité.

vérités logiques capables de satisfaire tous les logiciens ? L'auteur ne s'étale pas sur la question. Pour illustrer son argumentation, il propose une métathéorie de l'opérateur de conséquence logique : quelles sont les propriétés qu'un opérateur doit posséder pour être un opérateur logique, et pas simplement formel ? Cet opérateur est susceptible de définir les propriétés fondamentales d'une relation de conséquence logique et, par conséquent, d'apporter une réponse au problème ci-dessus du socle fondamental de vérités logiques. Partant d'une distinction entre les caractéristiques des modalités aléthiques L (ce qui est nécessairement vrai) et déontiques O (ce qui devrait être vrai), Kielkopf rappelle tout d'abord que la première implique la simple vérité : $L\phi \rightarrow \phi$ est logiquement vraie, mais pas la seconde : $O\phi \rightarrow \phi$ n'est pas logiquement vraie.[156] Puis il reporte les propriétés de ces deux opérateurs au niveau du métalangage afin de caractériser un méta-opérateur modal de *prouvabilité* $\vdash$: pour tout opérateur modal interprété, Kielkopf qualifie cet opérateur de *logique* uniquement s'il a un comportement métalogique de type *aléthique*, autrement dit, s'il satisfait la propriété de prouver ce qui est vrai.[157] Démontrabilité (ou prouvabilité) et vérité sont deux propriétés distinguées, ici, et c'est pourquoi la prouvabilité d'une formule dans un système S n'est pas considérée comme une condition suffisante pour faire de S un système *logique* : la vérité n'est pas une question strictement *interne* à un langage formel et des règles sémantiques arbitrairement choisies ; elle demeure une question externe et justifie dans ce cas la distinction entre validité dans une structure et vérité logique. Si donc un opérateur modal ne satisfait pas la propriété de réflexivité dans le métalangage, c'est qu'il n'est pas de nature logique mais *éthique* : il prouve ce qui devrait être vrai mais ne l'est pas, tandis qu'un système dit logique serait

> une *logique réelle* si son opérateur de prouvabilité est aléthique,[158]

sous-entendu qu'une formule valide doit être 'réellement' vraie si elle est logiquement vraie : ce qui est vrai dans un système doit être vrai *simpliciter*, c'est-à-dire dans n'importe quel système.

Or en distinguant vérité logique et prouvabilité (ou vérité formelle), l'auteur suppose qu'une vérité logique s'impose indépendamment de tout système dans lequel elle puisse être

---

[156] "(...) Je peux dire qu'un opérateur de nécessité est *normal* dans le système S si et seulement si ce qui suit est valable. (a1) $(Lp \rightarrow p)$ est prouvable dans S. (a2) $(L(p \wedge q) \equiv L(p) \wedge L(q))$ est prouvable dans S. (a3) $L(p \rightarrow p)$ est prouvable dans S. (a4) $(L(p) \equiv \neg M\neg p)$ est prouvable dans S. (a5) $(p \rightarrow Lp)$ n'est *pas* prouvable dans S. Un opérateur déontique (*ought-operator*) est normal dans un système S si et seulement si ce qui suit est valable. (d1) $(Op \rightarrow Pp)$ est prouvable dans S. (d2) $(O(p \wedge q)) \equiv (Op \wedge Oq)$ est prouvable dans S. (d3) $(Op \equiv \neg P\neg p)$ est prouvable dans S. (d4) $(Op \rightarrow p)$ n'est *pas* prouvable dans S. (d5) $(p \rightarrow Op)$ n'est *pas* prouvable dans S." (p. 5).

[157] "(cpc1) $(\vdash_{cpc}F \rightarrow F)$ est prouvable ; (cpc2) $(\vdash_{cpc}(F \wedge G) \equiv \vdash_{cpc}F \wedge \vdash_{pc}G)$ est prouvable ; (cpc3) $(\vdash_{cpc}F \rightarrow F)$ est prouvable ; (cpc4) $(\vdash_{cpc}F) \equiv \neg\neg(\vdash_{cpc}\neg\neg F)$ est prouvable ; (cpc5) $(F \rightarrow \vdash_{cpc}F)$ n'est *pas* prouvable. De façon similaire, si nous posons maintenant S comme le métalangage sémantique du système *doxastique* de Hintikka, considérons non-$\vdash_{dox}\neg F$ comme P(F), et symbolisons-le par: $\neg\vdash_{dox}\neg F$, montrer l'analogie entre $\vdash_{dox}F$ et un opérateur déontique revient à montrer ce qui suit. (dox1) $(\vdash_{dox}F \rightarrow \neg\vdash_{dox}\neg F)$ est prouvable. (dox2) $(\vdash_{dox}(F \wedge G) \leftrightarrow \vdash_{dox}F \wedge \vdash_{dox}G)$ est prouvable. (dox3) $(\vdash_{dox} F \rightarrow F)$ *n'est pas* prouvable. (dox4) $(\vdash_{dox}F) \equiv \neg\neg(\vdash_{dox}\neg\neg F)$ est prouvable. (dox5) $(F \rightarrow \vdash_{dox}F)$ *n'est pas* prouvable (pp. 5-6).

[158] [Kielkopf 1971], p. 11 (nous soulignons).

prouvée : peut-on parler de la vérité logique d'une formule hors de tout système formel ou modèle construit ? C'est ce que Kielkopf semble exiger, concluant de cette idée de vérité logique comme validité extra-systématique que la logique épistémique n'en est pas une ; elle contient en effet des relations de conséquence dont le comportement est de nature éthique ou normative (leurs propriétés métalogiques sont de type déontique) et non logique, ou aléthique. Cette caractérisation fait du calcul des propositions classique un système formel à part, le seul système proprement logique pour Kielkopf parce que capable de décrire les propriétés du vrai. On en revient ainsi à l'approche absolutiste de Frege, d'autant que Kielkopf assimile les *systèmes* intuitionniste et épistémique parce que leur méta-opérateur de prouvabilité affiche un comportement déontique commun.

Pour montrer que les inférences de [Hintikka 1962] n'ont rien d'épistémique, [Hocutt 1972] prétend quant à lui que la logique modale épistémique n'est qu'une logique classique déguisée, c'est-à-dire une logique qui tirerait toute sa légitimité du seul fait qu'elle conserve l'infrastructure *classique*. D'après lui, l'utilisation des conditions sémantiques telles que (C.K), caractéristique du système **T**, ou (C.K*), caractéristique du système **K**, aurait pour seule fin de réduire les formules modales à des formules non-modales et, sur la base de cette procédure sémantique, de réduire les inférences épistémiques à des fonctions de vérité :

> Par exemple, l'affirmation selon laquelle $[K_a(p \wedge q) \equiv K_a p \wedge K_a q]$ est auto-confirmante ne revient à rien de plus que l'affirmation que '$(p \wedge q) \equiv (p \wedge q)$' est une tautologie, et l'affirmation selon laquelle $[(K_a p \wedge (p \rightarrow q)) \rightarrow K_a q)]$ est un théorème équivaut à l'affirmation que '$(p \wedge (p \rightarrow q)) \rightarrow q$' est une tautologie. En bref, un théorème de logique épistémique est une affirmation quelconque qui s'avère être un théorème de la logique vérifonctionnelle lorsque les termes épistémiques ont été judicieusement éliminés.[159]

Dans les deux types d'arguments auxquels Kielkopf et Hocutt ont recours, les auteurs présupposent une compréhension pré-analytique des notions de vérité et de connaissance : les notions d'attitude épistémique ont des propriétés consacrées par l'usage, mais la logique formelle est incapable de les restituer sans postuler des agents idéaux ; la notion de vérité logique est réservée aux relations de conséquence de la logique classique, et tout langage formel n'est également logique que s'il se conforme à ces lois de l'être vrai. Si en revanche la logique classique n'a aucun monopole sur les propriétés du vrai, présupposer le contraire revient à admettre une forme de réalisme logique typique de la philosophie de Frege mais largement abandonnée depuis. Par opposition aux arguments de type [Kielkopf 1971] ou [Hocutt 1972], dans lesquels la vérité logique est assimilée à la validité universelle et n'est jamais relative à aucun modèle que ce soit, la vérité formelle d'une inférence épistémique se définit en fonction d'un ensemble de règles sémantiques et son adéquation informelle s'obtient relativement à un contexte d'usage de la langue naturelle. Pour user d'une

[159] [Hocutt 1972], pp. 442-3. [Cresswell 1972] prétend de même qu'il n'y a "aucun principe logique portant sur des affirmations de croyance qui ne soit déjà un principe portant sur les foncteurs de vérité seuls" (p. 11). L'interprétation de Hocutt est abusive, toutefois : (C.K) sert selon lui à biffer l'opérateur modal K, de sorte que la formule $(\phi \rightarrow K\phi)$ serait une conséquence logique du théorème d'introspection positive de **S4** : $(K)\phi \rightarrow (K)K\phi$. Hocutt a oublié de tenir compte du fait que la réduction $(\phi \rightarrow K\phi)$ est relative uniquement à ce que *a* sait et n'est pas valable en l'occurrence pour n'importe quel autre agent. En d'autres termes, ce n'est pas l'*énoncé* de forme générale '$\phi \rightarrow K\phi$' qui est déduit de $K\phi \rightarrow KK\phi$, mais l'*affirmation* de cet énoncé par le sujet connaissant.

métaphore, l'adéquation informelle d'un système de logique philosophique devrait être le résultat d'un 'équilibre réfléchi' entre l'usage admis pour un concept naturel et sa caractérisation au sein d'un langage formel quelconque. La logique philosophique n'impose donc ni un langage formel, ni un usage particuliers pour déterminer ses mots logiques ; sa valeur réside dans l'équilibre recherché entre une sémantique formelle et une interprétation naturelle.

Refuser une approche absolue ou extra-systématique de la vérité logique pose de nouveau la question du choix du juste système formel en vue de définir les termes formateurs des langues naturelles : ce choix des constantes logiques est-il purement arbitraire, ou suppose-t-on que la langue naturelle contient une algèbre qu'il nous reste à découvrir derrière ses structures grammaticales ? La question des critères d'adéquation (informelle) pour les théorèmes modaux avait été posée dans [Lemmon 1959], mais elle s'adressait encore à la logique des modalités en général : y a-t-il une et une seule logique modale 'correcte', entendue au sens l'adéquation informelle ? Son article concluait que le choix de la 'bonne' logique modale pure peut dépendre tout d'abord de l'interprétation des opérateurs modaux, selon que la modalité pure $\Box$ est interprétée comme la nécessité logique L, le temps H ou la croyance B ; mais même après interprétation, une formalisation exhaustive des propriétés d'une modalité lui paraît impossible :

> Comment devons-nous établir ce qu'est l'usage correct du mot, par opposition à l'usage ordinaire ? L'usage correct est une fonction très compliquée du modèle social de la communauté linguistique.[160]

La conclusion semble presque triviale de nos jours : la logique philosophique fait une distinction tranchée entre logique modale pure d'un côté, logiques modales interprétées de l'autre, et tout changement d'interprétation peut impliquer un changement dans le statut logique des formules modales ($\Box\phi \rightarrow \phi$ peut être admise ou non selon l'interprétation de $\Box$). L'idée de Lemmon autorise à penser qu'à chaque interprétation corresponde un système modal informellement (approximativement) adéquat : **S5** donnerait une analyse adéquate de la notion de vérité analytique (la L-vérité de [Carnap 1947]), **S4** une analyse adéquate de la notion intuitionniste de prouvabilité, etc. Le but d'une logique philosophique consisterait, semble-t-il, à construire un et un seul système approprié pour chaque concept philosophique –

[160] [Lemmon 1959], pp. 28-9. Pour Lemmon, la correction, c'est-à-dire l'adéquation informelle entre un langage formel et le fragment de la langue naturelle qu'il traite ne peut être qu'approximative : "le calcul propositionnel classique (...) peut toutefois prétendre être 'correct' en tant que simple modèle logique *proche de* convenir à l'usage correct" (*loc. cit.*, p. 28) ; il n'y a donc d'adéquation qu'approximative entre les deux types de langages, et "la chose est d'autant plus vraie dans le cas des logiques modales, parce qu'il semble peu probable que 'il est nécessaire que' ait une signification unique claire" (p. 26). Bien qu'opposé à un conventionnalisme radical de type [Carnap 1930] ou [Rougier 1941], Lemmon insiste sur l'idée que la signification 'réelle' d'un connecteur n'équivaut pas à son usage ordinaire : "Une enquête sociologique sur la pratique linguistique effective le sous-entend elle-même : les gens se basent-ils, ou ne se basent-ils pas, sur la loi du tiers exclu dans l'argumentation réelle ? Si oui, alors les mots 'soit ... soit' sont utilisés d'ordinaire de manière telle que les illustrations de la loi du tiers exclu sont vraies en vertu de cet usage ; sinon, elles ne le sont pas. En arriver à cette idée, c'est constater qu'elle ne fonctionnera pas non plus. Car nous voulons dire qu'une loi du raisonnement, aussi largement utilisée soit-elle dans la pratique, pourrait encore être fausse (la majorité pourrait se prononcer de façon incorrecte). En conséquence, il nous faut faire une distinction entre la façon dont les gens utilisent un mot *d'ordinaire* et la signification que le mot a *en réalité*" (p. 28).

analyticité, existence, temps, obligation, connaissance, etc. Mais la difficulté est plus profonde et jette un doute sur le projet même de logique philosophique : il peut y avoir plusieurs modèles pour un seul et même opérateur modal interprété. C'est le cas en logique modale épistémique, où plusieurs systèmes distincts peuvent être considérés comme adéquats : **S4** pour Hintikka, **S4.4** pour Kutschera, **S5** pour van der Hoeck, etc. Face à cette pluralité des interprétations épistémiques, tous liées à un usage distinct du concept de connaissance, "mon sentiment personnel", dit [Halpern 1996],

> c'est qu'il n'y a pas une bonne et unique notion de connaissance ; la notion appropriée dépend des applications.[161]

L'affirmation paraît triviale mais pose problème : si la relativité dans l'usage implique qu'il n'y a pas *un* système de logique philosophique adéquat pour chaque modalité interprétée, à quoi bon la logique en philosophie ? Il y a une réponse à cet argument de la variabilité des interprétations : pour reprendre la métaphore du fil d'Ariane, l'intérêt d'une logique philosophique ne réside pas tant dans le choix des prémisses (d'une structure modale) que dans les conséquences imprévues de ces prémisses (les conséquences logiques de cette structure) : un langage modal sera considéré comme adéquat si l'ensemble de ses validités est consolidé par l'usage. Hintikka exprime cette adéquation informelle en termes de plausibilité du système, ce qui

> veut dire l'accord avec toutes les intuitions que nous pouvons avoir concernant les concepts impliqués.[162]

Or nos intuitions pré-formelles sont susceptibles de s'opposer les unes aux autres, lorsqu'elles ne suivent pas la même interprétation informelle d'un opérateur ; nous le constaterons en ce qui concerne notamment la thèse KK (introspection positive) et le lien entre connaissance et justification. Face à cette complexité des mots de la langue naturelle, la tâche du logicien philosophe est de trouver un moyen de justifier leur transcription en mots logiques, simplifiés et schématisés au sein d'un modèle formel.[163] Alors que Quine voyait dans cette pluralité d'usages une raison suffisante pour considérer les verbes d'attitude propositionnelle comme 'trop colorés', c'est-à-dire dont le sens est trop variable pour constituer de véritables mots logiques, le but de [Hintikka 1968] est de justifier au contraire

---

[161] [Halpern 1996], p. 483.
[162] [Hintikka 1968], p. 3.
[163] Contre les critiques d'irréalisme adressées notamment par [Hocutt 1972], [McLane 1974] souligna que, comme dans n'importe quel modèle scientifique, une théorie logique est forcée de simplifier et schématiser la signification de ses objets d'étude : "Le système de Hintikka est en effet un idéal théorique qui n'a qu'une illustration partielle dans le monde 'réel'. Dire, comme le fait Hocutt, qu'un tel système ne peut pas être une 'logique' parce que les théorèmes ne sont pas toujours vrais du monde réel, c'est poser une restriction trop drastique sur la constructions des systèmes logiques. Le problème ici est similaire à celui soulevé par les géométries non-euclidiennes et les logiques multivalentes. La tendance générale parmi les logiciens consiste à choisir une définition plus 'libérale' de la logique que celle proposée par Hocutt" ([McLane 1974], pp. 562-3). Le problème est de savoir dans quelle mesure cette simplification ne dépasse pas les limites de la pertinence – un problème popperien, lié au critère de l'audace ou falsifiabilité de la théorie. Sur le rôle de la schématisation comme simplification en logique philosophique, voir notamment Sven Ove Hansson : "Formalization in philosophy", *The Bulletin of Symbolic Logic* **6**(2000), pp. 162-175.

un certain usage des verbes d'attitude épistémique. La prescription d'un certain usage nous renvoie à sa conception de la logique philosophique comme *modèle explicatif.*

2.2.3 La théorie du modèle explicatif de Hintikka

Il y a au moins deux raisons de rejeter l'idée d'une adéquation possible entre les règles déductives de la logique et les significations ordinaires des énoncés. Non seulement la logique a souvent pour but de mettre à jour et corriger les incohérences des expressions du langage ordinaire et, en cela, on ne saurait prétendre à une adéquation avec ce que l'on cherche à corriger. Mais les expressions de ce langage ordinaire paraissent également trop ambiguës pour être réductibles à un ensemble d'opérations abstraites, détachées du contexte de discours.

En raison de cette impossibilité apparente, la logique philosophique doit faire un choix pour formaliser certaines expressions du langage ordinaire : soit les constantes logiques sont en nombre limité mais invariables ; soit les constantes logiques sont en nombre potentiellement illimité mais variable. En d'autres termes, une constante logique représente soit la signification constante d'une expression naturelle au sein d'un unique langage formel, soit la représentation variable de cette expression au sein de plusieurs langages formels. Des cas d'étude du genre sont les expressions naturelles de la conséquence et de la négation : 'si ... alors' et 'ne ... pas' ont-elles un seul et même comportement logique dans tous les cas, ou plusieurs comportements relatifs au contexte dans lequel ces locutions sont utilisées ? Pour ces deux stratégies opposées, nous avons fait allusion aux positions de l'universalisme et du relativisme logique : soit les opérations qui définissent les constantes logiques sont invariables et s'attachent comme par essence à certaines expressions du langage ordinaire ; soit elles sont totalement variables et relatives à un ensemble de conventions d'usage choisies par le logicien. Nous avons déclaré par la suite qu'une logique philosophique ne peut se résoudre ni à la première, ni à la seconde approche : le critère de l'usage suggère un ensemble d'opérations susceptibles de formaliser au mieux certaines expressions du langage ordinaire, mais sans prétendre à leur universalité d'application. C'est cette voie médiane que Hintikka retient pour justifier sa logique épistémique : les opérateurs modaux épistémiques ne donnent pas une analyse fondamentale et irréfutable des verbes de croyance et de connaissance, certes, mais ils sont censés en retenir les traits *essentiels.*[164] Il nous reste à montrer en quoi les modalités épistémiques retiennent les caractéristiques 'essentielles' de ces attitudes.

La voie suivie par Hintikka reprend la notion de *jeu de langage* du second Wittgenstein : par opposition à l'espace logique immuable du *Tractatus*, dans lequel le langage logique est idéal et associe une et seule signification à tous les termes, [Hintikka 1968] invoque certains passages du *Livre Bleu* où Wittgenstein fait allusion à des langages de profondeur, sortes de proto-langages dans lesquels les expressions seraient pourvues de certaines règles d'usage mais qui ne prendraient forme que dans les langues de surface, c'est-à-dire les langues

[164] [Lenzen 1978] fait remarquer (p. 5) que "[Fisher 1964] définit comme but de la logique doxastique celui 'de formuler un ensemble d'affirmations qui sont vraies en raison de la signification du mot 'croit', i.e. en raison de la *nature* du concept de croyance" (p. 166). La notion de 'nature' est douteuse puisqu'elle fait songer à un ensemble de propriétés invariables du concept de croyance ; cette expression est contraire à l'idée de signification dans l'usage et de jeux de langage, que le modèle explicatif de Hintikka va prolonger dans ce qui suit.

ordinaires.[165] D'après Hintikka, le rôle de la logique philosophique consisterait à extraire de ces jeux de langage les significations *premières* des expressions ordinaires et à en détacher les significations *résiduelles*, produites par des contextes d'usage.[166] A partir de ce modèle des jeux de langage et de ses propriétés dites de profondeur, Hintikka légitime le recours à la logique formelle en vue de l'étude des concepts ordinaires ;[167] car tandis que la complexité des langues naturelles avait incité Wittgenstein à délaisser la logique formelle pour se tourner vers une étude plus prudente du langage (en termes informels de règles grammaticales), Hintikka prétend au contraire que la notion de jeu de langage ne donne pas définitivement raison au courant des philosophes du langage ordinaire de l'école d'Oxford (Austin, Strawson, Ryle). Selon lui, il est possible de formaliser des tournures aussi complexes que les attitudes propositionnelles, mais à condition d'assouplir le statut de la logique formelle et d'abandonner son image trop rigoureuse de notation canonique (observée dans la logique des fonctions de vérité et quantificateurs de Quine).

Contre cette utilisation normative de la logique, mais sans tomber pour autant dans un inventaire descriptif des différents usages d'une 'constante' logique, Hintikka propose sa conception intermédiaire, prescriptive, de la logique (philosophique) :

> Une branche de la logique, la logique épistémique par exemple, est la mieux conçue comme un *modèle explicatif* en termes duquel les mécanismes de notre langage ordinaire peuvent être compris. Dans certains cas, ce modèle explicatif peut être conçu comme faisant ressortir la 'logique de profondeur' qui sous-tend les réalités complexes de notre usage ordinaire des termes

---

[165] [Hintikka 1968], note 12 : "Dans le *Blue Book*, Basil Blackwell, Oxford, 1958, p. 17, Wittgenstein écrit 'L'étude des jeux de langage est l'étude des formes de langage primitives ou des langages primitifs. Si nous voulons étudier les problèmes de vérité et de fausseté, d'accord et de désaccord de propositions avec la réalité, la nature de l'assertion, de l'assomption et des questions, nous aurons grand avantage à regarder les formes de langage primitives dans lesquelles ces formes de pensée apparaissent sans l'arrière-plan confondant de processus de pensée hautement compliqués. Lorsque nous regardons de telles formes de langage simples, la brume mentale qui semble envelopper notre usage ordinaire du langage disparaît. Nous voyons des activités, des réactions qui sont nettement définies et transparentes. D'autre part, nous reconnaissons dans ces processus simples des formes de langage non séparées par une rupture de nos formes de langage plus compliquées. Nous pouvons constater que nous pouvons édifier des formes compliquées à partir des formes primitives, en ajoutant progressivement de nouvelles formes'. Ceci avait constitué en fait en grande partie le programme de Wittgenstein dans le *Livre Brun*. Le fait qu'il y incite le lecteur à ne pas concevoir 'les jeux de langage que nous décrivons comme des parties inachevées d'un langage, mais comme des langages achevés en soi, comme des systèmes de communication humaine achevés' (p. 81) ne dément pas cette simplicité. C'est simplement une injonction contre le fait d'introduire 'l'arrière-plan confondant de processus de pensée hautement compliqués', lorsque l'on considère un jeu de langage." Hintikka se sert ici de Wittgenstein pour répondre à ses détracteurs précédents.

[166] "La façon dont un modèle explicatif est supposé éclairer ce qui se passe dans le discours ordinaire pourrait être expliquée de la manière suivante : Nous appellerons la signification qu'une expression aurait dans le modèle explicatif sa *signification de base*. Or nous pouvons partir de cette signification de base et voir comment elle sera modifiée par les différents facteurs (...) Les significations qui en résultent, pour autant qu'elles diffèrent de la signification de base, pourraient être appelées les *significations résiduelles*. Si notre modèle explicatif est un modèle approprié, et si nous avons discerné comme il faut les facteurs pragmatiques et autres facteurs supplémentaires impliqués dans les différents cas, nous serons de cette façon en mesure d'expliquer ce qui se passe réellement aux différentes occasions de l'usage ordinaire" ([Hintikka 1968], pp. 6-7).

[167] Un autre enseignement du second Wittgenstein est la comparaison de la quantification au résultat d'une enquête : quantifier sur un énoncé, c'est identifier dans le monde un élément satisfaisant une question de départ. Hintikka verra dans cette interprétation une filiation directe avec la théorie de la quantification de Peirce et la théorie de l'assertion de Frege (engagement sur une question initiale), mais aussi et surtout une source d'inspiration pour sa future sémantique des jeux GTS (Game-Theoretical Semantics). Nous nous en tenons ici à son interprétation des attitudes propositionnelles.

épistémiques ('savoir', 'croire', etc.) et en termes duquel ces complexités peuvent être expliquées.[168]

En réponse aux détracteurs de [Hintikka 1962], l'auteur précise que la logique de profondeur des attitudes épistémiques ne correspond pas à leur usage courant ou prédominant, usage qu'il assimile à l'expression d''usage paradigmatique' de la linguistique empirique :

> La signification de base d'une expression n'est pas toujours, et peut-être pas même d'habitude, sa signification courante. Il peut même arriver qu'une expression n'ait *jamais* sa signification de base dans le langage ordinaire, du moins pas en dehors du discours des philosophes.[169]

Ce semble être le cas avec l'opérateur de connaissance K, dont les propriétés de distribution et de clôture sous implication heurtent notre 'intuition langagière'. Ce le sera également (quoique dans une moindre mesure) pour la propriété d'introspection positive du système modal **S4**, pour lequel on peut trouver des contre-exemples dans l'usage ordinaire du concept de savoir. La distinction entre une signification *de base* et une signification *résiduelle* des mots logiques repose essentiellement sur des distinctions contextuelles dans leur usage : la signification d'un opérateur modal tel que la croyance peut varier selon le but en vue duquel il est utilisé et la variété de ce but explique les réticences face à l'entreprise de formalisation de ce concept, jugé 'trop coloré' par Quine et auquel celui-ci refusait le titre de constante parce qu'il soutenait pour sa part une conception normative de la logique.[170]

Par contraste avec l'argument de Quine, Hintikka souligne que la polysémie d'un concept ordinaire n'est pas une raison suffisante pour lui refuser le titre de mot logique ; il en va de même après tout pour les fonctions de vérité, parmi lesquelles la négation :

> le problème que nous traitons ne se limite pas à la logique épistémique. Pourquoi sommes-nous justifiés d'incorporer la loi de la *double négation* dans notre logique propositionnelle ordinaire ? Nul doute qu'une expression doublement liée du langage ordinaire, s'il y a, a très rarement les

---

[168] [Hintikka 1968], p. 5. L'auteur ajoute que ce modèle explicatif, loin d'être donné par l'expérience du langage ordinaire, "doit être *inventé* plutôt que *découvert*" (*ibid.*) La logique de profondeur n'est donc pas une donnée expérimentale (pour la linguistique empirique) mais un modèle théorique.

[169] [Hintikka 1968], pp. 7-8. [Hintikka 1970a] confirme cette distinction entre la signification de base d'un concept et son usage courant : "En aucun cas l'acceptabilité de la thèse ne peut être décidée en faisant appel au 'langage ordinaire'. Il me semble en fait qu'une grande part de la posture méthodologique des philosophes analytiques est bien trop timide. L'ultime court d'appel pour décider si un principe logique gouvernant un concept donné est acceptable n'est pas l'usage ordinaire, aussi enrégimenté soit-il, c'est plutôt si le principe aide le concept en question à servir la ou les fins qu'il est en destiné en fait à servir dans notre répertoire conceptuel, et si ces fins elles-mêmes méritent notre effort. En décrivant ces fins et en les utilisant pour évaluer divers principes logiques, un analyste peut parfaitement bien être en désaccord avec l'usage ordinaire et même tenter de le réformer" (p. 141).

[170] Hintikka inclut dans les significations résiduelles les implications dites 'pragmatiques', comparables aux *implicatures* de Grice ; ces significations résiduelles sont produites par "les facteurs de la nature suivante : (i) D'autres objectifs, concurrents. Souvent, ce sont les objectifs vaguement généraux que pratiquement tout discours est supposé servir, ou du moins est supposé ne pas entraver. Ainsi, de nombreuses formes de discours servent à conserver la faculté qu'ont d'autres d'évaluer toute situation qui puisse se produire. Cet objectif n'est pas servi très efficacement si le locuteur ne fait pas des assertions aussi complètes (explicites) qu'il est en position de pouvoir faire. (ii) Diverses pressions pragmatiques, telles que la pression de ne pas user de circonlocutions sans une finalité spécifique. (iii) Diverses limites innées de l'esprit humain, par exemple les limites d'une mémoire à court terme. (iv) Les pressions dues au contexte particulier dans lequel une proposition est énoncée ou écrite" ([Hintikka 1968], p. 6).

> mêmes pouvoirs logiques que la proposition non niée de départ. Notre logique propositionnelle ne distord-elle pas grossièrement la logique du langage ordinaire ? La réponse est (très brièvement) que si la signification de base est supposée équivalente à celle de l'expression non niée de départ, nous pouvons expliquer les significations résiduelles qu'une expression doublement niée a à différentes occasions.[171]

Il reste maintenant à comprendre ce qui fait d'une explication d'un concept sa signification de base ou sa signification résiduelle : par exemple, l'implication stricte exprime-t-elle une signification résiduelle du conditionnel, produite par des facteurs pragmatiques ? Et quand bien même ce serait le cas, en vertu de quoi l'effet pragmatique d'une expression ne devrait-elle pas faire partie de ses caractéristiques logiques ?[172] Pourquoi considérer le comportement involutif de la négation comme caractéristique de sa signification de base ? Pour Hintikka, la théorie du modèle descriptif sert d'arrière-plan méthodologique visant à justifier sa logique épistémique en tant que logique philosophique ; si par exemple l'opérateur de connaissance est transitif et satisfait le système **S4**, c'est parce "le principe aide le concept en question à servir la ou les fins qu'il est en destiné en fait à servir dans notre répertoire conceptuel" et que "ces fins elles-mêmes méritent notre effort", ajoute [Hintikka 1970a]. Mais l'auteur ne précise pas ce qu'il entend par un 'répertoire conceptuel', ni pourquoi la propriété d'itération mérite de faire partie des propriétés caractéristiques du concept de connaissance. Ce travail méthodologique paraît pourtant nécessaire si l'on veut justifier l'exercice général de formalisation des langues naturelles. En quoi une telle formalisation permet-elle d''apporter un éclairage' à plusieurs concepts philosophiques dominants, pour reprendre les mots de [Hintikka 1969] ? Il s'agit de montrer dans ce qui suit que certaines propriétés logiques des modalités épistémiques peuvent être privilégiées afin d'en tirer des conséquences caractéristiques de l'usage des verbes 'savoir' et' croire'. Pour comprendre pourquoi Hintikka a choisi des caractérisations de type **KT4** pour le savoir et **KD4(5)**[173] pour la croyance,

---

[171] [Hintikka 1968], p. 7. On trouve dans [Hintikka 1962] des exemples de significations résiduelles de la négation : "Si l'expression doublement niée a les mêmes pouvoirs logiques que l'expression non niée, celui qui utilise la première doit probablement viser un effet spécial qui ne fait pas partie de la *logique de base de la négation*. Peut-être que l'effet le plus courant de ce genre consiste à signifier l'*hésitation* ou l'*incertitude* de la part du locuteur. Mais ce n'est pas le seul effet possible. Par exemple, le double négatif peut signifier la timidité ; à d'autres moments, il se peut qu'il soit ironique ("Avoir une masse d'argent ne va pas sans avantages") ou destiné à suggérer une réserve différente de l'hésitation. Ou bien la raison d'utiliser un double négatif peut-elle être simplement une préférence stylistique pour la construction en présence" (p. 115). Les significations résiduelles de la négation n'avaient pas à figurer parmi les propriétés de la négation en logique *mathématique*, dans laquelle '¬' a pour seule fonction de désigner la relation ensembliste de complémentation.

[172] [Gardiès 1979] associe le problème de l'implication à des raisons d'ordre pragmatique : "avant même l'expérience des diverses logiques modales, l'aventure de l'humble calcul des propositions suggérait que l'intuition joue avec les *structures* plus qu'elle ne s'y tient : la table de vérité de l'implication reste à nos yeux fondamentalement ce qui justifie le sens du *si* ; mais des raisons d'ordre *pragmatique*, qu'ont parfaitement analysée Russell et Tarski, ne permettaient à l'intuition de n'en avoir une conscience que partielle ; d'où ces 'paradoxes de l'implication' qui, depuis les stoïciens, éveillaient le soupçon sur la part de la logique au demeurant sans doute la moins contestable" (pp. 232-3 ; nous soulignons). Une raison d'exclure les aspects pragmatiques du domaine de la logique formelle tient probablement à la distinction créée par Charles Morris entre syntaxe (relation signe-signe), sémantique (relation signe-objet) et pragmatique (relation signe-locuteur) : seuls les deux premiers aspects sémiotiques du langage devraient être partie prenante en logique. Nous reviendrons et rejetterons plus loin cette séparation ; à titre d'illustration, la pragmatique formelle (Searle, Vanderveken) constitue une analyse logique des aspects pragmatiques du langage.

[173] Nous mettons (5) entre parenthèses parce que [Hintikka 1962] ne se prononce pas sur la validité ou non-validité du théorème S5 pour la croyance : $\neg Bp \rightarrow B\neg Bp$. Notre commentaire précédent sur la notion de doute

observons leurs propriétés qu'il a cherché à justifier. Nous proposons dans ce qui suit une étude des axiomes épistémiques admis dans [Hintikka 1962] : dans quelle mesure les axiomes monomodaux **K**, **T**, **S4** et d'autres théorèmes multimodaux[174] formalisent-ils la signification 'de base' de K et B ?

2.2.4 Problèmes en logique épistémique propositionnelle

La logique modale épistémique $\mathcal{L}_{f.ME}$ contient à la fois une syntaxe et une sémantique, un ensemble d'axiomes et une interprétation des formules. La propriété d'adéquation formelle garantit l'équivalence entre les formules déductibles et valides, mais [Hintikka 1968] a surtout insisté sur l'intérêt explicatif de l'approche sémantique : définir les opérateurs modaux en termes de structures de mondes possibles permet d'interpréter les modalités et de mieux comprendre les conditions dans lesquelles un agent sait ou croit quelque chose. Or ces conditions paraissent trop élevées dans un système de logique modale normale : en raison de leur aspect contre-intuitif, les axiomes de la logique épistémique ont souvent été qualifiés d'irréalistes. Nous allons passer en revue les différents axiomes épistémiques et doxastiques défendus par Hintikka, ainsi que les raisons de les maintenir malgré cet aspect contre-intuitif. Avec en point de mire le problème de l'interprétation et l'application des modalités épistémiques : quel est l'usage le plus adéquat des modalités épistémiques, susceptible de donner un sens acceptable à leurs comportement logique ?

2.2.4.1 L'omniscience logique

Il existe un moyen de défendre malgré tout les conséquences 'contre-intuitives' du système **K**. Si toutes les conséquences logiques d'une connaissance sont également connues, comme le veut la propriété de clôture épistémique sous implication logique : $K\phi, (\phi \Rightarrow \psi) \Rightarrow K\psi$, c'est dans la mesure où le sens de ψ est contenu 'dans' le sens de ϕ et en résulte. Cette interprétation de la conséquence logique comme intrinsèquement liée à sa condition est exprimée notamment dans le *Tractatus* de Wittgenstein, lorsque l'auteur prétend que "si p résulte de q, le sens de 'p' est contenu dans le sens de 'q'.[175] Or cette acception de l'implication ne semble pas concorder avec la vérité épistémique : l'ensemble de nos connaissances, mêmes rationnelles, ne devrait pas coïncider avec l'ensemble des vérités logiques en raison de nos ressources déductives limitées. L'omniscience logique constitue donc un problème fondamental pour une logique philosophique, puisqu'elle est contenue dans le système modal minimal **K** et empêche de rendre compte de nos limitations déductives.

---

nous inciterait à penser que ce théorème n'est pas valide, si la non-croyance est interprétée comme l'indifférence ou le doute par rapport à la vérité de p ; mais rien ne nous montre que cette signification de la non-croyance comme indifférence ne soit pas simplement 'résiduelle'. En revanche, il semble que la formule ci-dessus soit 'doxastiquement auto-confirmante'. Nous reviendrons plus loin sur le sens de cette expression, liée à l'analyse de la modalité S4.

[174] Un énoncé est *monomodal* s'il contient un seul type de modalité, par exemple : Kp, Bp, Fp, Op, etc. Il est *multimodal* s'il contient plusieurs modalités, par exemple : FBp, OBp, etc. Les modalités itérées de la logique épistémique : KBp, BKp, KBKp, etc. sont des modalités *mixtes* appartenant à la même famille conceptuelle ; mais on peut combiner également des opérateurs modaux issus de familles conceptuelles différentes : 'je sais que je dois faire p' (KOp), 'je croirai que p est interdit' (FB¬Pp), etc.

[175] [Wittgenstein 1969], §5.122.

Plusieurs travaux consécutifs à [Hintikka 1962] vont tenter de surmonter cet aveu d'impuissance de [Gardiès 1975] :

> La difficulté propre d'une logique épistémique tient à ce que le savoir ordinaire, celui que nous rencontrons dans notre expérience de la condition humaine, est à égale distance de l'irrationalité comme de la rationalité totale, sans qu'il soit possible de déterminer un niveau censé caractéristique de l'homme.[176]

La tâche du logicien épistémique consistera à corriger cet effet indésirable dû aux propriétés d'une logique modale normale (règle de nécessitation) et de modifier les règles de construction d'une logique épistémique. Mais à quel prix faut-il changer la logique modale de base, et à quelles propriétés faut-il s'attaquer ? Nous préciserons nos raisons pour lesquelles un changement dans l'interprétation des constantes logiques (leurs règles de transformation) serait toutefois mal avisé, et par quels autres moyens l'omniscience logique pourrait être évitée.

Le cœur de l'argument est le suivant. Le défaut du système modal minimal **K** et de toutes les structures de Kripke, qui reposent sur celui-ci, c'est qu'ils présupposent l'impossibilité d'ignorer une conséquence logique : chaque monde possible contenant la formule ϕ doit contenir également ψ si ϕ implique ψ, et il est impossible qu'il contienne à la fois ϕ et $\neg\psi$. Ce défaut est structurel et tient aux propriétés du système **K** (la *monotonie*, exprimée par la règle de nécessitation). Bien que l'on recense plusieurs types de possibilité : métaphysique, logique, épistémique, physique (causale), la possibilité épistémique désigne ici la consistance d'un ensemble des informations que possède un agent rationnel ; le critère de consistance montre ainsi que les mondes épistémiquement possibles sont inclus dans les mondes logiquement possibles, puisqu'il détermine les connaissances qu'un agent est en droit d'ajouter à son stock d'informations déjà acquises (son 'espace épistémique'). Or si la règle de nécessitation (ou nécessitation) paraît admissible pour les modalités aléthiques, elle ne l'est pas pour les modalités épistémiques. Ce résultat est contre-intuitif : les capacités déductives humaines sont trop limitées pour qu'un agent soit capable de déduire l'infinité des conséquences logiques de son stock de connaissance ; tout au plus, la capacité d'un agent à déduire ces conséquences varie en fonction de sa compétence logique pour 'percevoir' ces conséquences. Avant de remédier à ce défaut des structures de Kripke, revenons d'abord sur l'idée de compétence déductive (la capacité d'un agent à tirer les conséquences d'informations déjà connues) et tentons de mieux comprendre l'origine du problème : pourquoi l'omniscience logique est-elle considérée comme 'contre-intuitive' ?

Qu'est-ce qu'une vérité logique ?

Dans un tout autre registre que la question des compétences déductives, la clôture épistémique sous implication logique a été utilisée comme règle d'inférence destinée à étayer un argument sceptique : si je sais que j'ai une main (Kp), et si l'existence de ma main implique que je ne suis pas un cerveau dans une cuve ($p \Rightarrow \neg q$), alors je sais que je ne suis

[176] [Gardiès 1975], p. 124.

pas un cerveau dans une cuve ($K\neg q$). Or puisque je suis incapable de réfuter l'hypothèse de la cuve ($\neg K\neg q$), j'en conclus par contraposition que je ne sais pas si j'ai une main ($\neg Kp$) . Le même argument sera appliqué avec la clause de succès $K\phi \rightarrow \phi$ (section 2.2.4.3) pour réfuter la preuve de mon existence, mais le problème de la preuve porte ici sur la signification de $\Rightarrow$. Est-ce bien l'implication formelle $\Rightarrow$, c'est-à-dire la relation métalinguistique de déductibilité qui est en jeu dans cet argument sceptique, ou faut-il introduire une autre sorte de relation de conséquence pour l'exprimer ? Le cas du cerveau dans une cuve ne semble pas faire appel à l'implication formelle, et la question de l'omniscience logique ne sera pas abordée dans les mêmes termes : il ne s'agit pas de tirer des conséquences sceptiques du principe de clôture, mais, au contraire, de restreindre la validité de ce principe. En d'autres termes, le problème consiste à délimiter la classe des vérités logiques *connaissables* par un agent. Cela étant, l'argument sceptique a été discuté entre autres par [Nozick 1981] et a conduit à une solution, la théorie du pistage [*keeping track*] de la connaissance et sa contrepartie sémantique en termes de mondes possibles 'proches' : une inférence épistémique entre un antécédent ϕ et un conséquent ψ est satisfaite si ϕ et ψ appartiennent à des espaces que le sujet est en mesure de relier ou avec lesquels il peut garder le contact. La solution suppose une restriction des formules entre lesquelles la relation de déductibilité sera applicable par un agent, ce qui suppose la construction de mondes possibles 'incomplets'. Le problème porte sur les conditions de ce pistage : exprime-t-elle une relation logique entre un antécédent et un conséquent, ou ne renvoie-t-elle pas plutôt une à une sorte de nécessité causale, ou physique ? Interprété souvent en termes plus restreints d'implication stricte, non pas au sens de la nécessité logique de C. I. Lewis mais au sens d'une relation plus fine d'entraînement [*entailment*], l'argument du sceptique ouvre indirectement des débats sur les frontières de la vérité logique mais qui ne concernent pas directement le problème de l'omniscience logique. Pistage et compétence déductive sont liés tous deux à la même règle d'inférence épistémique mais s'intéressent à des aspects différents de la notion de connaissance. Le problème est celui de la nature de la relation d'accessibilité entre des mondes *épistémiquement* possibles. Est-ce une relation causale, physique, temporelle, etc.? Quoi qu'il en soit, l'argument du sceptique ne concerne pas la logique épistémique de Hintikka : dans celle-ci, les relations d'inférence épistémique permises par le système **K** concernent seulement les formules vraies dans toute structure, et la question de savoir si la relation entre avoir une main et ne pas être un cerveau dans une cuve est de nature logique ou déductive est laissée de côté. Le cas des mondes possibles proches de [Nozick 1981] devait être évoqué toutefois, puisque la règle de clôture épistémique pose indirectement la question de savoir ce qu'est une vérité logique. La différence entre vérités logiques et vérités *analytiques* est à l'origine de ce litige possible, mais ce qui suit ne va pas consister dans une défense de la règle de clôture. Au contraire, Hintikka va tenter d'éliminer l'omniscience logique en construisant un modèle dans lequel les mondes épistémiquement possibles ne soient *pas* inclus dans les mondes logiquement possibles : tout monde épistémiquement possible n'est pas forcément logiquement possible.

Les données du problème

[Dubucs 1991] et [Dubucs 1998] caractérisent les trois hypothèses à l'origine du paradoxe sous la forme d'un trilemme : si l'on souhaite le résoudre, la solution consiste à rejeter l'une

des hypothèses pour rétablir la cohérence. Soient trois hypothèses A, B et C, dont la conjonction aboutit à une absurdité :

***Hypothèse A** (vérité logique):* les vérités logiques sont des formules vraies dans *tous* les mondes logiquement possibles ;

Hypothèse B (non-omniscience logique) : l'agent épistémique ne connaît pas toutes les conséquences logiques de ce qu'il sait ;

Hypothèse C (logicité) : tous les mondes épistémiquement possibles sont logiquement possibles : $W_K \subseteq W_L$, donc que toute relation de conséquence logique entre des formules est connue par l'agent épistémique.

Quelle paire d'hypothèses doit être maintenue, et quelle hypothèse doit être abandonnée : maintenir AB et sacrifier C (noté : AB¬C) ; maintenir AC et sacrifier B (AC¬B) ; ou maintenir BC et sacrifier A (BC¬A) ?

La première hypothèse A ne sera pas contestée ici : par définition, la vérité logique est supposée valide dans n'importe quelle structure, donc dans n'importe quel ensemble de mondes possibles quelle que soit les relations d'accessibilité de la structure.[177] La solution restante est donc soit AB¬C, soit AC¬B. Si l'on rejette la seconde hypothèse B, alors un agent épistémique est censé être un logicien impeccable auquel aucune conséquence logique n'échappe, aussi complexes que puissent être les étapes déductives menant des prémisses à la conclusion. Hypothèse douteuse, à laquelle notre intuition nous enjoint de trouver une réplique. C'est cette hypothèse de la perfection déductive des agents épistémiques que la plupart des commentateurs considèrent comme rédhibitoire, parmi lesquels [Hintikka 1975a]. Puisque nous maintenons l'hypothèse A, la seule solution plausible qui reste est donc AB¬C, qui conduit au rejet de C : certains mondes épistémiquement possibles ne le sont pas logiquement, ce qu revient à dire qu'il est possible d'avoir des attitudes épistémiques inconsistantes. Cette dernière solution va être suivie par Hintikka, comme nous allons le voir plus loin avec son cadre des 'mondes possibles logiquement impossibles' : si certains espaces épistémiques sont exclus des espaces logiques, alors l'agent peut se comporter de manière inconsistante (incohérente).

Examinons les différentes combinaisons proposées, à l'exception de BC¬A : nous supposons incontestée l'hypothèse selon laquelle les vérités logiques sont vraies dans tous les mondes logiquement possibles. Le problème vient donc soit de B, soit de C. Quelle hypothèse rejeter ?

- Option A¬BC ?

Il semble envisageable de sauver malgré tout l'omniscience logique, ce qui revient à rejeter l'hypothèse B de non-omniscience. Deux conditions existent pour ce faire, qui ne sont pas du même ordre mais impliquent tous les deux le maintien des hypothèses A et C : pour rendre acceptable l'hypothèse de l'omniscience logique, soit l'on modifie le contenu de A, c'est-à-dire la définition des mondes *logiquement* possibles ; soit l'on modifie le contenu de C, c'est-

[177] La comparaison faite dans la Partie 1 entre mondes possibles et matrices nous permet de le dire autrement : si l'on compare les mondes logiquement possibles à des combinaisons de valeurs de vérité, une vérité logique est une formule vraie pour toute interprétation de ses éléments, c'est-à-dire qu'elle donne le vrai dans chaque dernière colonne de la matrice. La comparaison entre valuations et modèles confirme en d'autres termes l'hypothèse (A) : être logiquement vrai, c'est être vrai dans chaque colonne ou monde logiquement possible.

à-dire la définition des mondes *épistémiquement* possibles. Dans le premier cas, on assouplit les critères du concept de vérité logique de façon à ce qu'un ensemble épistémique inconsistant devienne logiquement possible ; dans le second cas, on assouplit les critères du concept de connaissance de façon à ce qu'une conséquence logique soit connue dans tous les cas. Nous allons considérer tout d'abord le cas où la notion de vérité logique est modifiée, avant de passer au cas de la notion de connaissance.

Dans une logique modale épistémique *normale*, l'agent épistémique ne peut pas admettre (donc ignore) certaines formules ψ si ces formules sont inconsistantes par rapport à ce qu'il sait déjà, par exemple que ϕ est le cas : $(K\phi, (\phi \Rightarrow \neg\psi) \Rightarrow \neg K\neg\psi)$; mieux que de les ignorer, il sait que ces formules ne peuvent pas être admises si elles produisent une inconsistance $(K\phi, (\phi \Rightarrow \neg\psi) \Rightarrow K\neg\psi)$. C'est le présupposé de *consistance* des vérités logiques qui produit ce résultat : un agent épistémique pense toujours logiquement, c'est-à-dire conformément aux règles de consistance (ou de normalité) qui définissent les vérités logiques et satisfont A. Mais il existe une astuce formelle, capable de valider des ensembles épistémiques inconsistants tels que $\{\phi, \phi \Rightarrow \psi, \neg\psi\}$ sans remettre en cause ni l'hypothèse C, ni l'hypothèse A : certains ensembles épistémiques deviennent à la fois inconsistants et logiquement possibles, contrairement à la condition fondamentale de consistance (C.¬) des logiques modales normales dont [Hintikka 1962] est un échantillon.

Lorsque un modèle épistémique contient un ensemble de formules inconsistant, il constitue ce que Hintikka appelle un 'monde possible logiquement impossible' : un espace d'information épistémiquement possible mais dans lequel se trouve un énoncé logiquement faux, donc impossible. Pareil scénario est-il concevable, ou l'inconsistance est-elle totalement incompatible avec la vérité logique ? Une possibilité de ce genre exigerait une révision au sein de la théorie des modèles, notamment en ce qui concerne la propriété de consistance (C.¬) : on pourrait imaginer à cet effet des conditions sémantiques de type *paraconsistant*, par exemple, permettant que certains modèles soient à la fois inconsistants et 'logiquement' possibles. Cette stratégie de révision des propriétés de la conséquence logique impliquerait le passage de la logique épistémique dans le rang des logiques déviantes, contrairement au système enrichi (basé sur des règles d'inférence classiques) de Hintikka. Dans ce cas de figure, la relation d'inclusion est maintenue entre les espaces logiquement et épistémiquement possibles puisque l'ensemble $\{\phi,\neg\phi\}$ n'est plus exclu : si l'on se permet quelques entorses à la condition de consistance (C.¬), auquel cas on tolère la conjonction d'une formule et de sa contradictoire au sein d'un espace logique, une 'incohérence' ou inconsistance n'est plus sanctionnée d'illogisme.

[Hintikka 1975] relève une suggestion semblable chez [Cresswell 1970], qui propose de redéfinir la négation afin de concilier inconsistance et vérité logique. La solution résiderait dans un changement des règles de la sémantique formelle. Cette solution peut séduire à la rigueur par son efficacité et sa simplicité : les vérités épistémiques (même inconsistantes) sont toutes des vérités logiques, mais au prix d'une redéfinition locale des règles d'inférence et d'un déplacement du problème. Hintikka accuse certaines théories d'être *ad hoc*, lorsqu'elles proposent de

> concevoir une sorte d'interprétation non-standard des constantes logiques. Cette démarche est très douteuse, toutefois. Le vrai problème a été créé par l'incapacité des gens à percevoir les conséquences logiques de ce qu'ils savent suffisamment loin. Ces conséquences logiques doivent être basées bien évidemment sur l'interprétation *classique* (standard) des connecteurs et quantificateurs. Une tentative d'interprétation non-standard est donc soit condamnée à être hors sujet, soit à *détruire* le problème au lieu de le *résoudre*.[178]

La paraconsistance est-elle la solution miracle, capable de faire admettre certains raisonnements inconsistants au sein de la classe des formules valides (et logiquement vraies) ? Nous ne choisirons pas cette voie de la révision des vérités logiques : la solution paraît arbitraire en tant que telle, puisqu'elle valide un nombre indéfini d'inconsistances sans préciser pour quels ensembles de formules cette clôture est admissible. Le problème concerne la frontière entre rationalité et irrationalité du point de vue logique : pour quels types de déductions l'agent est-il imputable de comportement irrationnel s'il ne perçoit pas leur relation de conséquence ? Dans un système paraconsistant, un ensemble $\{\phi,\neg\phi\}$ devient logiquement possible sans entraîner la vérité de n'importe quel autre énoncé ψ, mais la solution n'est que d'ordre technique ici et ne répond pas au véritable problème 'philosophique' de l'omniscience : comment fixer une ligne de démarcation entre les ensembles $\{\phi,\neg\phi\}$ épistémiquement possibles et impossibles ? Le véritable enjeu de l'omniscience logique n'est pas d'admettre l'inconsistance de n'importe quel système modèle; il est d'expliquer dans quelle mesure certains peuvent l'être et d'autres non. La solution tient dans la complexité de ϕ, comme nous allons le voir dans l'argumentation de Hintikka.

Changer les conditions de vérité dans un langage formel transforme certaines faussetés classiques en vérités déviantes et permet donc de considérer des inconsistances comme des mondes possibles épistémiquement et logiquement à la fois, mais à condition de redéfinir entre-temps la notion de monde *logiquement* possible. Or une telle solution technique élude selon nous le vrai problème philosophique qui sous-tend la question de l'omniscience logique (comment des inconsistances se produisent-elles dans nos raisonnements). Par conséquent, la limite entre les cas de consistance et les autres ne peut être fixée jusqu'ici que de façon arbitraire par le logicien paraconsistant, et cette solution est donc ajournée. Nous retiendrons l'objection présente à la paraconsistance comme une raison de rejeter la nature purement instrumentale de la logique : il ne suffit pas de changer l'interprétation des constantes logiques pour résoudre les difficultés posées en logique *philosophique*. Le même argument sera reproduit au sujet des critères logiques de rationalité (section 3.2).

Plutôt que de modifier les propriétés classiques de la vérité logique, supposons maintenant que l'on modifie le contenu de l'hypothèse C et les critères du concept de connaissance, en particulier la connaissance déductive. Pourquoi estime-t-on que certaines conséquences logiques de ce que l'on sait sont inaccessibles, ce qui nous conduit à affirmer $\neg B$? Dans le cas d'une tautologie de forme $p \Rightarrow (p \vee q)$, où p et q sont des formules atomiques et illustrent un cas de complexité minimale, la forme logique de la conséquence est suffisamment simple

[178] [Hintikka 1975], p. 477. Nous reviendrons à plusieurs reprises sur cette stratégie, consistant à changer les règles de transformation de la logique (l'interprétation des constantes) afin de régler les problèmes de logique épistémique. Cela revient à dire d'un problème qu'il n'existe pas en modifiant les données du problème.

pour paraître évidente et garantir le succès de la déduction par l'agent. Mais d'autres contre-exemples peuvent être invoqués pour rejeter l'hypothèse d'omniscience, lorsque la relation logique entre devient plus complexe. Quelle réponse y apporter pour sauver C ? Deux solutions favorables à l'omniscience logique vont être rejetées tour à tour dans ce qui suit.

Le première consiste à réinterpréter la notion métalogique de validité d'un théorème, ainsi que dans [Hintikka 1962], et déclarer que le paradoxe de l'omniscience logique n'en est pas un tant que l'on évalue dans un monde épistémiquement *idéal*, dans lequel les agents suivent suffisamment loin les conséquences de leurs connaissances. Or cette hypothèse a été rejetée en raison du présupposé d'idéalisation des agents (logiquement parfaits ou impeccables), alors que l'on attend avant tout des enseignements 'philosophiquement intéressants' de la part d'une logique philosophique telle que la logique épistémique. Pourquoi les agents réels ne sont-ils pas capables de tirer toutes les conséquences logiques de leurs connaissances ?

Une seconde solution consiste à déplacer le problème de la compétence logique sur le terrain de la compétence linguistique, c'est-à-dire de la *compréhension*. Dans l'exemple de conséquence logique ci-dessus, le cas de l'introduction de la disjonction, il y a échec déductif si l'agent ne comprend pas la signification de la disjonction $\vee$ ou s'il la confond avec la conjonction et croit que le conséquent est par conséquent faux. Mais cela revient à réduire le problème *logique* de l'omniscience à une question d'incompétence *linguistique* relativement au sens des signes logiques. Une telle solution est trop accidentelle pour être prise au sérieux, le problème de l'omniscience étant plus profond que celui d'une simple confusion conceptuelle.[179]

Le problème de fond ne provient pas de notre usage du symbolisme ; il se rapporte à une question de compétence logique. Il y a des choses que l'on sait sans les dire, d'autres que l'on pense ignorer tout en les connaissant ; tel semble être le cas pour les vérités logiques, d'après une doctrine logique répandue et à laquelle Hintikka va s'attaquer dans ce qui suit. L'hypothèse de la compréhension est plus intéressante si elle se définit en termes de disposition à agir : comprendre un énoncé auquel on croit, c'est être disposé à agir conformément à sa vérité. En vertu de cette disposition, il semble impossible de croire une contradiction *ex*plicite et d'agir conformément à ce genre d'énoncé : on peut dire une chose et son contraire (sa contradictoire), mais on ne peut faire une chose et ne pas la faire à la fois. Or la thèse de l'omniscience semble entraîner cette absurdité : si l'on admet un échec de la déduction $\phi \Rightarrow \psi$, cela veut dire que l'agent peut croire à la fois ϕ et la négation de sa conséquence : $\phi \wedge \neg\psi$. Or en vertu d'une règle standard de la logique classique propositionnelle, on obtient alors la conséquence absurde : $\phi \Rightarrow (\psi \wedge \neg\psi)$. Cette conséquence est inacceptable si l'on définit la compréhension en termes dispositionnels : personne n'est censé croire et agir conformément à une contradiction explicite, et un moyen de maintenir l'omniscience logique consistera à déclarer que la personne accusée d'inconsistance ne comprend pas en réalité ce qu'elle dit lorsqu'elle prétend ignorer que ψ. La paraconsistance éviterait certes cette preuve d'absurdité, puisqu'elle accrédite la contradiction explicite sans admettre n'importe quoi en conséquence. Mais encore une fois, la réponse n'est pas de type 'tout ou rien' : il ne s'agit pas d'imposer la connaissance logique ou, au contraire, de permettre l'ignorance logique pour n'importe quel type d'énoncé ϕ. La raison pour laquelle

[179] Des exemples d'incompétence linguistique de ce genre sont mentionnés dans [Lenzen 2004].

nous ignorons certaines conséquences logiques (mais pas toutes),[180] c'est que celles-ci peuvent ne pas être explicites en raison de la complexité de ϕ et ψ ; le comportement dépend alors de l'acuité logique [*logical acumen*] du sujet, c'est-à-dire de sa capacité à percevoir une relation inconsistante entre plusieurs prémisses. L'omniscience logique ne semble applicable qu'aux formules dont la forme logique est relativement simple et contient peu de constituants (de littéraux p ou $\neg$p). Peut-on limiter la vérité logique connaissable à une liste finie de formules, et à quel moment une vérité logique devient-elle inconnaissable ? Hintikka va restreindre ainsi le domaine des énoncés pour lesquels la clôture est satisfaite ; ce résultat entraîne une restriction du système modal minimal **K**, mais non une révision des propriétés générales de la logique modale normale. En d'autres termes, rejeter l'omniscience logique ne signifie pas admettre l'ignorance logique pour n'importe quelle formule, mais seulement pour certaines ; la complexité des relations logiques entre ces formules influe sur la constitution des mondes épistémiquement possibles, d'où le rejet de l'hypothèse C : $W_K \subseteq W_L$ n'est pas valable parce que les conséquences logiques de nos connaissances ne sont pas connues dans tout monde possible $w \in W$; il existe des mondes possibles w dans lesquels un agent peut supposer une vérité inconsistance dans l'ensemble de ses connaissances.

- Option AB$\neg$C

A la question : dans quelle *mesure* un agent est capable de (re)connaître des conséquences logiques et de satisfaire ainsi $W_K \subseteq W_L$, la réponse de Hintikka est que la classe des formules susceptibles de satisfaire cette inclusion correspond à la classe des tautologies dites *de surface*, c'est-à-dire les vérités logiques dont la structure inférentielle est suffisamment simple pour être perçue par n'importe quel agent. Or comment 'mesurer' ce qui sépare une forme logique simple d'une forme logique complexe ? Pour évaluer la compétence logique d'un agent, une logique philosophique doit permettre

> d'isoler les facteurs *objectifs*, structurels qui contribuent le plus à la difficulté pour percevoir les relations de conséquence logique.[181]

L'idée de base est que certaines conséquences logiques peuvent ne pas être perçues, lorsque la relation de conséquence $\phi \Rightarrow \psi$ est relativement compliquée ou qu'elle ne s'obtient qu'à partir d'une série d'étapes intermédiaires. Tenant compte des capacités déductives

---

[180] Précisons que la paraconsistance n'entraîne pas la possibilité d'admettre *n'importe quelle* ensemble contradictoire $\{\phi,\neg\phi\}$: le dialéthiste Graham Priest précise que la vérité d'une contradiction n'implique pas la vérité de toutes. L'erreur de démonstration du logicien classique (Aristote) serait d'imputer au paraconsistant (Héraclite) l'idée selon laquelle *toutes* les contradictions logiques sont vraies : $\models (\phi \wedge \neg\phi)$, alors qu'il ne le prétend que pour certaines : $\not\models \neg(\phi \wedge \neg\phi)$, et $\not\models (\phi \wedge \neg\phi)$. Mais notre problème demeure : quelles formes contradictoires peuvent être considérées comme vraies, et comment fixer la liste de formules susceptibles au contraire d'être ignorées par un agent épistémique ? Ce problème est-il d'ordre logique ? Le paraconsistant ne répond pas à ce problème mais fournit simplement des méthodes de preuve capables d'entériner l'hypothèse de l'ignorance logique.

[181] [Hintikka 1970b], p. 37. Les facteurs 'objectifs' ne seraient pas accidentels et relèveraient d'une raison computationnelle, relative aux capacité déductives générales des agents. La question à suivre est celle-ci : une logique épistémique est-elle capable de décrire ces limitations au sein d'un système formel ?

limitées des agents épistémiques, Hintikka va proposer des arguments issus de la science cognitive afin de résoudre le paradoxe de l'omniscience logique. Comment concevoir des mondes épistémiquement possibles mais logiquement impossibles ? Au niveau strictement formel, l'hypothèse des 'mondes possibles impossibles' répond à ce besoin mais reste à justifier dans ses applications. Considérons d'abord des solutions alternatives susceptibles de dissoudre le problème de l'omniscience, de le considérer comme un faux problème.

Une première solution consisterait à modifier la sémantique formelle, l'infrastructure sur laquelle les constantes logiques sont définies. Affaiblir la structure modale est concevable, puisque les systèmes de **S1** à **S3** sont plus faibles que **K** et ne contiennent pas la cause centrale de nos difficultés actuelles, à savoir la règle de nécessitation $\vDash\phi \Rightarrow \vDash\Box\phi$. Une telle solution paraît cependant trop radicale puisqu'elle invalide sans restriction la clôture épistémique, ce qui n'est pas souhaité ; de plus, affaiblir le système minimal **K** donne un système modal non-normal auquel ne correspond aucune structure : la solution de l'affaiblissement structurel revient donc à supprimer la sémantique relationnelle et limiterait par conséquent le traitement de l'omniscience logique à un niveau syntaxique. Il s'agit plutôt de distinguer les cas d'échec de clôture épistémique et les cas de succès, sans interdire la possibilité de connaître des conséquences logiques.

Une solution comparable à l'affaiblissement du système modal a été proposée également par [Girle 1969] : de même qu'en logique intuitionniste, sa solution à l'omniscience logique consisterait à restreindre les relations de dualité entre $\Box$ et $\Diamond$, notamment en modifiant la définition des conditions sémantiques (C.K$\neg$) et (C.P) ; si l'on rejette l'équivalence entre (C.$\neg$K) et (C.P), on rejette par là-même la règle de nécessitation et l'on remplace le système **K** par sa version affaiblie **S.05**. Le remède semble pire que le mal , encore une fois : si l'on perd la dualité des opérateurs modaux, on interdit également toute généralisation des relations modales. Toute théorie doit-elle être amendée dès qu'un cas particulier semble la falsifier ? [MacLane 1974] avait soulevé cette objection contre les critiques de [Hocutt 1972], mais la situation est telle que les dualités modales admises dans [Hintikka 1962] ont des conséquences intenables pour les ensembles épistémiques. Pour cette raison, un affaiblissement des propriétés structurelles du système **K** semble s'imposer malgré tout. Mais dans quelles proportions, et par quels moyens ?

Une condition proposée par [Girle 1969] consiste à rejeter la relation d'équivalence modale entre $\neg K\phi$ et $P\neg\phi$. Il y aurait certaines formules dont nous ignorons la vérité mais dont la négation est cependant impossible logiquement : $\neg P\neg\phi \wedge \neg K\phi$. Cela est impossible dans [Hintikka 1962], où l'agent ne peut pas ignorer des vérités logiques (celles dont la négation est logiquement impossible) : $\neg P\neg\phi \rightarrow K\phi$. Plutôt que d'affaiblir **K** et de choisir un système de type **S3**, ce qui nous empêcherait de construire un modèle épistémique correspondant, l'astuce de [Girle 1969] consiste à réviser certaines relations de dualité modale :

> [$\neg K_a\neg p$] est manifestement équivalente à '*a* ne sait pas non-p'. Un exemple de ceci serait '*a* ne sait pas $\neg(p \wedge \neg p)$'. Nul doute que nous pourrions affirmer ceci mais nier qu'il soit *possible*

pour autant que *a* sache que $p \wedge \neg p$. (...) En d'autres termes, si *a* sait que $\neg p$ alors il n'est pas possible, pour autant que *a* sache, que p. Mais la converse n'est pas vraie.[182]

La différence est faite ici entre ce qui est épistémiquement possible et ce qui est logiquement possible : certaines formules sont logiquement impossibles mais possibles épistémiquement. L'exemple choisi par l'auteur paraît curieux, toutefois, puisqu'il consiste à dire d'un agent qu'il peut ignorer la fausseté logique de '$p \wedge \neg p$' bien que '$p \wedge \neg p$' soit logiquement impossible. Puis-je sérieusement ne pas savoir que l'énoncé 'il pleut et il ne pleut pas' est indéfendable, à moins d'entrer dans des arguties de type paraconsistant ? Ce dernier type de logicien aurait plutôt tendance à admettre l'impossibilité de connaître certaines inconsistances logiques lorsque leur relation contradictoire n'est pas immédiatement perceptible, contrairement à la contradiction explicite ci-dessus. Dans ce cas, la raison de l'ignorance tient encore une fois à la complexité structurelle de la formule contradictoire. La méthode de [Girle 1969] anticipe la sémantique des mondes possibles impossibles, puisqu'il inaugure la distinction entre deux classes de mondes épistémiques : des mondes épistémiquement et logiquement possibles, satisfaisant la relation $W_K \subseteq W_L$; des mondes épistémiquement possibles mais logiquement impossibles, rejetant la relation d'inclusion ci-dessus et permettant d'invalider l'inférence $\neg P\phi \rightarrow K\neg\phi$.[183] Mais la solution n'est encore que formelle et ne répond pas au problème philosophique de la *raison* pour laquelle certaines impossibilités logiques sont ignorées.

D'autres interprétations du problème existent mais ne semblent pas donner une meilleure appréciation du problème de l'omniscience logique, qui est avant tout un problème de limitation des ressources déductives. On peut citer la logique de la conscience de [Fagin & Halpern 1988], dans laquelle la clôture épistémique n'est garantie que pour les formules gouvernés par un opérateur de *conscience* A : l'agent peut déduire la conséquence logique ψ de sa connaissance ϕ s'il a conscience de la relation logique entre ϕ et ψ : $A(\phi \Rightarrow \psi)$. Si l'on entend par conscience de ϕ la connaissance par l'agent de la vérité de ϕ, alors la solution est soit circulaire soit trop faible. Elle est circulaire, si elle présuppose que les mondes logiquement possibles sont limités aux relations logiques dont l'agent a conscience, puisque ces mondes sont épistémiquement possibles. Or l'enjeu du problème de l'omniscience n'est pas de confondre ces deux catégories de mondes mais de montrer au contraire en quoi ils diffèrent. Elle est trop faible, parce que la condition de clôture proposée ici s'apparente davantage à la règle de distribution modale qu'à la règle de nécessitation : $A\phi$ et $A(\phi \Rightarrow \psi)$ impliquent certes $A\phi$, mais le problème de l'omniscience concerne les cas de relation logique dont l'agent n'a pas forcément conscience : $\neg A(\phi \Rightarrow \psi)$. Bien que ce soit cette voie (de la conscience) que nous associerons plus tard au concept de connaissance (voir section 2.2.4.2), elle ne résout pas pour autant la difficulté actuelle concernant la compétence déductive des agents. [Meyer 2001] cite d'autres exemples similaires parmi lesquels le modèle des faisceaux de Chellas, dans lequel la croyance portant une contradiction explicite n'est plus logiquement fausse.[184] Ces révisions techniques (changer l'interprétation des constantes, ou ajouter

[182] [Girle 1969], p. 368. On a donc encore $(K\neg\phi \rightarrow \neg P\phi)$, mais $(\neg P\phi \rightarrow K\neg\phi)$ est rejetée. Donc $(\neg P\phi \wedge \neg K\neg\phi)$ doit être considéré comme défendable.
[183] [Girle 1969] établit respectivement cette distinction en termes d'alternative et d'alternative°.
[184] Voir [Meyer 2001], en particulier pp. 194-5.

d'autres contraintes aux inférences) obtiennent les résultats escomptés, à savoir le rejet de l'hypothèse C et la confirmation de l'hypothèse B ; mais elles ne rendent pas suffisamment compte des circonstances dans lesquelles les ressources déductives des agents humains sont trop limitées pour satisfaire la condition de clôture épistémique.

Il semble y avoir un malentendu d'ordre heuristique, concernant les problèmes de logique philosophique et les solutions proposées. Qu'est-ce qu'une solution à un problème de logique *philosophique*, porté sur la signification de certains concepts ordinaires ? Si une logique 'philosophique' n'est pas un pur langage formel mais un modèle explicatif, elle ne sert pas à imposer un usage normatif de ses axiomes mais à expliquer l'origine de leurs possibles défaillances dans la pratique. En conséquence, l'intérêt du calcul logique concerne surtout les logiques d'aspiration technique ou concernées par la programmation informatique, telles que les recense [Meyer 2001]. Mais elles n'expriment pas de véritables *raisons* aux problèmes posés par une logique philosophique, qui est plutôt du ressort du philosophe et nécessite une justification du formalisme employé. En conséquence, maintenir une liste d'inférences en l'état n'a pas plus d'intérêt que de les affaiblir de façon *ad hoc* : dans les deux cas, nous rejetons la logique pour ce qu'elle a tantôt de trop normatif, tantôt de trop descriptif. Nous reviendrons en fin de section sur ce problème qui concerne l'*adéquation* d'une solution à un problème posé en logique philosophique.

Comment fixer alors la limite entre le cas de succès et d'insuccès de la clôture épistémique ? Au premier abord, les raisons d'échecs paraissent trop relatives pour être déterminées : elles semblent dépendre aussi bien de la complexité structurelle des inférences que des capacités déductives de l'agent.[185] Mais [Hintikka 1970b] a prétendu plus haut que certains critères 'objectifs' sont disponibles pour mesurer la compétence déductive d'un agent épistémique, et cette mesure doit permettre de maintenir sous certaines conditions la caractérisation **KT4** de la logique modale épistémique. En résumé, la solution à venir consiste à montrer que l'échec de la clôture épistémique a un intérêt philosophique plus grand que son succès : d'après Hintikka, certaines conséquences logiques peuvent apporter une *information* et la conclusion d'une tautologie n'est pas toujours contenue dans ses prémisses. La nature de cette information acquise est-elle logique, donc objective, ou psychologique, donc subjective ? Pour illustrer l'intérêt philosophique (heuristique) de son système, Hintikka propose de faire remonter l'origine du problème de la clôture au *positivisme logique*, car c'est de ce courant que provient l'idée selon laquelle une tautologie n'apporterait aucune information. Les conséquences de ce point de vue ont un lien direct avec le problème de l'omniscience logique, puisqu'il impliquait tout à l'heure l'option AC¬B. Si la conséquence logique d'une connaissance particulière ajoute une information à celle-ci, il est possible que l'agent épistémique ignore cette conséquence et, contrairement à ce qui a été dit jusqu'à maintenant, ne la connaisse pas même s'il en connaît la condition logique. Le résultat sera à nouveau un rejet de l'inclusion $W_K \subseteq W_L$, donc une validation de l'hypothèse B. Le prix à payer pour cette défense semble lourd, toutefois, puisqu'il conduit Hintikka à redéfinir en profondeur la

[185] [Hintikka 1962] constate cette relativité apparente des capacités déductives d'un agent, selon que la relation de conséquence entre antécédent et conséquent est 'proche' ou 'lointaine' (de surface) : "Les critères de normalité présupposés ici ne sont pas forcément absolus. Lorsqu'un mathématicien écrit que quelque chose est 'facile à percevoir' (témoin le fameux 'Il est aisé à voir' de Laplace), il y a de fortes chances que son idée ne soit pas très évidente à percevoir même pour un profane perspicace" (note 6 p. 35).

notion de tautologie et à réviser la théorie probabiliste de l'information de Carnap : une vérité logique ne serait pas toujours une information dont la somme est nulle.

D'après le positivisme logique, au contraire, une *tautologie* n'a aucun contenu factuel et n'apporte aucune information nouvelle. Toutes les tautologies disent la même chose, c'est-à-dire rien : dans le *Tractatus*, Wittgenstein déclara ainsi que "La tautologie résulte de toutes les propositions : elle ne dit *rien*";[186] Carnap disait de même que la tautologie constitue une somme d'information nulle parce qu'elle ne nous apprend rien sur l'état du monde. En somme, la conséquence est contenue dans la prémisse "tout comme les plantes sont contenues dans leurs graines, et non comme les poutres sont contenues dans une maison".[187] Les positivistes logiques concluaient de cette définition de la tautologie que personne ne peut ignorer la conséquence logique d'une prémisse connue, dans la mesure où toute conséquence est déjà contenue dans la prémisse. C'est cette définition de la tautologie comme proposition analytique à information nulle que Hintikka va remettre en cause. Pour satisfaire l'idée selon laquelle une vérité logique apporte une certaine forme de connaissance supplémentaire, l'auteur propose d'

> esquisser un sens de l'information dans lequel le raisonnement déductif non-trivial augmente bel et bien notre information.[188]

Toute déduction ne serait donc pas triviale et pourrait transmettre quelque information qui ne serait pas donnée *a priori*, ou n'équivaudrait pas à ce que l'on qualifie ordinairement de vérité analytique. A travers le problème de l'omniscience logique, toute une liste de notions philosophiques liées appellent ainsi à être réévaluées : *a priori*, *a posteriori*, analytique, synthétique, tautologie.

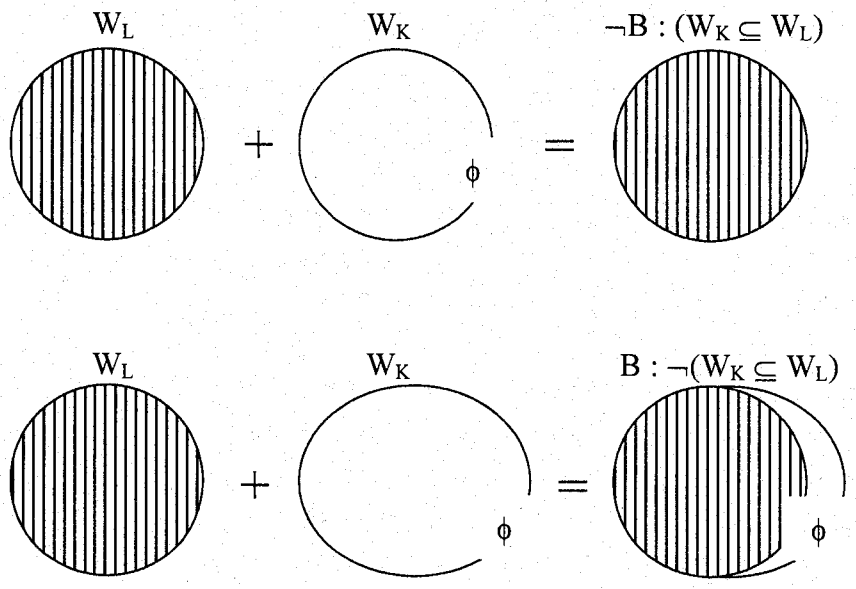
En référence à la théorie de l'information proposée par John von Neumann,[189] Hintikka va trouver dans la question de l'omniscience logique un exemple d'éclairage philosophique que la logique épistémique peut apporter à la notion d'*analyticité* ou de *tautologie*. Pour la résumer, l'idée décrite entre autres dans [Hintikka 1965] est que la notion d'analyticité ne s'applique pas indifféremment à l'ensemble des vérités logiques, lorsque l'on entend la notion d'analyse comme l'application de méthodes purement analytiques en vue de démontrer une tautologie. La distinction évoquée plus haut entre déductions triviales et déductions non-triviales correspond à la distinction entre des inférences *de surface* et des inférences *profondes*, dans lesquelles la relation de conséquence est plus ou moins complexe. Plus la chaîne déductive est longue entre une prémisse et la conclusion, ou entre une condition et sa conséquence, plus la compétence logique est mise à l'épreuve, et c'est entre les deux types d'information ci-dessus que se situe la réponse au problème de la connaissance logique.

Pour le résoudre, Hintikka n'a pas seulement proposé une autre interprétation formelle : la sémantique des 'mondes possibles impossibles', mais il a également apporté une explication (une interprétation informelle) de cette sémantique. D'un point de vue strictement formel, le procédé consiste à ajouter au triplet de la logique modale normale : $<v,W,R>$ un quatrième

[186] [Wittgenstein 1969], § 5.142.
[187] G. Frege: *Grundlagen der Arithmetik*, §88.
[188] [Hintikka 1970e], p 136.
[189] [Hintikka 1970b], p. 37.

élément non-normal W*, la classe des mondes impossibles. On obtient un quadruplet <v,W,W*,R>, dans lequel la valeur de vérité des formules complexes est considérée comme indéterminée dans W'. L'intérêt de la manœuvre est d'admettre la possibilité de vérité de certaines ensembles inconsistants (impossibilités logiques) dans un monde non-normal w*; si ce monde est accessible à partir des mondes normaux w, alors une tautologie n'est plus toujours connue puisqu'elle devient fausse dans un monde épistémiquement accessible (le monde w').[190]

Dans ce système modèle, l'ensemble W_K+W_K* des mondes épistémiquement possibles n'est plus inclus dans l'ensemble des mondes logiquement possibles, donc W_K et W_L deviennent des ensembles parfois disjoints. Pour illustrer ceci, la différence entre les modèles épistémiques normaux et non-normaux donne une situation suivante où une formule épistémiquement possible (incluse dans W_K) est logiquement impossible (n'est pas incluse dans W_L) :



Pour expliquer cet échec, Hintikka associe l'idée de mondes possibles impossibles à l'idée selon laquelle certains inférences pourraient ne pas être évaluées correctement en raison de leur complexité : des erreurs imperceptibles peuvent être commises lors de leur démonstration et empêcher l'agent de percevoir la vérité logique qu'elles contiennent. La solution rejette clairement la définition analytique des positivistes logiques : la conséquence logique d'une formule n'est pas 'contenue' dans celle-ci au sens où sa connaissance serait 'dans la tête' de l'agent. Pour donner un sens aux mondes impossibles, Hintikka s'inspire du modèle des urnes

[190] Plus précisément, on obtient les valuations suivantes pour l'opérateur K :
$v(K_a p,w) = 1$ si et seulement si $v(p,w^*) = 1$, pour tout $w^* \in W \cup W'$ tel que wRw^*,
$v(K_a p,w) = 0$ si et seulement si $v(p,w^*) = 0$ pour quelque $w^* \in W \cup W'$ tel que wRw^*.
L'explication de la non-omniscience logique est la suivante :
Si $\models \phi$, alors $v(\phi,w) = 1$ pour tout $w \in W$, mais non pour tout $w \in W \cup W'$, car il peut y avoir des w' où $v(\phi,w') \neq 1$.
Donc $\models \phi$ n'implique pas toujours $K\phi$, puisqu'il est possible d'avoir un w* dans lequel $v(\phi) = 1$ et $v(K\phi) \neq 1$. Une fois cette sémantique définie par le logicien, il incombe au philosophe la tâche d'expliquer la raison de l'indétermination de $v(\phi)$ dans w'.

variables de [Rantala 1982] où l'urne correspond à un modèle épistémique, c'est-à-dire un ensemble de mondes : si dans la structure contenant les mondes possibles impossibles la valeur des tautologie est indéterminée, c'est parce que

> ce sont simplement des mondes *changeants*. Ou pour le dire plus précisément, ce sont des mondes *insidieusement* changeants : ce sont des modèles dont le domaine peut changer pendant que nous l'examinons.[191]

L'évaluation d'une formule est comparée à un tirage de loterie, dans laquelle chaque boule symbolise un constituant (un littéral p ou ¬p dans une formule ϕ quelconque) et où chaque nouveau tirage propose une nouvelle combinaison de constituants. Le changement insidieux qui se produit dans les boîtes symbolise un changement dans les mondes épistémiques :

> Le fait n'est pas seulement que la composition de la boîte peut changer, mais qu'elle change entre nos tirages successifs des balles dans l'urne. Rantala utilise le terme judicieux de 'modèle d'urnes' pour ces nouveaux modèles des énoncés du premier ordre.[192]

L'information obtenue dépend du type de tirage effectué : si l'ensemble des boules (l'interprétation des mondes possibles) est constant, le modèle d'urnes est dit *invariant* et correspond à la sémantique modale normale ; s'il ne l'est pas, le modèle est dit variant et conduit à l'hypothèse B. D'après Hintikka, l'information apportée par un énoncé correspond à l'introduction d'un nouvel individu entre un premier et un second tirage, auquel cas les deux urnes sélectionnées ne sont pas invariantes.

Si l'on admet cette analogie et le diagnostic qu'en tire Hintikka, le problème de l'omniscience logique exige selon lui un traitement en termes de logique des prédicats et non des propositions, puisque celle-ci n'analyse pas suffisamment afin de révéler l'origine du problème. Dans la partie consacrée à la présentation de la logique épistémique, nous avons vu que toute formule peut être analysée comme une somme de possibilités d'états de choses ou de mondes possibles distincts, et l'*information* apportée par un énoncé dépend du nombre de ces possibilités. Les formes normales disjonctives[193] permettent d'exprimer la totalité des combinaisons possibles d'états de choses (de constituants d'une forme normale) en termes de conjonction de littéraux (de forme p ou ¬p, où p est une formule atomique). On dit des tautologies qu'elles représentent une information nulle parce qu'elles épuisent toutes les possibilités : tous les membres des formes normales qui les expriment sont *trivialement* vrais. Lorsque l'analyse passe au niveau de la quantification, l'analyse logique des formules est rendue plus complexe par la relation d'enchâssement entre les différents quantificateurs qui relient les individus des formes normales : chaque individu est en relation avec d'autres individus ou avec lui-même dans un prédicat à plusieurs places, et le degré de complexité de l'analyse varie en fonction de ce paramètre. La complexité des inférences logiques est directement liée à la structure logique de leurs termes, donc le problème de l'omniscience

[191] [Hintikka 1975a], p. 478.
[192] [Hintikka 1975a], p. 478.
[193] Le traitement du problème de l'omniscience en termes de formes normales et de constituants complexes montre une certaine continuité dans les travaux de Hintikka, dont la thèse avait porté sur le thème des formes normales distributives. Celles-ci servent à exprimer ici le problème de l'ignorance des vérités logiques.

logique est indirectement lié à la question de l'analyse des formules en termes de forme normale.

Hintikka appelle *profondeur* [*depth*] quantificationnelle d'une formule le nombre des variables d'individus liées à ses quantificateurs : par exemple, la formule $(\forall x)(\exists y)S(x,y)$ a une profondeur de degré 2 puisqu'elle contient deux variables x et y qui satisfont la relation S par la paire ordonnée $\langle x,y \rangle$. Pour représenter l'analyse d'une formule quantifiée du premier ordre, on construit un arbre inversé composé de branches et de nœuds (feuilles) aux extrémités de ces branches. Chaque nœud représente un individu et chaque branche représente les relations qu'un individu entretient avec d'autres individus. Sans entrer plus loin dans les détails, Hintikka exploite cette méthode de démonstration pour montrer que certaines tautologies ne sont pas 'perçues' lorsque, dans une inférence logique, l'antécédent et le conséquent n'ont pas le même degré de profondeur. La conclusion donnée par exemple dans [Hintikka 1970b] est que toute vérité logique de type $\phi \Rightarrow \psi$ est connue si et seulement si ϕ et ψ sont de même profondeur ; si leur profondeur est différente, alors il est possible que l'agent connaisse ϕ sans connaître ψ.

Ces résultats ont leur importance en philosophie de la logique, puisqu'ils relancent la discussion sur la signification de plusieurs notions telles que l'analyticité ou sur la distinction entre jugements logiques et mathématiques. Si un antécédent et un conséquent sont de profondeur différente, c'est qu'un nouvel individu a été introduit entre les deux. Or [Hintikka 1965] n'ignore pas que cette introduction correspond à la définition kantienne de la *synthèse* mathématique, où le sujet construit une démonstration par l'introduction de nouveaux individus. En quoi cette procédure est-elle logique, et n'y a-t-il pas dans la solution de Hintikka une confusion entre ce qui, depuis Kant et l'expression de jugements synthétiques *a priori*, distingue les énoncés logiques (déductifs) et les énoncés mathématiques (intuitifs) ? [Hintikka 1970b] explique notamment que les tautologies sont 'non-triviales' lorsque leur démonstration passe par une démonstration similaire à la construction par synthèse : c'est le processus d'ecthèse, qui démontre une formule générale par le biais d'un cas particulier. En somme, une tautologie non-triviale est une formule qui est vraie dans tous les mondes possibles mais dont certains constituants échappent aux procédures de démonstration purement analytiques (qui n'invoquent pas d'instance particulière au cours de la démonstration). Une analyse précise du problème nécessiterait des descriptions supplémentaires sur ces procédures (conditions de troncature, trivialité et non-trivialité, etc.), mais nous retiendrons surtout que la défense de l'hypothèse B passe par une reconsidération en profondeur du processus d'analyse logique.

A travers le problème de l'omniscience logique, Hintikka étend ainsi le domaine d'application des méthodes de logique et prétend montrer que l'intuition kantienne n'est rien d'autre qu'un processus d'instanciation dans un modèle : le sujet pensant construit un objet lorsque le sujet locuteur introduit un exemple dans une enquête. Cette analogie sera poursuivie dans sa sémantique des jeux, et l'on retrouve cette filiation entre mondes impossibles et profondeur quantificationnelle dans [Hintikka 1975a]. La solution de Hintikka repose sur une comparaison 'objective' du mécanisme du système nerveux central à une procédure déductive dont les opérations sont descriptibles en termes d'arbres, d'enchâssement, de profondeur. L'information transmise lors d'une démonstration repose sur

cet algorithme inspiré par von Neumann (recensé dans [Hintikka 1970b]. La logique des prédicats apparaît donc au cœur de la solution au problème de l'omniscience logique : il ne s'agit pas de changer les règles d'inférence logique (l'interprétation des constantes dans une sémantique formelle), mais de décrire les conditions dans lesquelles ces règles sont appliquées à des formules quantifiées.[194] Pour le dire autrement : ce n'est pas la logique classique qui est à réviser, c'est la condition d'application des inférences de la logique classique qui est à préciser.

En conclusion, un problème de logique philosophique a produit ici des résultats en philosophie de la logique, puisqu'elle ouvre des pistes en théorie de l'information et exige une réforme de l'analyse logique : l'analyse des formules polyadiques (les relations à plus d'une place) crée des difficultés et justifie selon Hintikka la distinction entre inconsistance triviale et inconsistance non-triviale : de façon surprenante, des formules tautologiques peuvent amener le sujet à en tirer des inconsistances lorsqu'elles ne sont pas simplement d'ordre monadique, mais ces inconsistances peuvent être éliminées par des clauses intermédiaires qui montreraient dès lors qu'une vérité logique n'est pas simplement analytique. La réponse à la question de [Hintikka 1965] : les vérités logiques sont-elles analytiques, conditionne ainsi la réponse de l'auteur au problème de l'omniscience logique. La pertinence de ce résultat correspond à l'idée que Hintikka se fait de la logique : un *modèle explicatif*, dont l'intérêt repose avant tout sur l'audace (au sens de Popper) des arguments et le champ d'application de leurs résultats. Dans le cas de l'omniscience logique, l'intelligence artificielle, la théorie de l'information et la logique épistémique se recoupent au sein d'un seul et unique cadre d'analyse : la logique classique, dont l'analyse des énoncés apporte un moyen de rejeter l'inclusion $W_K \subseteq W_L$.

Notons que le traitement ci-dessus de la non-omniscience donne une orientation computationnelle à l'interprétation de la notion de connaissance : celle-ci est définie par Hintikka comme un stock d'information déterminé par une série de calculs sur des formules, et la seule différence avec le modèle classique réside dans le nombre des modèles parcourus (le nombre des mondes imaginés par l'agent). Par la même occasion, le paradoxe de l'omniscience logique apporte ici une forme de démenti à l'*instrumentalisme logique*, en vertu duquel un changement d'interprétation des mots logiques est préférable dès qu'il permet de confirmer ou d'infirmer une formule. Or l'usage d'une sémantique des mondes possibles impossibles ne suffirait pas pour expliquer à elle seule l'ignorance de certaines inférences logiques. Ce que nous voulons dire par là, c'est qu'une solution à un problème de 'logique philosophique' proprement dit ne consiste pas à proposer une révision *ad hoc* de la sémantique formelle; la paraconsistance ou les différents systèmes non-classiques permettent de falsifier (construire un contre-modèle pour) une propriété indésirable telle que l'omniscience mais ne sont pas capables d'expliquer son origine, que Hintikka a située au

[194] On trouve dans [Hintikka 1965] d'autres hypothèses susceptibles de confirmer sa position : si la méthode de démonstration logique a été assimilée en général à une méthode *analytique*, dans laquelle aucune instance d'individu particulier n'est introduite au cours de la démonstration, c'est parce que la notion kantienne d''analytique' assimilait logique et syllogistique. Or la syllogistique se compose de relations monadiques qui n'exigent pas l'élimination de constituants *inconsistants*, contrairement aux relations polyadiques. La reconstruction de la procédure de démonstration produit non seulement une réponse plausible à l'omniscience logique : il y a une frontière objective entre les échecs et les succès d'une clôture épistémique ; mais elle permet de rediscuter également des distinctions léguées par la philosophie : l'analogie entre analytique et logique, d'une part, et celle entre synthétique et mathématique, d'autre part, est remise en cause.

cœur de la procédure de démonstration. Autrement dit, la solution ne serait pas dans la construction d'un modèle déviant mais dans l'étude des méthodes de démonstration que nous appliquons au sein du modèle classique. Changer les connecteurs booléens ou la méthode de valuation est donc hors sujet ou doit s'accompagner tout au moins de raisons préalables.

Faut-il conclure de la non-omniscience logique des agents épistémiques que la logique modale épistémique repose sur une infrastructure inadéquate, le système **K**, et que d'autres systèmes non-normaux doivent être envisagés pour obtenir une logique épistémique 'réaliste', c'est-à-dire conforme aux capacités déductives effectives des agents humains ? Cette voie est suivie de nos jours, mais ce ne sera toutefois pas la nôtre dans ce travail. Comment éviter le problème de la compétence déductive dès que l'on s'intéresse à l'analyse logique de la notion de connaissance ? Tout dépend de la signification de base proposée pour son opérateur, et le modèle explicatif que nous allons adopter par la suite va permettre d'éluder en quelque sorte le problème de la déduction effective en logique. Selon nous, une autre interprétation de la notion de connaissance peut être proposée et qui ne fait pas appel à la règle de nécessitation, esquivant de ce fait le problème de l'omniscience logique. Comment se détourner des problèmes de compétence déductive, au sein d'une logique épistémique ? Pour ce faire, nous allons examiner dans ce qui suit les arguments proposés par Hintikka pour justifier d'autres types d'inférences épistémiques, parmi lesquels les axiomes d'introspection. Nous verrons que ses arguments ne sont pas liés au problème des compétences déductives et de l'omniscience logique. L'analyse de ce dernier thème, majeur en logique épistémique, a produit notamment une série de programmes de recherche en philosophie de la logique : l'anti-réalisme radical et la question de la faisabilité en logique, par exemple, prolongent le thème de l'omniscience logique et sont consacrés au thème de la rationalité limitée.[195] Mais ce n'est pas cette orientation qui va occuper le reste de ce travail. Une autre interprétation des attitudes épistémiques va être proposée, dans laquelle le sens donné aux opérateurs modaux K et B nous permettra de supprimer la règle de nécessitation et, ainsi, d'éluder le problème qui a occupé toute cette section.

La logique modale épistémique à venir sera davantage portée sur l'aspect *linguistique* ou illocutoire des modalités et tentera de trouver une signification de base pour les modalités épistémiques. Il s'agira avant tout de répondre à la question suivante : quel système de logique modale est susceptible de saisir les propriétés formelles principales des notions de connaissance et de croyance ? Contre le choix d'une révision des constantes logiques, et contre l'idée d'un affaiblissement structurel des propriétés de K et B (en vue d'éviter le problème récurrent de l'omniscience), nous plaiderons pour une logique philosophique de type *explicatif* mais non descriptif. [Rantala 1979] a constaté qu'un traitement de l'omniscience logique n'était pas sans effet sur l'idée de logique en général, considérée d'habitude comme une norme de la pensée :

> Plusieurs solutions, basées sur diverses notions de mondes non-normaux, ont été proposées dans la littérature pour le problème de savoir comment éviter les engagements sur l'omniscience logique et des difficultés similaires causées par les logiques modales faibles en

[195] Voir notamment [Dubucs 1991,1998] et [Dubucs & Marion 2003]. L'accent est mis dans ces travaux sur les logiques substructurelles et la description des ressources limitées en termes de logique linéaire.

général (...) Les solutions proposées qui sont présentées en logique épistémique sont basées sur des considérations légitimes, aussi bien pragmatiques et herméneutiques que logiques. Cela montre, il me semble, que les développements en logique épistémique et dans les logiques des attitudes propositionnelles en général sont souvent motivées par des tentatives chez les philosophes de considérer ces logiques comme des théories explicatives et descriptives des attitudes, plutôt que comme des logiques qui idéalisent en un sens prescriptif le thème en question.[196]

Bien que nous dissocions pour notre part les aspects descriptif et explicatif d'une logique appliquée, Rantala insiste à juste titre sur le fait qu'une utilisation *descriptive* de la logique philosophique a pour effet indésirable de multiplier les systèmes selon les besoins, c'est-à-dire de proposer une logique pour chaque caractérisation possible d'un concept philosophique. Le résultat est une somme de systèmes formels qui risquent de nous faire perdre de vue la généralité de la discipline logique, contrairement au critère de généralité requis dans les sciences formelles. Par opposition à ce morcellement des systèmes de logique appliquée, une utilisation *prescriptive* de la logique tentera de maintenir une seule interprétation caractéristique des constantes logiques. [Hintikka 1962] est un exemple de logique prescriptive, malgré le défaut structurel imputé au système **K**.

Notre analyse des modalités épistémiques tentera ainsi d'être à la fois explicative et prescriptive : caractéristique des significations de base de la connaissance et de la croyance, d'une part ; applicable à un maximum de cas d'usage des attitudes épistémiques, d'autre part. Comment obtenir ainsi une logique à la fois réaliste et générale pour les modalités épistémiques ? La section qui suit propose une réponse à cette question d'adéquation informelle, en orientant l'analyse de 'K' et 'B' vers une analyse illocutoire des formules comme actes de discours. Cette interprétation particulière des attitudes épistémiques, qui reste à justifier, permettra ainsi d'éliminer l'effet indésirable de l'omniscience logique tout en préservant la simplicité de structure des logiques modales normales. Dans le même registre que l'opérateur de conscience de [Fagin & Halpern 1988], la remarque suivante de [Vernant 1997] préfigure notre analyse des modalités épistémiques comme actes d'assertion et, surtout, confirme l'idée que le problème de l'omniscience logique peut être évacué sous cette interprétation :

D'autre part, la considération du locuteur comme auteur de l'acte d'assertion conduit à limiter le jeu des inférences propositionnelles : il n'assume pas la totalité des conséquences logiques de ses assertions et l'ensemble de ses assertions ne doit pas être logiquement consistant. La logique illocutoire se déploie en un système d'*inférences pratiques* ne requérant qu'une cohérence minimale.[197]

Pour montrer qu'une explication des opérateurs K et B en termes d'actes de discours pourrait servir de signification de base, commençons d'abord par examiner les arguments de [Hintikka 1962] et voyons dans quelle mesure le système **S4** conduit à une analyse illocutoire des modalités.

---

[196] [Rantala 1979], p. 196.
[197] [Vernant 1997], p. 40.

2.2.4.2 L'introspection positive (et négative)

On trouve deux sortes d'explication des inférences monomodales de type K ⇒ KK ou B ⇒ BB. Lorsque leur explication repose sur des notions de conscience, d'esprit ou d'introspection du sujet connaissant, elle est de nature *psychologique* ; lorsqu'elle fait appel aux conditions de justification de la connaissance, elle est de nature *cognitive*. En raison de la connotation psychologique des verbes d'attitude, les axiomes **4** et **5** ont été qualifiés tous deux de thèses d'*introspection* : positive pour la première, négative pour la seconde. La thèse d'introspection positive, exprimée par la formule $K_a\phi \rightarrow K_aK_a\phi$, signifie que le fait même de connaître fait partie des connaissances d'un agent épistémique : il a accès à tous ses contenus de connaissance, y compris le fait qu'il sache ce qu'il sait. La thèse d'introspection négative, $\neg K_ap \rightarrow K_a\neg K_ap$, signifie que les ignorances de *a* font partie de ses connaissances : il ne sait pas tout, mais il sait tout ce qu'il ne sait pas. Dans les deux cas, les attitudes épistémiques exprimées dans le conséquent des schémas modaux **4** et **5** sont des attitudes portées sur des attitudes : $K(K\phi)$ et $B(B\phi)$ ne portent pas sur un état de choses ϕ mais sur une attitude K ou B à l'égard d'un état de choses ϕ ; autrement dit, les thèses d'introspection font des attitudes épistémiques un objet d'attitude à part entière. Alors que le schéma **4** est considéré comme caractéristique des attitudes de connaissance et de croyance dans [Hintikka 1962], le schéma **5** est considéré comme contre-intuitif : il sous-entend que les agents ont connaissance des limites de son savoir, ce qui est absurde ou témoigne en tout cas d'une forme de sagesse peu courante.[198]

Pourquoi associer la psychologie aux modalités épistémiques itérées, alors que l'analyse de K insiste en général sur une condition de justification qui distingue la connaissance de la simple croyance ? Pour reprendre l'expression de [de Sousa 1970], les croyances et connaissances itérées sont des modalités de *second ordre*, c'est-à-dire des attitudes dont l'objet n'est pas un simple énoncé factuel mais une attitude : pour cette raison, l'itération épistémique suggère une capacité de l'agent à prendre ses propres attitudes pour objets de connaissance ou de croyance. L'interprétation de l'itération des opérateurs en termes de conscience de soi ou d'introspection repose ainsi sur cette modalité de second ordre, mais elle ne sera cependant pas décisive dans l'argumentation de Hintikka. Bien que [Hintikka 1970a] fasse référence par ailleurs à un argument de la conscience transcendantale de Sartre pour étayer la validité du schéma **4** (dont la version épistémique est appelée 'thèse KK'), une telle référence ne peut avoir d'intérêt qu'historique et ne montre en rien que cette thèse réponde au 'répertoire conceptuel' de la notion de connaissance, ni qu'un tel usage de la notion de connaissance corresponde à nos 'besoins quotidiens du langage ordinaire', pour plagier Hintikka.[199] Telle est pourtant la condition à remplir pour intégrer **S4** parmi les structures

[198] C'est pourquoi l'introspection négative est souvent associée à la personne de Socrate et qualifiée également d'axiome de sagesse : 'si je ne sais pas certaines choses, au moins je sais que je ne les sais pas', marque de sagesse qui, hormis le philosophe, n'est pas caractéristique des agents humains.

[199] "A la fois dans 'La Transcendance de l'Ego' et dans L'Etre et le Néant, Jean-Paul Sartre déclare que toute conscience a conscience d'elle-même dans la mesure où elle a conscience de tout objet transcendant (au sens de Husserl)". ([Hintikka 1970a], p. 160). Ainsi que dans [Hintikka 1962], où l'auteur cite au Chapitre Cinq les références de Platon, Aristote, Saint Thomas d'Aquin, Spinoza et Schopenhauer en faveur de l'introspection positive, l'accent est mis ici sur la conscience de soi pour défendre la thèse KK : si je sais que ϕ est le cas, alors je sais que je le sais. Mais nous verrons plus loin que cette notion sera rabaissée au titre d'usage résiduel de la notion de connaissance. La signification de base de l'opérateur K repose selon Hintikka sur un autre critère.

modales caractéristiques d'une logique de la connaissance. Avant de passer à la démonstration non-psychologique (ou non-introspective) de Hintikka, revenons d'abord sur les principaux arguments utilisés pour ou contre l'itération positive, et voyons pour quelles raisons certains arguments tantôt favorables tantôt défavorables ne sont pas retenus par Hintikka.

Par définition, la connaissance $K\phi$ est considérée comme une croyance vraie justifiée $B\phi \wedge \phi \wedge J\phi$; donc par distribution, la connaissance de la connaissance $KK\phi$ est censée satisfaire les conditions $BK\phi$ et $JK\phi$. Dans quelle mesure une connaissance est-elle à la fois un objet de croyance et de justification ? Pour répondre à cette question, nous présenterons tout d'abord les deux formes d'analyse traditionnelle de la connaissance en termes psychologiques ou cognitifs : la première méthode (internaliste) exige une faculté subjective d'introspection ; la seconde méthode (externaliste) repose sur un critère objectif de preuve ou justification. Puis nous en viendrons à un troisième critère qui, selon nous, est irréductible aux deux précédents et permet de justifier de façon convaincante l'axiome **4** (section 2.2.4.2.2). Cette dernière ligne d'argumentation, *pragmatique*, donnera raison par ailleurs à l'affirmation précédente de [Hocutt 1972] selon laquelle la logique modale épistémique n'est épistémique (n'est liée à l'analyse de ce qui concerne la cognition) que par le nom.

2.2.4.2.1 Les approches cognitives de la connaissance

L'argument principal porté contre l'introspection positive est illustré dans [Radford 1966] et [Radford 1967], dans lesquelles deux conditions possibles de la connaissance montrent leur insuffisance : non seulement il est possible de savoir quelque chose sans en être absolument *certain* (Radford s'inspire ici de l'analyse de Ayer en vertu de laquelle la connaissance implique le sentiment, justifié ou non, de certitude), mais on peut également savoir quelque chose sans en avoir la preuve. Dans les deux cas, dit Radford, 'savoir' n'implique pas 'savoir que l'on sait' : dans le premier exemple, Radford interprète l'expression 'je sais que je sais' par 'je suis *certain* que je sais' ; dans le second exemple, il l'interprète par 'j'ai une *preuve* que je sais'. L'auteur constate que l'antécédent peut être vrai sans le conséquent et en conclut que la 'thèse KK' n'est pas caractéristique du concept de connaissance.

Soit 'p' l'énoncé 'la voiture est fermée à clef'. Je peux savoir que la voiture est fermée à clef, si je dispose d'une preuve à cet effet ; donc Kp. Mais je ne serai pas certain de le savoir pour autant et il se peut que je revienne sur mes pas pour vérifier ; donc $\neg KKp$. Or si être certain de ce que l'on sait est, comme Ayer le prétend, une des conditions nécessaires à la connaissance, alors je devrais dire que je sais que ma voiture est fermée à clef mais ne suis pas certain de le savoir ($Kp \wedge \neg KKp$). Deux conséquences peuvent être tirées de cet exemple : soit l'on maintient l'inférence $K\phi \rightarrow KK\phi$ et, par contraposition, j'en conclus que je ne sais pas si la porte est fermée à clef : $[(Kp \rightarrow KKp) \wedge \neg KKp] \rightarrow \neg Kp$. Ce n'est pas ce raisonnement que Radford emploie. L'introspection positive n'est pas considérée comme une prémisse mais comme la conclusion à réfuter : je peux savoir quelque chose sans avoir la certitude de le savoir, donc Kp peut être vraie sans KKp, donc ($Kp \wedge \neg KKp$) est satisfiable lorsque la condition de certitude est incluse dans la caractérisation de la connaissance.

Dans le second exemple, je peux donner une bonne réponse à une question qui m'a été posée par hasard (la date de décès de la reine Elizabeth 1ère), mais cette bonne réponse ne constituera pas une connaissance et pour deux raisons liées : je n'ai pas de preuve adéquate en

sa faveur et, qui plus est, je ne peux pas croire que ma réponse est juste si je l'ai donnée au hasard, parce que la probabilité d'erreur est largement supérieure dans ce cas à la probabilité de bonne réponse. Je ne sais donc pas que je connais la bonne réponse à la question posée, donc je ne sais pas que je sais. Les deux exemples que Radford a donnés ont utilisé deux conditions distinctes de la notion de connaissance : la certitude d'une part, la justification d'autre part. Bien que la définition platonicienne de la connaissance n'ait pas recouru à la première condition, la démonstration de Radford visait à montrer qu'il existe des usages de la notion de connaissance qui contredisent l'argument de l'introspection positive et réfutent ainsi la connaissance de second ordre. Aux contre-exemples précédents, nous répondrons que l'existence d'autres usages de la connaissance ne suffisent pas à rejeter par l'absurde le programme de logique modale épistémique.

Comme dans la plupart des argumentations critiques du genre, les caractérisations que Radford a proposées ne correspondent pas en effet à la signification de base de la connaissance comme croyance vraie *entièrement* justifiée. D'une part, le cas de la réponse chanceuse de [Radford 1967] n'est qu'une simple opinion vraie et, à cet égard, ne satisfait pas le savoir de Hintikka. Il existe certes des usages déviants de K, mais il en existe également pour la double négation et cela ne fait pas moins de la négation classique une forme de signification de base dans notre répertoire ordinaire. D'autre part, le critère de certitude n'apparaît jamais parmi les conditions nécessaires de la connaissance chez le Finlandais : la certitude ou droit d'être sûr a été proposée par Ayer dans la définition de K,[200] mais soit celle-ci est trop subjective pour être acceptable (être borné ou intransigeant peut impliquer la certitude, mais cela n'implique pas un cas de savoir), soit elle insère une notion de 'droit' d'être sûr qui réintroduit de fait une sorte de justification dans l'analyse de la connaissance. Mais la justification qui est suggérée par la notion d'introspection n'est pas une 'évidence' au sens d'une preuve objective : il ne s'agit pas de trouver un argument A en faveur de B, car l'évidence en question ne fait pas appel à une autre connaissance que celle qui est à justifier. L'évidence s'obtient par introspection et se décrit variablement en termes de connaissance de soi [*self-knowledge*], de transparence de la connaissance ou d'illumination intérieure [*self-illumination*].[201] Ces arguments nous amènent désormais à la controverse classique entre *internalisme* et *externalisme* en épistémologie, sur laquelle la défense de la thèse KK repose partiellement.

[Chisholm 1982] et [Pritchard 2001] défendent la thèse KK par un argument de l'internalisme, position épistémologique en vertu de laquelle les conditions d'accès à la justification d'une connaissance sont *internes* au sujet et accessibles par une réflexion sur ses contenus de pensée. Si l'internaliste dit vrai, alors toute connaissance peut être considérée comme un objet accessible en toute transparence, toujours à portée de son possesseur. Si au contraire la justification d'une connaissance repose sur des conditions extérieures au sujet,

[200] J. Ayer : *The Problem of Knowledge*, Penguin Books, Harmondsworth, Middlesex (1956), sec. I, v.

[201] 'Evidence' est une expression ambiguë qui recouvre ces deux sens distincts : tantôt le sens objectif d'une preuve, d'un témoignage ou d'un argument externe au sujet ; tantôt le sens subjectif d'une certitude intérieure du sujet. Le premier sens n'est pas psychologique et correspond à la notion anglaise d''evidence' ; le second sens est psychologique et correspond à la notion française d''évidence'. Il est préférable de traduire le premier sens par 'données empiriques' ou 'témoignages', pour ne pas le confondre avec le sens courant de l'évidence comme certitude. L'appellation d'*introspection* positive oriente l'interprétation du schéma **4** vers un sens psychologique, mais rien n'y oblige.

telle qu'une preuve annexe ou une relation causale avec son environnement, alors le sujet n'a pas besoin d'un accès privilégié à ses connaissances et peut même être dans l'incapacité d'y accéder.

Que penser de ce critère de connaissance interne, dont semble dépendre ici la défense de l'introspection positive ? Il n'est pas question d'entrer dans un débat d'épistémologie, ici, pas plus qu'il ne serait question d'entrer dans un débat d'ontologie pour défendre des théorèmes de logique classique. [Hintikka 1962] défend la condition (C.KK*) en ces termes laconiques :

> Une personne peut être justifiée de dire 'je sais' seulement si ses raisons sont 'décisives' ou 'adéquates' en un sens. Je ne suis pas en position de dire 'je sais' tant que mes raisons de le dire sont telles qu'elles ne me donnent pas le droit de négliger toute évidence ou information supplémentaire.[202]

Ce passage élude le débat d'arrière-plan entre internalistes et externalistes, puisqu'il considère comme intégrée à la notion de savoir la condition décisive de justification ; s'il revient aux épistémologues la tâche de discuter la nature et les conditions d'acquisition de ces raisons décisives, en revanche le logicien épistémique ne s'intéresse qu'aux inférences logiques, c'est-à-dire aux relations de conséquences validées en vertu de la caractérisation des constantes logiques. Et compte tenu de ce qui vient d'être dit, il s'agit pour Hintikka non pas de discuter ce qui permet d'attribuer une connaissance à quelqu'un mais de discuter ce que l'on peut inférer d'une telle attribution, une fois cette attribution faite.

Comme l'indique cet extrait, la solution proposée par Hintikka pour défendre la thèse KK ne se situe pas dans une analyse de la condition de justification J, pourtant nécessaire à l'acquisition de K, mais dans les conséquences du fait de dire 'je sais'. D'une interprétation psychologique du savoir, on passe ainsi à une interprétation performative de la *déclaration de* savoir. L'analyse à venir nous permettra de donner une orientation globale à la suite du travail et, par ailleurs, d'évoquer d'autres formules et axiomes épistémiques tels que la thèse de l'entraînement : $K\phi \rightarrow B\phi$, ou les modalités mixtes : $K\phi \rightarrow KB\phi$ et $B\phi \rightarrow BK\phi$. Chacune d'elles trouvera une explication simple en termes de déclaration de savoir ou de croyance.

[Hintikka 1956] manifeste très tôt son opposition aux arguments de type psychologique, même lorsque ceux-ci vont dans le sens de son système logique :

> Je tiens à nier que le savoir et la croyance soient des *états mentaux* pouvant être reconnus par introspection.[203]

Mais comme nous venons de le signaler plus haut, son argumentation n'insiste pas non plus sur la condition de justification et n'aborde presque jamais le débat entre internalistes et externalistes. Dans un premier temps, l'auteur rejette ce que Gilbert Ryle a appelé le caractère

---

[202] [Hintikka 1962], p. 20.

[203] [Hintikka 1962], p. 56 ; nous soulignons. Le savoir a une connotation objective qui la distingue d'un simple état mental, par ailleurs. Considérer K comme un état mental revient donc à assimiler la connaissance à un sentiment de certitude ou de croyance forte ; or cette approche quantitative n'est pas celle de Hintikka ni celle de la définition classique de K, qui distingue la connaissance de la croyance non en termes d'intensité de la croyance mais en termes de justification objective. Quoi qu'il en soit, ni les propriétés de B ni celles de J n'expliqueront dans ce qui suit l'implication allant de B à BB et de K à KK.

auto-suggéré [*self-intimating*] de la connaissance parce qu'il conduirait à des conséquences inacceptables. Hintikka présente son argument à partir d'un rejet de l'axiome épistémique mixte $B\phi \rightarrow KB\phi$: en raison du caractère *décisif* (irréfutable) du savoir selon Hintikka, admettre cet axiome signifierait que rien ne peut contrarier la croyance de ϕ et qu'il est impossible d'imaginer que l'agent finisse par ne plus le croire. Or puisque une croyance n'est jamais décisive et que seul ce qui implique la vérité l'est, c'est-à-dire la connaissance, la condition sémantique (C.BK) de [Hintikka 1962] doit être rejetée : la formule ($B\phi \rightarrow KB\phi$) n'est pas un théorème, donc croire n'implique pas toujours que l'on sait ce que l'on croit. Mais l'intérêt de l'argumentation de Hintikka réside dans sa manière d'interpréter cette inférence. Pourquoi sommes-nous tentés d'admettre malgré tout l'idée que nous savons ce que nous croyons ? C'est parce que, déclare Hintikka, notre interprétation des verbes psychologiques est intimement liée au processus d'introspection, malgré les conséquences absurdes que cela peut entraîner. Dont la suivante :

> Si je peux découvrir en cherchant dans mon esprit ce que je sais ou ce que je crois, je dois de même être capable de découvrir ce que je ne sais pas ou ce que je ne crois pas. (Cela a été déclaré en effet par des personnes qui se basent sur des arguments introspectifs.) En d'autres termes, il devrait en résulter, entre autres, que chaque fois que je ne sais pas quelque chose je sais virtuellement que je ne le sais pas.[204]

Etant donné que l'introspection suppose notre capacité à (re)connaître l'état de nos connaissances, elle implique également que nous puissions avoir conscience de ce que nous ne savons *pas*. Hintikka rejette cette conséquence et sa raison par l'absurde, caractérisées par l'axiome **5** : $\neg K\phi \rightarrow K\neg K\phi$.

Même si l'interprétation psychologique des verbes d'attitudes épistémique est reléguée au second plan, en qualité de signification résiduelle, non seulement on ne peut contester son usage prégnant ; qui dit attitude dit état mental, donc suggère la psychologie. Mais de plus, on peut concilier également cette interprétation avec la signification de base du savoir 'fort' ou décisif : lorsque l'on dit que le savoir suppose une justification décisive de ce que je déclare savoir, la condition essentielle pour rendre compte de **S4** et rejeter **S5** ne réside pas dans la justification du savoir mais dans la déclaration de savoir. C'est dans ce sens performatif que [Hintikka 1962] va orienter la suite de sa démonstration, et que nous orienterons également la suite de la thèse. Dans quelle mesure cette tournure pragmatique de la logique épistémique permettra-t-elle de saisir la signification de base de K, et quel rapport y a-t-il entre une analyse du concept de savoir et de la déclaration de savoir ?

Notre interprétation performative prend sa source dans l'analyse d'un théorème épistémique contesté :

(1) $$K_aK_b\phi \rightarrow K_a\phi$$

En vertu des conditions sémantiques données dans [Hintikka 1962], cette formule s'avère auto-confirmante.[205] N'est-ce pas un résultat sinon contre-intuitif du moins étonnant, dans la

---

[204] [Hintikka 1962], pp. 53-54.

[205] La preuve donne ceci :

1. $K_aK_bp \wedge \neg K_ap \in w_0$	contre-hypothèse
2. $K_aK_bp \in w_0$	1., (C.$\wedge$)
3. $\neg K_ap \in w_0$	1., (C.$\wedge$)

mesure où l'attribution d'un savoir à un autre que soi n'apporte pas la justification de cette connaissance ; et si le critère de la justification manque à l'appel, comment prétendre qu'un autre 'sait' la vérité d'un énoncé pour lequel nous n'avons nous-mêmes aucune preuve ? Un tel théorème insiste davantage sur la notion de garantie, introduite par l'acte de déclaration, que sur celle de justification, basée sur le contenu de la déclaration : une déclaration de savoir est *transmissible*, au sens où l'attribution de la connaissance de ϕ à un autre vous donne le droit de déclarer savoir vous-mêmes que ϕ est vraie. Une autre formule modale mixte peut être rejetée en revanche parce qu'elle ne repose pas sur cet argument d'autorité des déclarations de savoir : croire que *b* sait que ϕ ne transmet pas la même autorité à *b* et laisse le droit à l'agent *a* de contester la connaissance de ϕ, donc $B_aK_b\phi \rightarrow K_a\phi$ n'est pas un théorème épistémique au même titre que (1).

Mais le rapport entre cet argument de transmissibilité du savoir par déclaration et l'introspection positive n'est pas encore tout à fait clair. Tout d'abord, la première formule précédente était de forme $\Box\Box \rightarrow \Box$ (où $\Box$ était interprété par K), mais cette forme logique correspond à la relation inverse de celle que nous voulons défendre ici, c'est-à-dire les axiomes d'introspection de forme logique $\Box \rightarrow \Box\Box$ (où $\Box$ est soit K soit B). Or aucune des formules qui suivent ne sont des théorèmes dans [Hintikka 1962] :

(2) $$K_b\phi \rightarrow K_aK_b\phi$$

n'est pas un théorème épistémique de $\mathcal{L}_{f.ME}$, pas plus que

(3) $$K_a\phi \rightarrow K_aK_b\phi.$$

Dans (2) et (3), l'antécédent peut être vrai sans le conséquent puisque rien ne présume dans leur antécédent que les agents ont des connaissances *communes*. Dans la formule (1) de type $\Box\Box \rightarrow \Box$, en revanche, la connaissance de *b* implique la connaissance de *a* parce que ce dernier a déclaré que *b* sait que ϕ. Pour qu'un savoir propositionnel soit transmissible d'un agent à un autre, il faut donc supposer que ces agents disposent d'informations en commun sur leurs propres connaissances ou croyances, et ce partage se traduit formellement en termes d'enchâssement d'un opérateur dans un autre : *a* sait que *b* sait que ϕ ($K_aK_b\phi$), ou *b* sait que *a* sait que ϕ ($K_bK_a\phi$) ; cette condition préalable n'étant pas satisfaite par les antécédents des formules (2) et (3), elles ne sont donc pas valides. D'où l'orientation dialogique[206] et performative que va prendre la suite de l'argumentation : la thèse KK (et BB) trouve sa meilleure explication dans les actes de déclaration de savoir (ou de croyance). Le débat entre externalistes et internalistes n'a de lien qu'indirect avec l'argumentation de Hintikka : dans $K\phi$, le contenu de savoir ϕ est accessible par introspection, certes, mais dans la mesure où il est manifesté au préalable au sein d'un acte de discours 'ϕ'. C'est cette approche explicite de K et B qui va nous conduire au paradoxe de Moore, lequel donne sens à la thèse KK et

---

4. $K_bp \in w_0$	2., (C.K)
5. $p \in w_n$	4., (C.K), pour tout w_n alterne à w_0
6. $P_a\neg p \in w_0$	3., (C.$\neg$K)
7. $\neg p \in w_1$	6., (C.P*), pour au moins un w_1 alterne à w_0

Les étapes 5 et 7 violent la condition (C.$\neg$p), donc $\models (K_aK_bp \rightarrow K_ap)$.

[206] [Hintikka 1973a] donnera une perspective définitivement dialogique à son analyse logique, passant d'une sémantique des modèles à une sémantique des jeux de langage ou GTS (Game-Theoretical Semantics). Nous n'aborderons cependant pas cette partie de son œuvre, poursuivant l'analyse des actes de discours en termes d'opérateurs modaux plutôt qu'en termes de coups dans un dialogue. Dans la conclusion finale, nous examinerons la possibilité de traduction entre modalités illocutoires et dialogique.

explique également notre inclination à analyser les modalités itérées en termes psychologiques.

2.2.4.2.2 Une approche pragmatique : le paradoxe de Moore

Le paradoxe qui suit fut présenté par G. E. Moore sous deux formulations voisines : la première s'énonce 'Je suis allé au cinéma mardi dernier, mais je ne crois pas que j'y suis allé', et elle aboutit tout comme la seconde à une formule absurde mais non-contradictoire : 'Je crois avoir été au cinéma mardi dernier, mais je n'y suis pas allé'.[207] La formulation la plus connue de ce paradoxe est l'énoncé 'Il pleut, mais je ne le crois pas'. On obtient dans les deux cas les formes respectives : (4) $\phi \wedge \neg B\phi$, et (5) $B\phi \wedge \neg\phi$.[208] D'après Hintikka, le paradoxe correspond à la question suivante : comment ces énoncés peuvent-ils être absurdes et non-contradictoires à la fois, et il trouve une explication si l'on considère (4)-(5) comme des énonciations plutôt que des énoncés : les énoncés représentés par les formules (4) et (5) ne sont pas logiquement faux, mais aucun locuteur ne peut les prononcer honnêtement (dans le but de dire la vérité). Dans [Harnish 1980], l'auteur remarque que [Searle 1969] suit cette même démarche :

> D'après Searle, parler un langage c'est s'engager dans une forme de comportement intentionnel, gouverné par des lois, ou c'est réaliser différents actes de discours.[209]

Parmi les hypothèses générales qui gouvernent les règles illocutoires de discours, Harnish recense la suivante :

> Chaque fois qu'il y a un état psychologique spécifié dans la condition de sincérité, la performance de l'acte vaut comme une expression de cet état psychologique

et Searle voit dans cette règle d'expression des croyances une solution au paradoxe de Moore.[210] Or puisque l'état psychologique qui est exprimé dans un énoncé déclaratif est une croyance, (4) et (5) donnent respectivement les formes itératives

---

[207] Sur l'origine du paradoxe, voir G. E. Moore, *Ethics* (Londres, 1912), p. 125, éd. épuisée (Londres, 1947), p. 78 ; ainsi que G. E. Moore, "A Reply to My Critics", in *The Philosophy of G. E. Moore*, éd. par A. Schilpp (Evanston, Ill., 1942), en particulier pp. 541-543. Une partie de la littérature produite sur le paradoxe de Moore est recensée dans [Hintikka 1962], p. 64.

[208] [Heal 1994] signale que l'analyse de l'assertion proposée par Wittgenstein dans les *Recherches Philosophiques* ne permet pas de révéler la nature contradictoire des *deux* versions du paradoxe de Moore: d'après l'Autrichien, la croyance ou assertion de ϕ ($\vdash\phi$) signifie la vérité de ϕ : $B\phi$ présuppose à la fois l'affirmation 'ϕ' et la vérité de ϕ, en précisant bien que la relation entre $B\phi$ et ϕ n'est pas ici une implication logique : ($B\phi \rightarrow \phi$) n'est pas valide, et la relation entre les deux expressions $B\phi$ et ϕ correspond plutôt à une forme de présupposition ou d'implication conversationnelle (implicature) à la Grice, de forme 'ϕ' $\Rightarrow \phi$, ou $B\phi \Rightarrow \phi$. Cette règle de conversation signifie que ce que l'on asserte soit considéré comme vrai ; donc (5) donne $(B\phi \wedge \neg\phi) \Rightarrow (\phi \wedge \neg\phi)$. En revanche, (4) ne produit pas de contradiction parce qu'elle associe à ϕ une non-croyance $\neg B\phi$ qui n'implique pas $\neg\phi$, contrairement à $B\neg\phi$.

[209] [Harnish 1980], p. 73. [Austin 1962] distingue trois types d'actes effectués dans les langues naturelles : les actes 'locutoires', 'illocutoires' et 'perlocutoires'. L'acte locutoire correspond à ce que [Searle 1969] appelle l'acte propositionnel et correspond au contenu d'un énoncé ; l'acte illocutoire désigne l'intention du locuteur exprimée par son acte propositionnel ; l'acte perlocutoire désigne l'effet produit par un acte illocutoire sur l'interlocuteur. L'aspect illocutoire renvoie ainsi à l'*intention* du locuteur lorsqu'il exprime un énoncé.

(4B) $B(\phi \wedge \neg B\phi) \equiv B\phi \wedge B\neg B\phi$

et

(5B) $B(B\phi \wedge \neg\phi) \equiv BB\phi \wedge B\neg\phi$.

En vertu de quoi cette transformation des formules de Moore les rend non seulement absurde, mais aussi contradictoire ? En vertu de la condition sémantique (C.BB*) des structures de Kripke transitives : si $B\phi \in w$, alors $B\phi \in w^*$ pour tout monde w* alterne à w, on peut démontrer que les formules ci-dessus sont indéfendables.[211] La thèse d'introspection positive sera défendue plus loin de la même façon avec la notion de connaissance, à partir de la condition sémantique (C.KK*).[212] Une telle analyse suppose que deux conditions soient satisfaites entre-temps. Premièrement, que la question de la condition de *vérité* des énoncés soit remplacée par la question des conditions de *véridicité* ou de *véracité* des énonciations. Deuxièmement, que l'effet illocutoire d'une énonciation puisse être intégré dans le domaine de l'analyse logique et susceptible d'être formalisé. Précisons d'abord où se situe la distinction entre ces trois conditions du discours des assertions.

On distinguera tout d'abord l'aspect locutoire de l'énoncé, considéré comme un simple contenu propositionnel, de deux de ses aspects illocutoires. La *vérité* d'un énoncé est l'aspect locutoire dont s'occupe la logique standard et qui n'indique rien quant aux intentions du locuteur. Les deux aspects illocutoires sont la véridicité et la véracité. La *véridicité* d'une énonciation indique que le locuteur a l'intention de dire la vérité et y parvient ; on retrouve parfois cette appellation dans l'axiome **T** de la clause de vérité, appelée également 'clause de véridicité' pour insister sur la relation logique entre connaissance et vérité. La *véracité* d'une énonciation, enfin, indique l'intention du locuteur de dire la vérité quelle que soit l'issue de son acte de discours (satisfait ou non) ; on la retrouve à travers la 'condition de sincérité' des assertions. Une énonciation peut être sincère sans avoir un contenu vrai, donc elle peut être vérace sans être véridique ; en revanche, une énonciation qui n'est pas vérace ne peut pas être véridique, car celui qui n'est pas honnête dans son énonciation ne peut pas avoir l'intention de

---

[210] Voir [Searle 1969], chapitre 1.4, p. 65, cité in [Harnish 1980], p. 74. Ce dernier précise toutefois que l'hypothèse de Searle n'est pas suffisante et suppose des conditions plus sévères pour résoudre le paradoxe de Moore. Ces conditions indexicales supplémentaires seront précisées plus loin, car on les retrouve également dans [Hintikka 1962].

[211] La preuve d'indéfendabilité de (4B) procède ainsi, présentée dans les termes de [Hintikka 1962] :

1. $B(\phi \wedge \neg B\phi) \in w$	prémisse
2. $\phi \wedge \neg B\phi \in w^*$	1., (C.b*), pour au moins un w* alterne à w
3. $B(\phi \wedge \neg B\phi) \in w^*$	1., (C.BB*), pour tout w* alterne à w
4. $\neg B\phi \in w^*$	2., (C.$\wedge$)
5. $C\neg\phi \in w^*$	4., (C.$\neg$B)
6. $\neg\phi \in w^{**}$	5., (C.C*), pour au moins un w** alterne à w*
7. $\phi \wedge \neg B\phi \in w^{**}$	3, (C.B*), pour tout w** alterne à w
8. $\phi \in w^{**}$	7., (C.$\wedge$)

Les étapes 6 et 8 violent la condition (C.$\neg$), donc la prémisse est indéfendable.

La preuve de (5B) repose sur la substitution de $\neg B\phi$ à $B\neg\phi$, ce qui la rend d'autant plus indéfendable : si l'étape 8 montre qu'il est inconsistant de ne pas croire ϕ, c'est-à-dire de supposer la vérité de $\neg\phi$ ne serait-ce que dans *au moins un* monde alterne à w, alors il est d'autant plus inconsistant de croire ϕ, c'est-à-dire de supposer la vérité de $\neg\phi$ dans *tous* les mondes alternes à w.

[212] L'indéfendabilité de la version *épistémique* du paradoxe de Moore est encore plus évidente : en vertu de (C.K), la formule (4K) $K(\phi \wedge \neg K\phi) \equiv (K\phi \wedge K\neg K\phi)$ est réductible à la forme contradictoire $(K\phi \wedge \neg K\phi)$. La forme (5K) ne donne pas une relation de contradictoires mais de contraires : (5K) $KK\phi \wedge K\neg\phi \equiv K\phi \wedge K\neg\phi$. L'implication $(K\neg\phi \rightarrow \neg K\phi)$ permet toutefois d'aboutir au même résultat contradictoire.

la vérité par elle. On considère comme allant de soi que l'assertion est toujours vérace en tant que telle : on n'asserte pas proprement dit ce que l'on ne croit pas ; en revanche, l'avis des auteurs est plus nuancé lorsqu'il s'agit de décider si un acte d'assertion présuppose la condition de véridicité ou se contente d'une simple condition de véracité : soit l'acte d'assertion doit porter sur un énoncé vrai pour être réussi, soit il lui suffit d'être une déclaration de vérité dans laquelle la condition de vérité ne fera pas partie des conditions de satisfaction. L'assertion telle que nous la présenterons dans ce qui suit ne présupposera pas la vérité du contenu propositionnel pour être réussie : une assertion agit comme une déclaration d'autorité sur un interlocuteur ou un auditeur, mais la vérité du contenu de discours n'est pas essentielle pour produire un effet sur le destinataire et influencer son comportement. La distinction de [Hanson 1952] entre déclarations factuelles et déclarations de fait sera utile à rappeler pour nos besoins. Une déclaration de fait signifie qu'une énonciation porte sur un fait avéré, contrairement aux déclarations factuelles :

> les déclarations de fait sont des déclarations factuelles vraies ; ce sont des déclarations des faits. Dire de celles-ci qu'elles sont fausses, ou qu'elles sont corrigibles, c'est dire qu'elles ne sont pas ou peuvent ne pas être des déclarations de fait. Ce qui est une réfutation en soi.[213]

La formalisation des aspects illocutoire du langage est admise de nos jours ; mais elle n'allait pas de soi pour certains commentateurs de [Hintikka 1962]. [Hocutt 1972] reprochait notamment que l'on parlât de 'logique' épistémique là où il ne voyait qu'une simple analyse linguistique des déclarations de savoir : pour reprendre les termes de [Chisholm 1963], il reprochait la confusion produite entre logique de la connaissance [*logic of knowledge*] et logique de l'acte de connaître [*logique of knowing*]. Une telle démarche paraissait inacceptable pour Hocutt dans la mesure où, selon lui, la logique traite des conditions de vérité et la notion de vérité exprime une relation de satisfaction entre un énoncé et des faits. Cette restriction de la logique à l'aspect locutoire du langage conforte le rôle central de la logique classique en rejetant d'emblée ce que [Searle 1969] appellera par la suite des opérateurs de force illocutoire. Mais sans parler de force illocutoire, [Hintikka 1968] conteste cette distinction tranchée entre les aspects sémantique et pragmatique du langage :

> Une grande part de confusion a été créée par des applications exagérément simplifiées de la trichotomie carnapienne syntaxe-sémantique-pragmatique (...) Ce genre d'usages simplifiés à l'excès a favorisé l'illusion que toute étude des *usages* du langage devrait se situer au-delà du champ des méthodes logiques et appartenir à la psychologie ou à la sociologie du langage plutôt qu'à la logique ou à la philosophie. Mais il n'y a pas une once de raison pour laquelle les structures générales exposées par le langage ne pourraient pas être étudiées également dans l'*usage* par des moyens logiques et mathématiques. (...) les structures logico-formelles sont souvent en mesure d'éclairer les travaux sur le discours ordinaire dans la mesure où elles peuvent être conçues comme des manières possibles d'utiliser le langage –ne l'utilisant pas seulement pour le plaisir, mais dans un but précis. Bien que Carnap ne doive pas être blâmé

[213] [Hanson 1952], p. 24. Si l'on symbolise l'intention de dire la vérité (croire que le contenu propositionnel ϕ est vrai) par $BT\phi$, la différence entre les deux types de déclaration peut être rendue ainsi : ($BT\phi \wedge \phi$), pour les déclarations de fait ; $BT\phi$, pour les déclarations factuelles.

pour la confusion malheureuse que sa trichotomie a engendrée, il me semble qu'il est temps de remplacer la trichotomie par un schéma plus flexible.[214]

Des auteurs tels que Hocutt ou Kielkopf caractérisent au contraire la logique par ses seuls aspects syntaxique et sémantique : une syntaxe définit les combinaisons admises entre les signes d'un langage ; une sémantique définit les relations entre les signes et leurs objets de référence. Quant à la relation pragmatique entre les signes d'un langage et leurs utilisateurs (le locuteur et les interlocuteurs), elle est exclue du domaine logique dans la mesure où elle ne porte pas sur des conditions de *vérité* proprement dites.

Mais la notion de vérité devient plus flexible lorsqu'elle est considérée à l'aune de la sémantique relationnelle et de ses modèles multiples : dans un espace épistémique, par exemple, être vrai dans un modèle signifie être vrai pour un sujet et dans un champ d'information extensible. La différence entre la logique modale épistémique et la logique classique ne réside pas dans la signification de la notion de vérité, mais dans les modalités de son application : bien que la vérité désigne dans les deux cas la satisfaction d'une fonction par un ensemble d'objets du modèle, cette satisfaction ne dépend pas d'un domaine de référence unique et objectif mais multiple et subjectif. Dans le cas des actes de langage tels que les présente le paradoxe de Moore, l'indéfendabilité de (4) et (5) ne signifie pas que les énoncés sont logiquement faux : un énoncé peut être vrai même si je ne le crois pas, et un énoncé peut être faux même si je crois à sa vérité. En revanche, elle signifie que ces énoncés sont impossibles à prononcer pour le locuteur : je ne peux pas croire qu'un énoncé est vrai si je ne le crois pas, et je ne peux pas croire qu'un énoncé est faux si je crois à sa vérité. Ce ne sont donc pas les énoncés mais les énonciations qui sont logiquement fausses, et l'impossibilité de ces énonciations peut être analysée en termes référentiels de vérité des énoncés *du point de vue de leur locuteur*. Et puisque l'assignation de vérité dans un modèle épistémique signifie la vérité du point de vue du sujet, la relation *pragmatique* entre un énoncé et son locuteur peut être exprimée formellement par le biais d'une théorie des modèles. C'est le cas en logique modale épistémique, où les formules épistémiques peuvent être considérées comme l'expression d'actes d'énonciations ; la preuve avec le traitement précédent des 'énoncés mooréens'.

Pour montrer que (4) et (5) sont indéfendables, [Hintikka 1962] formalise leur impossibilité pragmatique en termes d'indéfendabilité doxastique et épistémique : prononcer l'expression 'ϕ' signifie pour son locuteur 'je crois que le cas est le suivant : ϕ (est le cas)'.[215] La procédure consiste à paraphraser les énoncés de Moore en termes d'énonciation, à partir d'une *internalisation* des modalités assertoriques au sein du langage-objet : donner le sens interne à un acte assertorique, c'est symboliser l'état psychologique ou force illocutoire dans lequel se trouve son auteur.[216] Bien que le sens externe des énoncés ne paraisse pas

[214] [Hintikka 1969], note 3 p. 3. L'origine de la distinction entre syntaxe, sémantique et pragmatique est attribuée au logicien américain Charles W. Morris, in *Foundations in the Theory of Signs*.

[215] Cette paraphrase des actes d'énonciation en termes de démonstratif 'c'est le cas que –' avait été utilisée au début de [Hintikka 1962] pour évoquer un problème de traduction des expressions épistémiques : lorsque l'on dit 'il ne sait pas *que* ϕ', la complétive 'que ϕ' est l'objet d'une assertion et présuppose que ϕ est vraie : $\neg K_{il}\phi \wedge \phi$, contrairement à l'expression voisine 'il ne sait pas si ϕ' qui laisse le doute sur la vérité de ϕ : $\neg K_{il}(\phi \vee \neg\phi)$.

[216] Les expressions de sens externe et de sens interne sont utilisées entre autres par [Åqvist 1964] : le sens externe de 'je déclare ϕ et je déclare $\neg\phi$' correspond à la forme locutoire '$\phi \wedge \neg\phi$', et son sens interne donne la

logiquement inconsistant, la condition de sincérité qui est associée aux actes de langages tels que l'assertion permet de montrer que le sens interne de (4) et (5) donne soit une contradiction : $B(\phi \wedge \neg B\phi)$, soit une inconsistance : $B(B\phi \wedge \neg\phi)$. La solution au paradoxe de Moore repose donc sur la traduction formelle des assertions ou déclarations [*statements*] de vérité en termes de modalités épistémiques :

> Ce qui est violé dans le fait de *prononcer* [$\phi \wedge \neg B\phi$], ce n'est pas la consistance logique (défendabilité), c'est plutôt la présomption générale que le locuteur croit ou, tout du moins, peut croire de façon concevable ce qu'il dit.[217]

Ainsi n'est-il pas logiquement contradictoire mais contraire à l'intention du locuteur (dire la vérité) d'asserter 'ϕ mais je ne crois pas que ϕ'. La notion pragmatique d'*auto-annulation* [*self-defeat*] constitue le pendant illocutoire de la contradiction locutoire : un énoncé auto-annulant ϕ correspond à une énonciation contradictoire ou logiquement impossible, c'est-à-dire $B\phi \rightarrow \bot$. L'impossibilité illocutoire de prononcer des énoncés tels que (4) et (5) est qualifiée par Hintikka d'*indéfendabilité doxastique* (ou *épistémique*) : un ensemble d'énoncés $\{\phi_1 \wedge \phi_2 \wedge \ldots \wedge \phi_k\}$ est doxastiquement indéfendable à énoncer pour l'agent *a* si et seulement si l'énoncé '$B_a(\phi_1 \wedge \phi_2 \wedge \ldots \wedge \phi_k)$' est indéfendable tout court [*simpliciter*]. Dans les cas de (4) et (5), on a affaire à deux formules doxastiquement indéfendables dans la mesure où $B(\phi_1 \wedge \neg B\phi_1)$ et $B(\neg\phi_1 \wedge B\phi_1)$ sont indéfendables. En termes généraux, le paradoxe de Moore peut être étendu du cas particulier de la première personne du singulier 'je' à n'importe quelle variable de personne *a* : (4) et (5) sont *doxastiquement* indéfendables à énoncer pour quelque locuteur, pour autant qu'il énonce ses propres croyances et non celle d'un autre.[218]

Il faut noter toutefois que la condition de sincérité (on ne dit que ce que l'on croit être vrai) ne suffit pas pour rendre compte pleinement de l'absurdité des énoncés de forme (4) et (5) : les conditions *indexicales* du temps (à quel moment du temps ces énoncés font-ils référence ?) et de l'identité du locuteur (par qui l'énoncé est-il prononcé ?) sont deux autres critères de paradoxalité qui doivent être précisés pour révéler une contradiction interne.[219]

---

version illocutoire (doxastique) '$B\phi \wedge B\neg\phi$', considérant qu'une déclaration exprime une croyance de la part du locuteur. [Åqvist 1964] attribue également différents degrés aux sens internes, selon l'état psychologique qu'ils incarnent : assertion, certitude, conviction, supposition, etc. Les degrés de force illocutoire que nous considérons pour notre part sont au nombre de deux : la connaissance et la croyance. On retrouve cette dualité externe/interne dans [Bochvar 1938] et [Searle 1969] : partant de la distinction entre métalangage et langage-objet, le sens interne d'une formule correspond à l'introduction d'un jugement du métalangage dans le langage-objet ; c'est le cas en logique modale épistémique, où croyance et connaissance sont des formes de jugement dont on peut calculer le sens.

[217] [Hintikka 1962], p. 67.

[218] Le paradoxe réside dans l'usage *performatif* des modalités épistémiques, c'est-à-dire l'intention du locuteur qu'elles servent à exprimer : "Bien que l'usage des pronoms de la première personne du singulier nous permette de parler d'*énoncés* indéfendables, et pas simplement de *déclarations* indéfendables, l'héritage logique de cette notion 'absolue' d'indéfendabilité doxastique est révélé ainsi par le fait que leur caractère d'être contraire au but recherché ne se manifeste que lorsqu'ils sont énoncés de manière *assertive* ou professés. En un sens, l'absurdité des énoncés doxastiquement défendables est de *caractère performatif*; elle est due à l'*acte* plus qu'au moyen (l'énoncé) employé pour ce faire ([Hintikka 1962], p. 77 ; nous soulignons).

[219] "La notion d'indéfendabilité doxastique, à la différence des autres notions que nous avons définies jusqu'ici, n'a pas été définie tant pour les *énoncés* [*sentences*] (ou pour des ensembles d'énoncés) que pour des ensembles d'*affirmations* [*statements*]. Elle ne dépend pas uniquement de la forme de mots énoncée ; elle dépend également du locuteur (ou auteur) et des façons dont il fait référence à lui-même. Elle ne peut être définie pour des ensembles d'énoncés qu'en la relativisant à un nom ou un pronom figurant dans les énoncés en question.

D'une part, une formule de type (4) ne sera pas paradoxale si le temps de l'énoncé et le temps de l'énonciation ne concordent pas, c'est-à-dire si les événements qu'ils expriment donnent lieu à des incohérences logiques. Pour donner deux exemples, les énoncés tels que 'Il pleuvait, mais je ne le croyais pas', ou 'Il tuera son père, mais je ne le croirai pas' ne sont ni contradictoires ni absurdes à prononcer par le locuteur référé par l'indexical 'je'.[220] La logique des énoncés mooréens repose donc sur un facteur temps : pour qu'une assertion soit conformes aux règles du discours, le locuteur doit s'engager sur la vérité de ce qu'il dit et croit *maintenant* ; à l'inverse, il lui est interdit de s'engager à la fois sur ce qu'il ne croit *plus* et sur ce qu'il ne croit pas *encore*. En effet, les énoncés exprimant des opinions passées ou futures peuvent décrire des comportements dont la vérité n'est plus ou pas encore assumée par le locuteur ; en ce sens, leur énonciation n'aura rien de paradoxale tant qu'elle ne viole pas la condition de sincérité des actes d'assertion.[221] Dans le cas de l'énonciation présente d'un énoncé au passé, 'Il pleuvait, mais je ne le crois pas' est absurde à prononcer parce qu'il revient pour le locuteur à rejeter la vérité d'un événement passé dont il est en train d'affirmer la vérité. On peut expliquer cette absurdité par le caractère irrévocable des vérités passées : ce qui était vrai est vrai et sera toujours vrai, donc une vérité passée est nécessairement vraie à partir du moment où le fait qu'elle exprime a eu lieu. Dans le cas de l'énonciation présente d'un énoncé au futur, les vérités futures sont considérées en général comme des vérités contingentes au temps présent, c'est-à-dire indéterminées au moment de leur énonciation. La contradiction d'une telle énonciation ne sera donc pas aussi flagrante que dans le cas précédent, mais elle en reste une pour la raison que l'acte d'énoncer une vérité future a pour effet de considérer cette vérité comme déterminée : un énoncé tel que 'Il y aura de la neige à Noël, même si je ne le crois pas pour l'instant' est absurde à prononcer pour le locuteur.[222]

Lorsque le locuteur (ou auteur) visé est manifeste, (...) l'identité du locuteur (ou auteur) doit être alors *déduite* [*gathered*] du contexte" ([Hintikka 1962], p. 74).

[220] "De même qu'un changement du temps grammatical supprime l'absurdité de [(α)], témoin un énoncé tel que 'Il était à la maison mais je ne le croyais pas" ([Hintikka 1962], p. 65). Puisque la logique modale épistémique évoquée ici est monomodale, l'auteur constate que "nous n'avons aucun moyen de prouver sa défendabilité formellement, car nous ne traitons pas de la *logique des temps*, mais la remarque est suffisamment évidente" (p. 69). Du fait que la condition de sincérité repose sur la concordance entre les temps de l'énonciation et du contenu de l'énoncé, nous n'aurons pas besoin de logiques multimodales à la fois épistémiques et temporelles. Nous nous en passerons également dans le cadre du paradoxe de Fitch (section 3.1.1.2), auquel le même traitement pragmatique sera appliqué.

[221] Une des clauses de l'analyse de [Hintikka 1962] est l'unité de temps : les conditions de consistance d'un ensemble d'énoncés réfèrent à un seul et même moment. Lorsque une énonciation porte sur des attitudes épistémiques futures ou passées, celles-ci peuvent indiquer des convictions que le locuteur a rejetées entre-temps ou qu'il prévoit de rejeter bien qu'il en assume la vérité maintenant, c'est-à-dire au moment de l'énonciation. Ce phénomène oppose le plus souvent le passé au présent, mais il est possible également de prévoir un changement d'avis pour le futur : "Celui qui n'est pas socialiste à vingt ans n'a pas de cœur ; celui qui est socialiste à quarante n'a pas de tête", selon la formule d'Aristide Briand recensée dans [Bovens 1995] ; une telle énonciation est impossible à prononcer pour un jeune idéaliste, dans la mesure où le faire consisterait à admettre qu'il finira par rejeter tôt ou tard ce en quoi il croit fermement aujourd'hui.

[222] L'hypothèse de la nécessité des vérités passées constitue la première prémisse de l'Argument Dominateur (χυριευον λογος), un trilemme antique dont la résolution avait opposé plusieurs philosophes tels que Diodore Cronos, Cléanthe ou Chrysippe. La définition extensionnelle de la nécessité diodorienne en termes de vérité omnitemporelle (vue au Chapitre 1) est liée à l'origine de ce problème. Quant aux énonciations de vérités futures, le seul cas dans lequel elles peuvent être paraître doxastiquement défendables est celui de la *self-deception* (traduit par 'aveuglement') : même si je dois croire ce que je considère comme vrai, il y a certaines vérités que je n'arrive pas à admettre. La dénégation freudienne est un cas similaire, mais nous ne considérerons

D'autre part, le paradoxe de Moore est pragmatique et dépend aussi de l'identité du locuteur : (4) est paradoxal si le temps grammatical de l'énoncé est le présent (de l'indicatif, bien entendu) ou si le locuteur parle de ses propres croyances ; il ne le sera plus en revanche si le locuteur évoque les croyances d'un autre. Ainsi, l'énoncé 'il pleut, même si ces abrutis ne le croient pas' n'a rien de paradoxal parce que l'identité du locuteur ne concorde pas avec les sujets de son mépris. Tout comme dans l'analyse du paradoxe en termes de temps grammaticaux, la concordance dont il est question ici ne signifie pas un critère d'identité : de même que les énoncés mooréens peuvent être paradoxaux sans que le temps grammatical de l'énonciation soit identique au temps grammatical de l'énoncé, ils peuvent l'être également sans que l'auteur de l'énonciation soit identique au sujet de l'énoncé. Hintikka signale à cet effet que (4) et (5) ne sont pas paradoxaux uniquement lorsqu'ils figurent à la première personne du singulier : ils le restent lorsque 'je' est remplacé par les pronoms personnels 'tu', 'nous' ou 'vous', c'est-à-dire dans tous les cas où se produit une relation de locuteur à interlocuteur avec le sujet de l'énoncé ; ce qui est vrai pour les pronoms personnels l'est également pour les noms propres, puisque le mode de référence du sujet ne compte pas dans l'évaluation du paradoxe.[223] Le rôle de cette relation de locuteur à interlocuteur renvoie à la propriété de transmissibilité épistémique, déjà vue plus haut : lorsque un locuteur parle à son interlocuteur, son intention communicative est de transmettre une information vraie et son assertion produit un effet d'adhésion sur son vis-à-vis.[224] Cela dit, cet effet *perlocutoire* de l'énonciation dépend de la *force illocutoire* que lui transmet son auteur : 'ϕ, mais tu ne crois pas que ϕ' semble plus ou moins paradoxale selon qu'elle exprime une déclaration de vérité ou une simple supposition réfutable. Dans ce dernier cas, le locuteur donne son opinion mais laisse à l'interlocuteur le droit de la contester.[225] Il en va autrement lorsque l'assertion désigne une déclaration de savoir : 'ϕ, mais tu le ne crois pas' est absurde à prononcer pour le locuteur dans la mesure où l'assertion de ϕ transmet à son interlocuteur une connaissance que celui-ci ne peut réfuter en tant que telle, et ce dernier ne peut continuer de croire $\neg\phi$ dès que le locuteur lui a 'fait savoir' que ϕ.[226] Ce processus à double face (illocutoire et perlocutoire) de

---

pas ce phénomène *patho*logique de résistance à l'évidence comme l'expression d'une contradiction logique. Nous reviendrons sur ce cas de figure au Chapitre 3.

[223] [Hintikka 1962] cite les cas originaux de Charles de Gaulle et Napoléon, réputés pour parler d'eux-mêmes à la troisième personne du singulier ou par leurs noms propres. Mais ces cas singuliers de mise à distance entre le locuteur et le sujet de l'énoncé ne sont que fictifs, parce que le locuteur sait qu'il parle de lui à la troisième personne ou par le biais d'un nom propre. Cette précision permet à Hintikka de généraliser la notion d'indéfendabilité doxastique (ou épistémique) : pour tout locuteur *a*, '$\phi \wedge \neg B_a\phi$' est indéfendable à prononcer par *a* si celui-ci sait qu'il est référé par *a*. Cette dernière condition avait été mise en évidence par [Castañeda 1970] et donné lieu à une reformulation la condition (C.KK*) : 'si a sait que ϕ, alors a sait qu'*il* sait' est valide seulement si '*a*' et 'il' sont co-référentiels (si 'il' a un usage anaphorique dans ce cas de figure).

[224] "Le fait que ['$(K_a p \wedge \neg K_b p)$'] soit épistémiquement indéfendable à énoncer pour le porteur de *b* peut apporter un éclairage sur les problèmes examinés dans [le fait de simplement croire que vous savez]. Il montre qu'une implication épistémique a lieu entre 'vous savez que' (ou 'il sait que') et 'Je sais que' ([Hintikka 1962], p. 81).

[225] Le droit de contester est typique d'une logique dialogique : lorsque le Proposant P propose un énoncé ϕ, il ne coupe pas court à toute discussion et ne force pas son interlocuteur à adhérer à la vérité de ϕ ; il ouvre au contraire un débat polémique avec lui, lequel est un Opposant O et non un Adhérent A. Ce que O conteste est donc une supposition de départ et non une assertion dirimante, comme dans le cas des énoncés mooréens.

[226] En ce sens, l'assertion de ϕ provoque une *connaissance commune* entre le locuteur et son auditoire : la propriété de transmissibilité du savoir préfigure ainsi la logique modale épistémique multi-agents de van Benthem, et son lien avec le paradoxe de Moore a été exprimé ainsi par [Williams 2000] : "L'assertion implique [*involves*] une tentative de faire d'une proposition une connaissance commune" (p. 32), et la tentative est

la transmissibilité du savoir par le biais d'une déclaration ne fonctionne pas si le pronom figurant dans l'énoncé est 'il' ou 'ils', dans la mesure où le locuteur n'est pas alors en contact avec la personne référée dans l'énoncé et ne peut donc lui transmettre aucune connaissance.

Pour résumer l'analyse ci-dessus des énoncés mooréens, les conditions transitives (C.BB*) et (C.KK*) permettent d'invalider un type de formules dont l'absurdité repose sur des critères indexicaux de temps grammatical et de personne référée dans l'énoncé. Mais le résultat de cette analyse peut varier en raison d'un troisième critère de concordance entre énonciation et énoncé, en plus des critères de temps et de la personne : même si le locuteur et le sujet de l'énoncé sont épistémiquement liés, le premier produisant sur le second un effet perlocutoire (parce que le sujet de l'énoncé et le locuteur ne font qu'un, parce que le sujet est son interlocuteur, ou parce qu'il fait partie de son auditoire), la formule de type (4) ou (5) ne sera pas paradoxale si l'attitude exprimée par son énonciation concorde avec celle exprimée dans l'énoncé. Si (4) et (5) sont paradoxaux, alors on constate que les formules qui suivent le sont d'autant plus :

(4KB) $K(\phi \wedge \neg B\phi)$

et

(5KB) $K(B\phi \wedge \neg\phi)$

sont indéfendables parce que le locuteur y déclare de deux façons différentes (par l'acte d'énonciation ou dans l'énoncé, respectivement) ce qu'il *sait*, à savoir que ϕ est le cas ou n'est pas le cas. Or puisque que ce dont la vérité est impossible pour de simples croyances le sera *a fortiori* pour des connaissances, l'indéfendabilité de (4) et (5) implique respectivement celle de (4KB) et (5KB).[227] Si en revanche l'énonciation exprime une simple croyance ou supposition, plutôt qu'une conviction ou assertion, le paradoxe n'aura pas toujours lieu. C'est le cas notamment si l'attitude exprimée dans l'énoncé exprime une conviction *plus forte* que celle de l'énonciation. Dans la formule

(4BK) $B(\phi \wedge \neg K\phi)$,

l'attitude du locuteur est celle d'un agent parfaitement cohérent qui tient un certain énoncé pour vrai mais n'oublie pas qu'il demeure réfutable en tant que tel, du fait que l'introduction d'une information supplémentaire pourrait modifier son point de vue sur ϕ. Or contrairement à (4BK), défendable si le locuteur considère que ses croyances sont réfutables, celui-ci ne peut prononcer en toute cohérence des formules mixtes telles que

(4KB) $K(\phi \wedge \neg B\phi)$

ou

(4K) $K(\phi \wedge \neg K\phi)$

---

supposée fructueuse par l'indéfendabilité des énoncés (4) et (5). D'autres logiques d'assertion sont encore plus exigeantes sur la condition d'adhésion à la vérité de ϕ : dans la logique discussive de Jaskowski ou la logique assertionnelle de [Łoś 1948], la vérité d'un énoncé n'est pas transmise par l'annonce d'un seul mais exige l'accord préalable de *tous* les interlocuteurs. La connaissance commune est le résultat d'un accord intersubjectif, dans ce cas, et non l'effet d'une annonce faisant autorité.

[227] (5BK) sera trivialement indéfendable : $(K\phi \wedge \neg\phi)$ est absurde *et* contradictoire, donc le rôle de la force illocutoire B dans $B(K\phi \wedge \neg\phi)$ est superflu. Autrement dit, l'aspect paradoxal de (5BK) n'est pas d'ordre pragmatique.

Dans ces deux cas de figure, les énoncés de forme mooréenne sont paradoxaux parce que l'énonciation exprime une force illocutoire *supérieure ou égale* à celle de l'énoncé.[228] Selon le contexte de discours et quelques indices propres au cadre du discours (l'intonation de la voix, le rôle du locuteur au sein d'une conversation, etc.), on considérera que la force illocutoire associée à un énoncé est comparable soit à une simple croyance, soit à une affirmation de connaissance : dans le premier cas, l'affirmation est réfutable et s'apparente davantage à une supposition ; dans le second cas, l'affirmation est péremptoire et vaut pour son auteur comme une vérité acquise. L'énonciation de ϕ par le locuteur implique alors *épistémiquement* que l'interlocuteur acquiert la connaissance de ϕ : dire que ϕ produit une déclaration de savoir et annule toute supposition du 'contraire' $\neg\phi$;[229] en conséquence, les déclarations de forme '$p \wedge \neg K\phi$' et '$p \wedge \neg B\phi$' sont toutes deux paradoxales, au sens où elles sont épistémiquement indéfendables à prononcer par leur locuteur. Le seul cas où le paradoxe n'apparaît pas est celui où la force illocutoire sera inférieure à l'attitude indiquée dans l'énoncé, ce qui revient à dire que tous les énoncés évoqués jusqu'ici seront épistémiquement défendables à prononcer.

Le passage de l'implication locutoire à l'implication illocutoire (de force soit doxastique soit épistémique) permet ainsi d'expliquer l'absurdité des énoncés mooréens en termes d'énonciation. Mais outre ce détail, l'analyse en termes de déclaration (supposition ou assertion) permet d'expliquer également l'inclination des épistémologues à confondre itération et introspection, assertion et psychologie : lorsqu'un exemple d'énoncé mooréen est donné, celui où le pronom est à la *première personne*. Or dans ce cas de figure, le locuteur et le sujet évoqué dans l'énoncé sont une seule et même personne et, pour cette raison, les croyances suggérées par l'acte d'énonciation ne peuvent être contestées par le locuteur sans que celui-ci ne se contredise *ipso facto*. La croyance d'un agent peut être contestée en toute cohérence lorsqu'elle est prononcée par une tierce personne, c'est-à-dire une personne dont les connaissances n'interfèrent pas avec celles de l'agent (voir la formule (1)). Les formules (4)-(5) ne sont validées que lorsque le locuteur et le sujet de l'énoncé sont épistémiquement liés : pour tout *b* en contact avec *a*, $(\phi \wedge \neg B_b\phi)$ et $(B_b\phi \wedge \neg\phi)$ sont doxastiquement et, *a fortiori*, épistémiquement indéfendables pour tout *a*. La leçon à tirer de cette analyse en demi-teinte des énoncés mooréens, c'est que [Hintikka 1962] propose un cadre sémantique adéquat lorsque les attitudes épistémiques sont prononcées dans le cadre d'une conversation, c'est-à-dire lorsqu'un échange d'informations se produit entre le locuteur *a* et le sujet de l'énoncé *b*. Si l'on restreint l'interprétation de *b* au seul cas de la première personne du singulier, la conséquence est une confusion entre l'acte public d'assertion et l'acte privé d'introspection :

---

[228] On pourrait généraliser ainsi les conditions d'absurdité des énoncés mooréens. Pour tout $\Box$ symbolisant une attitude épistémique à l'égard d'un énoncé ϕ (certitude, conviction, supposition, doute, etc.), la formule $\Box_1 (\phi \wedge \neg\Box_2)$ est indéfendable seulement si $\Box_1$ indique une force illocutoire supérieure ou égale $\Box_2$. L'ordre des forces est croissant ou décroissant selon que l'attitude tend vers la considération de vérité ou fausseté absolue.

[229] L'interprétation modale des actes épistémiques justifie notre tendance à parler couramment du 'contraire' plutôt que du 'contradictoire' : bien que 'contraire' ne soit pas le terme approprié pour opposer les énoncés ϕ et $\neg\phi$, il l'est en revanche pour opposer les énonciations 'ϕ' et '$\neg\phi$' dans la mesure où une énonciation correspond à une attitude épistémique de forme $\Box\phi$. La supposition contraire à 'ϕ', c'est-à-dire $\Box\phi$, n'est donc pas l'énoncé $\neg\phi$ mais l'énonciation '$\neg\phi$', c'est-à-dire $\Box\neg\phi$. Or $\Box\phi$ et $\Box\neg\phi$ sont effectivement en opposition contraire ; la même chose vaut pour les expressions quantifiées : 'Tout le monde est là ?' ($\forall xFx$). Si l'on répond : 'Au contraire', on sous-entend que personne n'est là ($\forall x\neg Fx$) et non que certains ne sont pas là ($\neg\forall xFx \equiv \exists x\neg Fx$).

> Compte tenu de l'aspect fallacieux des arguments de l'introspection, il est important de se rendre compte qu'aucune des conditions ou règles que nous avons adoptées n'est basée sur elles. Les arguments que nous avons donnés en leur faveur concernaient tous les circonstances dans lesquelles on peut raisonnablement *dire* d'un ensemble de déclarations [*statements*] faites *explicitement* qu'elles sont défendables. Aucune référence n'a été faite à ce que l'on peut savoir en recherchant dans son *esprit*. Par exemple, la condition (C.KK*), peut-être la moins évidente de nos conditions, ne devrait pas être défendue en termes d'*introspection* (...) elle est basée sur ce que l'on pourrait peut-être appeler l'aspect *quasi-performatif* du verbe *savoir*.[230]

L'introspection est considérée ainsi comme un symptôme de l'implication épistémique : asserter que ϕ implique le fait d'avoir conscience que ϕ, mais cette prise de conscience ne prend effet que si la connaissance ou croyance est prononcée. Nous verrons par la suie (Chapitre 3) que l'axiome de 'sagesse' (introspection négative) peut être admis sous cette interprétation : on sait ce que l'on ne sait pas, à condition que l'on ait déclaré explicitement ce que l'on ignore au préalable.

La tendance à confondre assertion et introspection explique également pourquoi d'autres formes inférentielles paraissent auto-confirmantes lorsqu'elles sont interprétées en termes psychologiques : 'j'ai conscience que ϕ' implique 'j'ai conscience d'avoir conscience que ϕ', dans la mesure où la prise de conscience résulte d'un acte d'assertion explicite. Par opposition, 'il a conscience que ϕ' n'implique pas 'il a conscience d'avoir conscience que ϕ' tant que la tierce personne en question n'asserte rien.[231] Pour la même raison, les conditions sémantiques mixtes (C.BK) et (C.KB) citées dans [Hintikka 1962] deviennent intuitivement correctes lorsqu'elles sont interprétées en termes d'implications doxastiques ou épistémiques : 'je crois que ϕ' implique 'je sais que je crois que ϕ' seulement si la personne référée par 'je' a *déclaré* au préalable que ϕ ; dans ce sens performatif, l'implication devient légitime mais ne l'est pas tant que l'implication est considérée comme simplement locutoire (porte sur le contenu des énoncés).[232] L'autre condition mixte (C.BK) est d'autant plus défendable épistémiquement, puisque 'je sais que ϕ' implique 'je crois que je sais que ϕ' aussi bien aux sens locutoire et illocutoire de l'implication.

L'analyse performative des attitudes épistémiques permet de justifier aisément la thèse d'entraînement ou de conséquence, selon laquelle toute connaissance implique une croyance

---

[230] [Hintikka 1962], p. 55 ; nous soulignons. L'introspection ou conscience de soi est une conséquence de l'acte public d'assertion, étant donné que personne ne peut ignorer ce qu'il dit ; en revanche, certaines pensées inexprimées peuvent demeurer à l'état inconscient, et c'est pour cette raison que **5** n'est pas retenu comme caractéristique de K : l'antécédent $\neg$K n'indique pas un acte d'assertion mais une absence d'assertion. Rien d'explicitement *reconnu* ne peut être établi sur la base de $\neg$K, donc.

[231] La tendance fautive à dissocier l'introspection de l'acte préalable d'assertion est rappelée par Hintikka en termes du 'sophisme du psychologue' (attribué à Williams James in *The Principles of Psychology: The Long Course* (New York, 1890), I, 196-197, qui consiste à "identifier son point de vue 'avec celui d'un *fait* mental' (plutôt qu'à celui du *sujet* dont le savoir ou la croyance est examiné)".(p. 86; nous soulignons). Cette objectivation des états du sujet en termes de faits mentaux a pour effet d'assimiler nos propres états de connaissance à ceux des autres, mais en oubliant entre-temps la place intermédiaire et centrale que tient l'assertion dans la prise de conscience des états mentaux. Le rôle de l'*énonciation* et de l'implication *épistémique* dans l'analyse des actes mentaux est de mettre tous ces effets secondaires en évidence.

[232] [Hintikka 1962] conclut de son analyse que "l''idée fausse selon laquelle '$B_a p$' devrait impliquer '$K_a B_a p$' dans un sens plus fort peut être partiellement (ou totalement) due à une confusion entre les différentes implications" (pp. 84-85), c'est-à-dire entre $B\phi \rightarrow KB\phi$ et $K(B\phi \rightarrow KB\phi)$.

correspondante : $K\phi \rightarrow B\phi$. En effet, l'absurdité de (4) vient de ce que l'énoncé '$\phi \wedge \neg B\phi$' est épistémiquement indéfendable : (4KB) est indéfendable et sa négation est donc auto-confirmante : $\neg K(\phi \wedge \neg B\phi) \equiv K(\phi \rightarrow B\phi) \equiv (K\phi \rightarrow KB\phi) \equiv (K\phi \rightarrow B\phi)$. Etant donné que K et B sont deux modalités fortes $\Box$ logiquement indépendantes (elles ne sont pas interdéfinissables), la relation de conséquence entre K et B est un axiome et non un théorème, il doit révéler une signification de base et non dérivée. Bien que la définition classique (platonicienne) du savoir soit en mesure d'étayer $K\phi \rightarrow B\phi$ (puisque la croyance est une condition nécessaire à la connaissance, en plus de la vérité et de la justification), elle ne permet pas de surmonter les objections précédentes de [Radford 1967], contrairement à l'analyse illocutoire qui suit : si la connaissance correspond à un acte d'*assertion* et ne s'exprime pas par un simple acte de conjecture ou de supposition (ce qui est le cas du candidat chanceux de Radford), alors toute déclaration de connaissance implique la croyance puisqu'une assertion est une force illocutoire de degré supérieur à la simple croyance. Tout ce qui est asserté est supposé vrai *a fortiori*, d'où l'apparence paradoxale d'une formule telle que (4) et l'intérêt de son interprétation en termes de (4KB).[233]

A noter toutefois que, selon certains, *déclarations* de connaissance et *déclarations* de croyance s'excluent mutuellement : celui qui affirme 'je sais' n'affirmera pas 'je crois', ($K\phi \rightarrow \neg B\phi$) ; celui qui affirme 'je crois' n'affirmera pas 'je sais' ($B\phi \rightarrow \neg K\phi$). [Lehrer 1968] fait état de cet argument contre la thèse d'entraînement, ainsi que [Blanché 1968] : si je dis 'je crois que ϕ', ma déclaration de croyance suggère que je ne sais pas que ϕ, sans quoi je dirais directement que je sais que ϕ ; puisque l'énonciation 'je ne le crois pas, je le sais' est courante, elle semble montrer également que la signification dans l'usage de B et K exclut mutuellement leur vérité. Mais appliquer cet argument, c'est confondre encore une fois vérité et affirmation de vérité, ce qui revient à confondre les propriétés des énoncés avec les propriétés des énonciations. On peut donc rejeter ces deux arguments contre la thèse d'entraînement, et pour deux raisons liées à la rhétorique du langage ordinaire : dire 'je ne crois pas que ϕ, je le sais' ne signifie pas que l'on ne croit pas que ϕ au sens strict, mais que l'on n'est pas *seulement* en mesure de le croire. L'usage fait ici de la négation n'a pas pour effet de nier ϕ mais d'atténuer l'importance de la déclaration de croyance dans le discours : son rôle n'est pas d'affirmer la fausseté de ϕ mais, tout au contraire, de produire l'effet d'une *litote* et de renforcer ainsi la vérité ϕ. [Lehrer 1968] compare cet usage de la négation à l'énoncé 'Ce n'est pas une maison, c'est un manoir', sous-entendu que ce qui est un manoir est *a fortiori* une maison et ne l'exclut en rien. En conséquence, celui qui sait que ϕ ne fait pas que croire que ϕ mais le croit dans le même temps. Cet effet rhétorique des expressions épistémiques, que Lehrer attribue à la rhétorique et distingue de la logique des notions de croyance et de connaissance, est lié également à ce que [Blanché 1968] considère comme un

[233] Dans le même sens que ce que nous avons suggéré plus haut (note 216), [Åqvist 1964] exprime ceci par des indices d'engagement gradués : $\Box_5$ désigne une supposition ou simple croyance, tandis que $\Box_{10}$ désigne une assertion ; d'après lui, l'analyse correcte de (4) donne le sens interne contradictoire : $\Box_{10}\phi \wedge \Box_5\neg\phi$, en violation avec la règle $\Box_{10}\phi \rightarrow \Box_5\phi$ (contrepartie de notre règle d'entraînement). La converse est défendable, en revanche : $\Box_5\phi \wedge \neg\Box_{10}\phi$ signifie que l'on croit quelque chose sans en avoir la preuve ou la certitude et correspond à notre formule (4BK). Le cas est différent avec la version (5BK) $\Box_5\phi \wedge \Box_{10}\neg\phi$: on ne peut pas croire ce dont on asserte en même temps la fausseté. Mais ce cas correspond à $B\phi \wedge K\neg\phi$ et diffère de notre règle de conséquence.

principe d'économie informationnelle : celui qui *dit* croire que ϕ sous-entend par la force illocutoire modérée de son énonciation (une simple supposition) qu'il ne sait pas si ϕ est le cas. Mais en dehors de cette règle de pertinence informationnelle à la Grice (ce qui est dit doit révéler l'état psychologique précis du locuteur et ne pas faire usage de circonvolutions), une déclaration de croyance n'est pas logiquement incompatible avec une déclaration de connaissance. Une personne certaine de la vérité de ϕ a le droit de déclarer qu'il croit que ϕ est vraie, ou bien l'inverse, de sorte que les déclarations 'je crois que ϕ' ou 'je sais que ϕ' peuvent impliquer (épistémiquement) 'je sais que je crois que ϕ' ou (doxastiquement) 'je crois que je sais que ϕ' ; cela étant, la règle de pertinence ci-dessus incite simplement à associer de préférence (par souci d'économie des informations au cours d'un dialogue) une implication doxastique à toute énonciation de croyance, une implication épistémique à toute déclaration de connaissance. Mais hormis cette règle implicite, la combinaison des deux sortes d'implication pragmatique avec K et B reste logiquement défendable.

Toutes les autres expressions qui suivent peuvent être analysées et validés ou invalidées par le même type d'explication pragmatique de K et B : 'je crois que ϕ, mais je ne le sais pas' et 'ϕ, mais je ne sais pas si ϕ' sont deux variantes de (4BK) ; quant à 'si ϕ, alors je crois que ϕ' et 'si ϕ, alors je sais que ϕ', elles illustrent la négation de toutes les versions indéfendables de (4) où la force illocutoire est au moins égale à celle indiquée dans l'énoncé, c'est-à-dire (4B), (4K), (4KB) et (4KK). La relation d'implication entre déclaration de vérité et déclaration de savoir est formulée ainsi par l'implication épistémique de l'énoncé, 'si ϕ, alors je sais que ϕ'; bien que cette implication ne soit pas acceptable dans un cadre locutoire, sous peine de trivialisation du concept K, elle prend tout son sens dans un cadre illocutoire et représente notamment la conséquence que [Hocutt 1972] avait rejetée par l'absurde tout à l'heure (section 2.2.2) : ($K\phi \rightarrow KK\phi$) est effectivement une conséquence de ($\phi \rightarrow K\phi$), mais c'est une conséquence *pragmatique* ou illocutoire, au sens où l'acte d'énonciation 'si ϕ, alors je sais que ϕ' ($\phi \rightarrow K\phi$) implique *épistémiquement* ($K\phi \rightarrow KK\phi$) : de la déclaration ($\phi \rightarrow K\phi$) on infère $K(\phi \rightarrow K\phi)$, ce qui équivaut à ($K\phi \rightarrow KK\phi$). Si l'on distingue l'implication logique par la présupposition ou implication pragmatique (implicature), on symbolisera cette relation d'inférence particulière par 'ϕ' ⇨ '$K\phi$', c'est-à-dire $K(\phi \rightarrow K\phi)$. Cette dernière implication, illocutoire, est décrite en termes de mondes possibles par l'appartenance de $\phi \rightarrow K\phi$ à tous les ensembles modèles compatibles avec la connaissance du sujet-locuteur. La condition (C.KK*) ne justifie donc pas la thèse cartésienne de l'illumination de soi [*self-illumination*] ou transparence de la connaissance, contrairement à ce que dit Hocutt ; elle exprime la transparence des *déclarations* de connaissance, uniquement.[234]

Si l'énonciation 'ϕ, donc je sais que ϕ' est *épistémiquement* auto-confirmante, elle l'est *a fortiori* sous une forme affaiblie telle que 'ϕ, donc il se peut que je sache que ϕ' : $\phi \rightarrow PK\phi$, et sous une forme négative '$\neg\phi$, donc je sais que je ne sais pas que ϕ'. La première formule, même si elle annonce le cas du paradoxe de Fitch (voir section 3.1.1.2), ne se confond pas

[234] A noter que [Hintikka 1962] propose une certaine réfutation de (4B) qui ne recourt pas à (C.BB*) : lorsque le paradoxe de Moore est interprété par 'il pleut, mais je ne le crois pas ; et je crois ce que je viens de dire', de forme logique : $(p \wedge \neg Bp) \wedge B(p \wedge \neg Bp)$, la preuve d'indéfendabilité se contente des conditions plus faibles (C.B*) et (C.b*). Mais l'utilisation des conditions itératives (C.BB*) ou (C.KK*) permet d'insister sur le lien logique existant entre ($\phi \rightarrow B\phi$) ou ($\phi \rightarrow K\phi$) et $B(\phi \rightarrow B\phi)$ ou $K(\phi \rightarrow K\phi)$: les premières impliquent *doxastiquement* ou *épistémiquement* les secondes, donc elles justifient l'itération des opérateurs.

avec la définition anti-réaliste de la vérité comme connaissabilité ; ce n'est pas la vérité mais la déclaration de vérité d'un énoncé qui est toujours connaissable. La seconde formule semble entériner l'axiome **5**, mais seulement en termes d'implication épistémique : c'est le fait préalable d'asserter '$\neg\phi$' qui permet au locuteur d'en conclure qu'il sait que $\neg\phi$ et, par voie de conséquence, qu'il ne sait pas que ϕ. Or si l'énonciation de '$\neg\phi$' implique épistémiquement $K\neg\phi$, $\neg K\phi$ (par dualité) puis $K\neg K\phi$, le passage de $\neg K\phi$ à $K\neg K\phi$ n'est donc valable que pour les cas de déclaration explicite d'ignorance : dire que ϕ est faux, c'est prouver de fait que l'on ignore *a fortiori* la vérité de ϕ. Le traitement illocutoire de ces énoncés permet ainsi d'expliquer pourquoi leur énonciation apparaît intuitivement correcte, malgré la non validité de leurs énoncés dans [Hintikka 1962]. L'analyse performative d'une logique des déclarations de croyance et de connaissance justifie ainsi l'axiome d'introspection positive en termes d'énonciations et de force illocutoire : '$\phi \rightarrow K\phi$' et '$\phi \rightarrow B\phi$' implique $K(\phi \rightarrow K\phi) \equiv K\phi \rightarrow KK\phi$ et $B(\phi \rightarrow B\phi) \equiv B\phi \rightarrow BB\phi$; mais elle rejette celle de l'introspection négative : $\neg K\phi \rightarrow K\neg K\phi$ n'est pas une conséquence pragmatique de '$\neg\phi \rightarrow \neg\phi$', car une modalité niée telle que $\neg K$ ne présuppose pas un acte de discours préalable. La seule forme dans laquelle un analogue à **5** serait justifiable est $K(\neg K\phi \rightarrow K\neg K\phi) \equiv K\neg K\phi \rightarrow KK\neg K\phi$, qui ne donne rien d'autre par réduction que l'énoncé trivial d'identité $\neg K\phi \rightarrow \neg K\phi$.

De plus, l'équivalence entre B et BB ou K et KK permet d'éluder une objection de régression à l'infini des itérations : que signifie le fait de croire que l'on croit que l'on croit … que l'on croit, où B serait répété 10 fois ? La question est posée par [Williams 2000] et considérée par [Cargile 1967] comme une raison de rejeter les modalités itérées par l'absurde :

> Le problème, avec [l'axiome d'introspection positive], c'est qu'il a lui-même besoin d'une explication. En l'appliquant à plusieurs reprises, je peux dériver du fait que je crois que Johnson est le Président des Etats-Unis la conséquence que je crois que je crois que je crois que je crois que je crois que Johnson est le Président des Etats-Unis. Et il serait normal de vouloir savoir ce que cela *veut dire*. Peut-être que ce genre d'inférences peut être éliminé en restreignant le nombre de fois que l'axiome peut être appliqué à une croyance donnée, mais nous devrions nous demander alors la raison de cette restriction *arbitraire*.[235]

Or c'est l'inverse que permet l'argument d'itération, c'est-à-dire non pas une multiplication indéfinie des opérateurs mais la réduction à un seul : pour toute itération $\Box\ldots\Box_n\phi$ répétée n fois sur un énoncé ϕ, cette formule itérée équivaut en vertu de (C.$\Box\Box$*) à $\Box_1\phi$ ou $\Box\phi$.[236]

Quelques objections centrales ont été relevées précédemment et peuvent être résumées dans ce qui suit, en ce qui concerne la valeur sémantique des énoncés mooréens et leur analyse performative.

---

[235] [Cargile 1967], p. 183 ; nous soulignons.

[236] "Ceux qui supposent que le savoir implique le savoir que l'on sait ont été accusés parfois de générer une régression à l'infini (ou peut-être mieux, un progrès infini des niveaux ou ordres de savoir). Ils semblent postuler une infinité d'actes de savoir ayant lieu simultanément. Mais c'est l'inverse qui se produit, en réalité. Dire que le savoir (dans son sens de base) implique logiquement le savoir que l'on sait, c'est dire qu'une déclaration de savoir que l'on sait n'ajoute absolument rien à une déclaration de savoir ordinaire, qu'elle le fait simplement d'une manière détournée" ([Hintikka 1968], p. 14)

L'énoncé de forme (4) est-il aussi paradoxal pour les déclarations de croyance que de connaissance ; autrement dit, qu'y a-t-il d'absurde dans 'ϕ, mais je ne sais pas que ϕ' ou 'ϕ, mais je ne sais pas si ϕ' ? Bien que Hintikka signale que la version doxastique est plus paradoxale que la version épistémique (on peut affirmer des choses sans les savoir), cette dernière devient tout aussi répréhensible lorsque l'on considère une assertion comme une déclaration de savoir à part entière. Si aucune information disponible pour le locuteur ne l'incite à croire qu'il peut se tromper sur sa déclaration, alors il est en droit d'*asserter* ϕ, et pas simplement de supposer ou croire sa vérité. Pour cette raison, l'énoncé 'ϕ, mais je ne le sais pas' n'apparaît pas paradoxal lorsque la déclaration 'ϕ' manifeste un degré de force illocutoire inférieur à l'acte d'assertion : si le locuteur suppose simplement la vérité de ϕ, alors il ne déclare pas le savoir et peut se tromper à ce sujet.[237] En revanche, le locuteur ne peut pas admettre la possibilité de se tromper lorsque son énonciation a un degré de force illocutoire identique à celui de l'assertion : $K\phi \wedge P\neg K\phi$ est épistémiquement indéfendable, et c'est dans ce cas de figure précis que le paradoxe de Moore sévit aussi bien avec des déclarations de connaissance que de croyance. D'après [Hintikka 1962], le fait que les énonciations paradoxales de (4) et (5) portent aussi bien sur B que sur K montre que, contrairement aux propos de [Black 1952], le paradoxe de Moore n'est pas réductible à un problème d'*honnêteté* de l'énonciation :

> l'explication de Black ne nous permet pas de comprendre pourquoi [(4BK)] est logiquement étrange. (...) Ce ne peut pas vraiment être une présupposition de la notion d'assertion honnête que le locuteur *sache* ce qu'il dit ; on peut se tromper honnêtement. Par conséquent, [(4BK)] ne viole pas les présuppositions d'assertion honnête comme le fait (4B).[238]

Faut-il en conclure qu'une assertion doive satisfaire la condition de *véridicité*, en plus de la condition de *véracité* (ou de sincérité) ? Pas selon nous : une déclaration 'honnête' ne signifie pas uniquement que le locuteur croit à son contenu mais qu'il satisfait les conditions minimales de l'acte de discours *quel qu'il soit* : une déclaration de croyance est honnête si le locuteur croit à ce qu'il dit ; par extension, une déclaration de savoir est honnête si le locuteur exprime une croyance justifiée ; une déclaration d'amour est honnête si son auteur éprouve des sentiments sincères envers son interlocuteur ; une déclaration de guerre est honnête si elle exprime un état d'hostilité entre deux partis, etc. Il ne faut pas restreindre la notion d'honnêteté au discours de croyance, même si la condition de croyance est toujours incluse dans l'ensemble des actes de discours honnêtes. L'honnêteté signifie ainsi l'intention de dire la vérité, elle suffit en tant que telle pour expliquer le caractère faussement paradoxal de (4BK) : contrairement à ce que dit Hintikka, ce *doit* être une présupposition de la notion d'assertion honnête que le locuteur sache ce qu'il dit lorsqu'il effectue une déclaration de *connaissance*, d'où la distinction entre le faux paradoxe (4BK) et les véritables paradoxes (4B), (4K) et (4KB).

---

[237] C'est le sens donné à l'expression 'pouvoir se tromper' : pouvoir ne pas savoir, soit $P\neg K\phi$. Hintikka explique que l'on peut se tromper sur ce que l'on croit, mais par sur ce que l'on sait : $B\phi \wedge P\neg K\phi$ est défendable, mais $K\phi \wedge P\neg K\phi$ est indéfendable.

[238] [Hintikka 1962], p. 97.

Au total, notre explication des paradoxes mooréens repose sur une explication en termes de *degré de force illocutoire* d'un acte de discours : un acte de discours est paradoxal (impossible à satisfaire) lorsque la négation du contenu propositionnel est plus forte que son affirmation. On peut déclarer ignorer ce que l'on croit, parce que la déclaration d'une croyance est un acte d'assertion moindre que la déclaration d'une connaissance, par exemple. A l'inverse, un acte de discours n'est pas paradoxal lorsque la négation du contenu propositionnel est moins forte que son affirmation. On peut exprimer une croyance sans prétendre 'savoir', de même que l'on peut apprécier une personne sans prétendre l''aimer' ou 'aider' une personne sans prétendre la secourir. Dans tous les cas, la condition de vérité contenue dans la notion de véridicité est absente : c'est l'intention et uniquement l'intention du locuteur qui détermine le succès de son entreprise. Les implications doxastiques et épistémiques dont parle [Hintikka 1962] sont l'expression formelle de cette règle du degré de force illocutoire. Elle engage une réflexion sur la nature des actes de discours, sur l'intention qui les caractérise et, surtout, sur la nature des déclarations de connaissance et leur relation aux déclarations de croyance. Leur but est-il d'exprimer un état mental plus ou moins fort (de la simple conjecture à la certitude), ou la connaissance reste-t-elle irréductiblement liée à une condition objective de connaissance ? Notre élimination de la condition de véridicité en faveur de la seule condition de véracité indique notre penchant pour une assimilation de la connaissance à l'état mental de certitude ; mais une réduction totale n'est pas envisageable étant donné la diversité des usages du concept (de ses 'jeux de langage'). Parce que la notion de connaissance est partagée entre la croyance vraie justifiée et la certitude, il semble difficile sinon impossible de lui associer une 'signification de base', c'est-à-dire une seule d'intention d'usage caractéristique.

Peut-on réduire ainsi la psychologie à une pragmatique des actes de langage, le discours des états mentaux à un discours d'assertions ? Il existe après tout des croyances *implicites* qu'une tierce personne peut 'déduire' sur la base du [*gather from*] comportement du sujet croyant : ses désirs ou ses actions doivent permettre d'en conclure un certain état mental chez lui. Mais selon nous, l'attribution d'une croyance sur la base d'un comportement implicite n'apporte pas une transparence aussi nette des états mentaux que les actes de discours explicites. Sans en venir immédiatement au contre-exemple de la duperie de soi ou conscience aveugle [*self-deception*], dans lequel le sujet fait des déclarations dont il persiste à contester la vérité, celui-ci peut être considéré comme un cas limite de la psychiatrie et ne révèle en rien un exemple typique de croyance rationnelle, ni ne révèle la signification de base des actes de discours assertoriques. De même que la pensée est comparable à un acte de discours intérieur,[239] l'analyse des attitudes épistémiques en termes d'actes de discours explicites évite d'attribuer des états mentaux sans justification à l'appui. C'est notamment une attribution de croyance d'un agent à un autre et sans déclaration *explicite* à l'appui qui crée le problème d'opacité référentielle en logique épistémique quantifiée, comme nous le verrons plus loin.

L'avantage de l'analyse des modalités épistémiques en termes illocutoires d'assertion est qu'elle apporte une condition *suffisante* à la validité des énoncés : l'occurrence d'un état mental chez *a* est garantie dès qu'un acte de discours a été effectué par *a*. Sur la base de cette

[239] "Une tentative intéressante dans cette direction a été faite par Peter Geach [in *Mental Acts*, London 1957, pp. 75ff.]. Penser, de ce point de vue, n'est pas tant un monologue intérieur qu'un *dialogue intérieur* ; un point de vue anticipé par la formule de Platon, 'dans l'acte de penser, l'âme s'entretient avec elle-même'. Des idées étroitement similaires sont exprimées également par C. S. Peirce". ([Hintikka 1962], pp. 92-93).

réduction du psychologique au performatif, les énoncés de savoir ou de croyance deviennent intuitivement évidents et ne font plus appel à des conditions de justification discutables (sont-elles internes ou externes à l'agent ?). Dans l'analyse performative, la justification est à la fois externe et interne : interne, puisque le sujet locuteur est conscient de ce qu'il croit ou sait lorsqu'il le dit explicitement ; externe, puisque le locuteur fait savoir aux autres ce qu'il sait ou croit. L'implication épistémique ou doxastique permet ainsi de *présupposer* l'existence d'un état psychologique chez le locuteur, conformément à la condition de [Searle 1969] évoquée plus haut, et elle garantit en termes d'énonciations les conditions (C.BB*) et (C.KK*) parmi les propriétés caractéristiques des *énoncés* de croyance et de savoir.[240]

Si toute assertion correspond à une déclaration de connaissance, combien de nuances faut-il introduire entre les actes de tenir pour vrai, d'accepter, d'affirmer, d'asserter, de croire, de connaître, de supposer ? La connaissance a été assimilée ci-dessus à une assertion, et la croyance à un degré de force illocutoire inférieur. La première exclut la fausseté de l'énoncé asserté, contrairement à la déclaration de croyance. Sur la base d'une distinction entre asserter, tenir pour vrai et accepter, [Engel 1998] prétend notamment que l'on peut faire des énonciations sans force assertive correspondante : la vérité d'un énoncé pourrait être acceptée sur la foi de témoignages partiels ou de suppositions générales mais, dans ce cas, accepter n'équivaut plus à croire, et le sens interne de (4) (c'est-à-dire (4B)) ne donne plus une contradiction. La remarque d'Engel instaure des degrés de force illocutoire variables entre les actes d'énonciation, et à juste titre ; mais prononcer la vérité d'un énoncé sans être en mesure de le croire soi-même présuppose toutefois qu'une déclaration n'exprime une croyance que si elle s'accompagne d'une justification minimale. Toute croyance est-elle à justifier, et l'analyse des attitudes épistémiques en termes de déclarations est-elle insuffisante pour déceler la nature précise de ces attitudes ? Quoi qu'il en soit de notre approche formelle, la solution ci-dessus au paradoxe de Moore ne prétend pas à une applicabilité sans faille (des significations résiduelles existent pour les déclarations de connaissance et de croyance, telles que la duperie de soi, le mensonge ou la conjecture) et son rôle consiste seulement à capter un sens adéquat dans lequel les modalités épistémiques sont utilisées dans notre discours ordinaire. Compte tenu de l'efficacité de l'explication pragmatique ou performative, nous orienterons une bonne part de la logique modale épistémique vers une logique des actes de discours [*logic of knowing and believing*], conformément au fameux mot d'ordre de [Russell 1905] concernant la fonction de l'analyse logique : apporter une solution à des paradoxes, et déterminer l'origine conceptuelle du problème.[241]

Au total, nous avons distingué trois manières de défendre les axiomes d'introspection : en termes psychologiques d'états mentaux (conviction ou supposition, assertion ou assomption), en termes cognitifs de justification, puis en termes performatifs ou pragmatiques d'actes de discours. L'argumentation psychologique a été utilisée en large partie pour des usages

[240] L'idée de présupposition d'existence véhiculée par l'acte d'énonciation fut reprise notamment par Hintikka pour réinterpréter la preuve du *cogito* de Descartes : la preuve de l'existence proviendrait de la même source que le paradoxe de Moore, dans la mesure où 'je pense, donc je suis' devient en quelque sorte 'je dis quelque chose, donc j'existe'. Voir [Hintikka 1985, 1990], ainsi que [Mitton 1972] pour un commentaire de cette interprétation performative du *cogito*.
[241] Voir section 2.1.2.2, note 46.

résiduels, selon Hintikka : les psychologues insistent sur la signification de K en termes de conscience et, par l'emploi d'un sophisme dit du psychologue (vu plus haut), déduisent à partir des déclarations de discours que les actes mentaux sont des faits objectifs ou que KK incarne la thèse introspective de la conscience de soi. Les épistémologues inclinent pour certains d'entre eux (les internalistes) vers la signification de K en termes de certitude et, pour cette raison, considèrent la thèse KK comme l'expression du caractère décisif du savoir, un savoir fort et incontestable. Mais que ce soit en termes psychologiques de conscience ou en termes cognitifs de justification, nous avons montré dans cette section que le fait de présupposer des actes de discours pour l'analyse des modalités épistémiques rend nos intuitions plus claires à leur sujet : c'est la présupposition d'une déclaration de savoir ou de croyance qui permet d'impliquer la présence d'une preuve décisive ou d'un acte de la conscience, tandis qu'aucune de ces deux dernières conditions classiques du savoir ne précède l'acte d'assertion. La suite de la thèse va tenter de montrer ainsi le rôle primordial d'une analyse illocutoire des modalités épistémiques, aussi bien dans l'application de ses axiomes qu'en vue de la résolution des paradoxes épistémiques.

A commencer par la clause de vérité : pour quelle raison le concept de connaissance fait-il partie de la classe des verbes ayant une 'grammaire de succès', c'est-à-dire tels qu'ils valident la formule $\Box\phi \rightarrow \phi$?

2.2.4.3 Un seul axiome 'évident' : la clause de vérité ?

Supposons qu'une implication de forme $\phi \rightarrow \psi$ soit considérée comme vraie. Elle peut l'être pour deux raisons distinctes : parce que l'énoncé ϕ est faux et que tout peut être impliqué de la fausseté, donc ψ ; parce que ϕ est vrai et que, en vertu des propriétés logiques de ϕ, ψ doit l'être aussi. C'est ce dernier sens de l'implication qui nous intéressera ici. Deux question seront posées : pour quelle raison dois-je déduire la vérité de ϕ de la vérité de $K\phi$? Si l'on admet cette conséquence logique, que veut-on dire par le fait que l'énoncé ϕ est 'vrai' ? Ce sont les conditions nécessaires *et suffisantes* à l'application de la clause de vérité qui vont être discutées ici.

A première vue, il n'y a pas d'axiome plus 'évident' que $K\phi \rightarrow \phi$, dans la mesure où aucune personne ne dirait qu'elle sait quelque chose de faux : $\neg\phi \rightarrow \neg K\phi$. [Lemmon 1959] y voit le seul axiome acceptable pour la logique modale épistémique et rejette tous les autres, y compris l'introspection positive.[242] Si l'interprétation 'quasi-performative' de Hintikka nous a paru la plus à même de justifier $K\phi \rightarrow KK\phi$ sans entrer dans des considérations toujours discutables (internalisme versus externalisme), elle avait permis également de proposer une formalisation de l'aspect pragmatique du langage. Dans cette perspective, l'objection de Lemmon à l'introspection positive peut être éliminée. Mais nous allons voir que l'avantage de cette analyse pragmatique du schéma **4** crée un problème lorsqu'elle applique à l'axiome **T**.

La distinction essentielle entre croyance et connaissance a été fondée sur ce dernier axiome, puisqu'il est possible de croire quelque chose de faux, et la logique modale

[242] "Absolument aucun des autres principes n'est valable. Dans le cas de l'axiome [**K**], X peut très bien savoir qu'un conditionnel est vrai, et savoir que son antécédent est vrai sans en tirer la conclusion, et donc sans savoir que le conséquent est vrai. L'axiome [**4**] (...) a été sollicité par des philosophes, il est vrai. Mais il y a un sens clair dans lequel il n'est pas valable : il y a un grand nombre de choses que les gens savent sans *réaliser* qu'ils les savent" ([Lemmon 1959], p. 38 ; nous soulignons).

épistémique établit cette distinction par l'idée selon laquelle l'opérateur B n'est pas caractérisable par la classe des structures de type réflexif ; mais selon [Girle 1974], il serait possible également de falsifier des connaissances. Cela ne revient-il pas à rejeter la validité de la formule $K\phi \rightarrow \phi$? Girle déclare ainsi qu'aucune preuve *logique* n'est capable de justifier à elle seule une déclaration de savoir, ajoutant que

> les déclarations de savoir peuvent être falsifiées. Cela tient à la structure du concept et s'explique par $[K_a\phi \rightarrow \phi]$, ainsi que les autres axiomes et règles. Mais quant à ce qui justifie réellement les déclarations de savoir, cela ne peut être laissé au mieux qu'à ceux qui s'intéressent à la sociologie de la connaissance.[243]

Cela dit, Girle ne veut pas dire ici que la formule **T** est invalide en logique épistémique : une déclaration de savoir exclut la fausseté *par définition*, mais le contenu de cette déclaration (l'énoncé supposé connu) peut toutefois être réfuté par la suite et ne constitue une forme de connaissance que tant qu'aucune preuve n'a été apportée en faveur de sa négation. L'absence de preuve décisive pour un énoncé de ce genre

> ne veut pas dire qu'il n'y ait aucune possibilité de justifier une déclaration de savoir, mais que nous ne pouvons pas prouver déductivement qu'une affirmation que *a* sait que ϕ est vrai à partir des prémisses de croyance, de preuve et de vérité factuelle. Cela veut dire également que critiquer le concept de connaissance sur la base du fait qu'il n'y ait pas de conditions suffisantes pour la connaissance revient à critiquer dans le vide.[244]

Dire que le contenu d'une déclaration de savoir est falsifiable peut vouloir dire deux choses distinctes : que la formule $(K\phi \rightarrow \phi)$ n'est pas auto-confirmante, donc que sa négation $(K\phi \wedge \neg\phi)$ est satisfiable ; que $(K\phi \wedge \neg\phi)$ est indéfendable mais qu'une preuve de $\neg\phi$ peut être trouvée indépendamment, auquel cas le sujet doit reconnaître qu'il ne sait *finalement* pas que ϕ. Nous rejetons la première explication, mais nous admettons la seconde. Or c'est cette seconde interprétation que Girle adopte ici. On peut formuler autrement le problème des connaissances falsifiables : la possibilité de réfuter ϕ porte-t-elle sur le savoir que ϕ en tant que tel, ou sur la déclaration de savoir de ϕ ? Dans notre approche performative de la connaissance, notre avis serait qu'il est impossible de dire 'je sais que ϕ' tout en concédant que l'on puisse de tromper : '$K\phi \wedge \neg\phi$', et que l'argument de Girle ci-dessus consiste à admettre la faillibilité d'une connaissance en tant que telle, c'est-à-dire hors de toute déclaration dans le cadre d'un discours assertif. Nous allons diviser l'analyse de la clause de vérité en deux approches : d'une part, examiner les conditions dans lesquelles une *connaissance* peut être considérée comme justifiée et impliquer la vérité ; d'autre part, mesurer les conséquences d'une *déclaration de connaissance* sur le sens de la notion de vérité. Nous évoquerons dans ce dernier cas un paradoxe issu de cette justification pragmatique de la clause de vérité.

Bien avant la sémantique relationnelle, la distinction entre connaissance et croyance est consacrée dans le *Théétète* de Platon : le savoir ou connaissance correspond à une croyance

---

[243] [Girle 1974], p. 372.
[244] *Ibid.*

vraie justifiée, formalisée $K\phi =_{df} B\phi \wedge \phi \wedge J\phi$, dans laquelle on constate que la connaissance *présuppose* entre autres la croyance. Elle présuppose aussi et surtout la vérité,[245] et c'est sur le sens attribué à cette vérité que porteront les objections ci-dessous. Bien que $J\phi$ ne soit pas un opérateur utilisé dans le langage-objet de la logique épistémique, il sert implicitement à distinguer la simple croyance vraie de la croyance vraie justifiée, et c'est sur cette distinction que Hintikka a appuyé notamment son interprétation du 'savoir fort' des philosophes : la justification qui accompagne la croyance vraie (le *logos*, chez Platon) est implicitement présente dans l'axiome d'introspection positive, puisque celui qui connaît doit être capable d'apporter une preuve adéquate en faveur des déclarations de connaissance auxquelles il croit. Les contre-exemples de [Radford 1966,1967] à l'introspective positive et la thèse d'entraînement furent considérés comme un type de croyance vraie qui ne retranscrit pas la signification 'de base' de K.

Or bien que la clause de vérité semble aller de soi, elle s'est heurté à des objections liées à son appellation : comment garantir au juste la vérité de ce que quelqu'un prétend connaître ? Ce n'est certes pas au logicien de justifier la vérité des énoncés élémentaires, mais la question ci-dessus concerne la possibilité de donner un sens pratique à l'axiome **T**. La suite porte sur la relation entre justification et vérité, notamment. Comment s'effectue la transition entre la justification d'un énoncé et sa vérité, c'est-à-dire dans quel sens peut-on dire d'un énoncé qu'il est vrai, et pourquoi ce sens implique-t-il qu'il est impossible de connaître quelque chose de faux ? Lorsque philosophes et logiciens avancent des objections contre la clause de vérité, leur critique porte plus précisément sur la relation d'ordre entre les notions de connaissance, de justification et de vérité. Nous considérerons tout d'abord les sens éventuels dans lesquels il serait impossible de savoir quelque chose de faux, avant de constater le lien créé par ce débat entre les schémas **T** et **4** ; puis nous aborderons le paradoxe de [Zemach 1969], qui porte sur la relation conceptuelle entre vérité et connaissance. Ce dernier point sera mis en valeur, parce qu'il concorde avec l'interprétation *pragmatique* des opérateurs modaux que nous avons retenue plus haut dans la signification de base de K.

Comment contester la validité de $(K\phi \rightarrow \phi)$? Deux attitudes sont possibles face à cette implication : si on la récuse en admettant la possibilité de sa négation : $(K\phi \wedge \neg\phi)$; soit on l'admet et l'on en conclut par l'absurde l'impossibilité de connaître quoi que ce soit : $[(K\phi \rightarrow \phi) \wedge \neg\phi] \rightarrow \neg K\phi$. Cette dernière attitude est reprise par les sceptiques : s'il est impossible d'affirmer absolument la vérité d'un énoncé quelconque, alors cela signifie que nous ne savons absolument rien. Pour éviter un malentendu créé par les sceptiques, [Lemmon 1959] accuse d'abord ces derniers de dévoyer le véritable sens de la clause de vérité : lorsque le logicien épistémique évoque cet axiome, il songe à

[245] Sur le raisonnement par contraposition : $(K\phi \rightarrow \phi) \Rightarrow (\neg\phi \rightarrow \neg K\phi)$, voir note 152. [Hintikka 1962] fait un rapprochement similaire avec Parménide et cite également le *Gorgias* : "Je suis tenté de qualifier [(C.K)] de 'loi de Parménide'. Le vieil homme semble l'avoir reconnu lorsqu'il disait que 'tu ne peux pas savoir ce qui n'est pas (cela est impossible), ni le prononcer' (...) Il est intéressant de constater qu'il était très naturel pour Parménide de passer de cette idée (valide) à sa fameuse thèse selon laquelle 'ce qui peut être exprimé et pensé doit nécessairement être' (...) tout ce qu'il avait à faire, c'est remplacer le verbe pour le savoir (*γιγνωσκω*) par un verbe pour la pensée (*νοεω*) et contraposer. La transition a été rendue particulièrement séduisante par les significations entrelacées des deux verbes grecs. Pour une autre reconnaissance précoce de l'impossibilité de la 'connaissance fausse', cf. Platon, *Gorgias* 454d" (note 7 p. 22).

> la vérité contenue derrière la formule 'si vous savez, vous ne pouvez pas avoir tort'. Mais les sceptiques considèrent celle-ci comme signifiant 'si vous savez que p, alors il est nécessaire que p', et concluent que si vous pouvez avoir tort vous ne savez pas : puisque rien n'est nécessairement le cas (ou très peu de choses), rien ne peut être connu (ou si peu). Mais la déclaration devrait être interprétée par 'il est nécessairement le cas que si vous savez que p, alors p' : la contraposition sceptique échoue. En effet, il existe (probablement) un grand nombre de propositions vraies, et chacune d'elles peut être connue.[246]

Autrement dit, l'erreur d'interprétation du sceptique résiderait dans la portée qu'ils accorde à l'opérateur de nécessité : la nécessité qu'une connaissance soit vraie ne serait pas exprimée par la portée étroite ($K\phi \rightarrow \Box\phi$), mais uniquement par la portée large $\Box(K\phi \rightarrow \phi)$. Si le sceptique disait vrai, alors son argument consisterait à montrer que l'extrême majorité de nos connaissances porte sur des faits empiriques révocables, ce qui lui permet de conclure par contraposition sur *sa* version de la clause de vérité (c'est-à-dire : $K\phi \rightarrow \Box\phi$) que les gens ne savent quasiment rien : ($\neg\Box\phi \rightarrow \neg K\phi$). On retrouve ici le même genre d'argumentation sceptique que celui évoqué au sujet de la clôture épistémique du système **K**.

L'objection sceptique aux axiomes modaux épistémiques nécessite deux remarques, à la fois sur le type de nécessité qui est associé à la connaissance et sur la signification de cette nécessité.

D'un côté, Lemmon semble avoir montré ci-dessus que le sceptique confond à tort la nécessité de la vérité d'un contenu de connaissance avec la nécessité de ce contenu lui-même : dans l'interprétation de portée large, la formule $\Box(K\phi \rightarrow \phi)$ signifie que ϕ est nécessairement vraie si elle est connue ; dans l'interprétation de portée étroite, la formule ($K\phi \rightarrow \Box\phi$) signifie que ϕ est nécessaire si elle est connue. Contrairement à ce que veut montrer le sceptique, Lemmon explique ainsi que la nécessité porte sur la formule entière du schéma **T** et non sur le contenu propositionnel de cette formule. A moins de considérer les interprétations de portée large et de portée étroite comme des formulations équivalentes, ce qui n'est pas notre cas ici, l'objection du sceptique est rejetée parce qu'elle est sans incidence réelle sur le théorème épistémique. C'est la grammaire de succès du concept de connaissance qui est nécessaire, et non le contenu d'une connaissance. Nous admettons d'autre part l'interprétation de portée large $\Box(K\phi \rightarrow \phi)$, dans la mesure où elle consiste à appliquer la règle de nécessitation sur un théorème épistémique admis.

D'un autre côté, la nécessitation de la clause de succès ci-dessus pose problème parce qu'elle introduit un autre opérateur de nécessité en plus de la nécessité épistémique K. Puisque la boîte $\Box$ désigne une modalité pure et sans interprétation particulière, la formule ci-dessus exige que nous lui trouvions une contrepartie dans la langue naturelle. Quelle peut être la signification précise de l'opérateur $\Box$, lorsque nous prétendons que la clause de succès est un 'fait nécessaire' et que l'on ne peut pas penser autrement ?[247] Pour autoriser la

---

[246] [Lemmon 1959], p. 38.

[247] "Si l'on sait, *peut*-on se tromper ? Il est tentant de répondre à cette question par 'non', voire d'utiliser cette exigence comme une caractérisation d'un certain sens suffisamment fort du savoir. Mais la question n'est pas très utile, cependant, du moins pas directement, car elle est sérieusement ambiguë. Le sens dans lequel 'peut' doit être conçu n'est pas clair, et il se peut même que l'interprétation recherchée de la question ne soit pas tout à fait claire" ([Hintikka 1968], pp. 149-50 ; nous soulignons).

nécessitation de la clause de vérité et confirmer ainsi le diagnostic de [Lemmon 1959], deux choses doivent être montrées dans ce qui suit : que ($K\phi \rightarrow \phi$) exprime une nécessité conceptuelle,[248] et qu'elle peut être considérée en tant que telle comme nécessairement vraie. Dans cette optique, $\Box$ ne désigne une modalité forte qui n'est ni logique, ni physique ; $\Box$ sera compris selon nous de la manière la plus appropriée qui soit dans $\Box(K\phi \rightarrow \phi)$ si on l'interprète comme la nécessité épistémique, mais cela reste à justifier. De cette façon, la défense de l'axiome **T** se fera sur le même type d'argumentation que la défense ci-dessus de **4**, à savoir l'interprétation pragmatique de K comme acte d'énonciation. Une telle réduction de la nécessité conceptuelle à la nécessité épistémique réhabilitera *en un certain sens* la formulation du sceptique, mais pas l'objection qu'il tentait d'imposer par son biais : remplacer $\Box$ par K fait de l'interprétation de portée étroite une formule trivialement vraie telle que ($K\phi \rightarrow K\phi$), mais le sens pragmatique que nous associerons à la nécessité épistémique permettra de montrer que cette formule n'est pas aussi triviale que sa forme logique le laisse paraître pour l'instant : asserter que tout ce qui est connu est vrai et asserter ce qui est connu sont deux actes de discours équivalents et, en ce sens, les interprétations de portée large et étroite de $\Box$ reviennent à dire la même chose lorsque la nécessité conceptuelle désigne un acte d'assertion ou de déclaration de vérité.

La justification de cette approche sera mise en évidence par le paradoxe pragmatique de la connaissance de [Zemach 1969], illustrant la relation ambiguë entre les notions de vérité et de connaissance : la vérité apparaît tantôt comme un *critère* préalable à l'acquisition d'une connaissance, tantôt comme le *symptôme* d'une connaissance acquise. La vérité est-elle un critère de la connaissance et la connaissance un symptôme de la vérité, ou bien l'inverse? C'est la difficulté de trouver une réponse à cette question qui sera reflétée dans le paradoxe de Zemach et résolue à nouveau par une distinction entre énoncé et énonciation, (déclaration de) vérité et (déclaration de) connaissance.[249]

Tout en laissant de côté pour l'instant le sens à donner à $\Box$ en vue de contrer l'argument du sceptique, précisons que le problème initial de ce dernier concerne la justification de la connaissance et la possibilité de ($K\phi \wedge \neg\phi$), c'est-à-dire $\Diamond(K\phi \wedge \neg\phi)$: dans quelle mesure est-il impossible de se tromper lorsque l'on sait que ϕ, $\neg\Diamond(K\phi \wedge \neg\phi)$? [Hintikka 1970a] s'interroge sur la forme logique appropriée de la question 'peut-savoir quelque chose et se tromper?' : "La dernière question est alors pour l'essentiel celle de savoir si nous traitons de (6) $\Diamond(K\phi \wedge \neg\phi)$ ou de (7) $K\phi \wedge \Diamond\neg\phi$",[250] et traite les deux formules sur un pied d'égalité. Contrairement à la remarque précédente de [Lemmon 1959], (6) et (7) seraient donc également acceptables *a priori* tant que l'interprétation de $\Diamond$ n'est pas fixée et ne signifie pas

[248] [Hintikka 1970] symbolise par M la 'possibilité logique' et par K la 'nécessité épistémique', par exemple. Bien que nous ayons toujours assimilé jusqu'ici les notions de 'possibilité' et de 'nécessité' à l'unique catégorie de la modalité logique (ou aléthique) faible et forte, nous admettrons un usage interchangeable des notions de nécessité/possibilité et de modalité forte/faible.

[249] A est un *critère* de B si A est une condition nécessaire de B, c'est-à-dire si B ne peut pas être vrai sans A : $B \rightarrow A$; A est un *symptôme* de B si A est une conséquence de B, c'est-à-dire si B peut être vrai sans A : $A \rightarrow B$. A peut-il être critère *et* symptôme de B sans devenir *ipso facto* équivalent à B : $A \equiv B$? Ces relations conditionnelles seront revues plus bas.

[250] [Hintikka 1970a], p. 150. Dans [Hintikka 1962], la section 4.16 évoque la possibilité de se tromper relativement à une croyance : la formule '$B\phi \wedge C\neg\phi$' est indéfendable, contrairement à '$K\phi \wedge P\neg\phi$'. La section présente va chercher à comprendre dans quel sens $K\phi$ implique $\neg\Diamond\neg\phi$.

la vérité dans *n'importe quel* monde possible. La formule ($K\phi \rightarrow \Box\phi$) signifie-t-elle que ce qui est su est vrai dans chaque monde logiquement possible, physiquement possible, épistémiquement possible ? Si $\Box$ désigne un autre opérateur que K ou même B, alors la tâche reviendra à une logique multimodale de s'interroger sur l'interprétation sceptique de la clause de vérité : $K\phi \rightarrow \Box\phi$.

Supposons par exemple que $\Box$ désigne la modalité aléthique, celle de la *nécessité logique* symbolisée par L. Alors l'objection à la clause de vérité du sceptique porte sur la consistance logique d'une déclaration de connaissance : si $K\phi$, est-il logiquement possible que $\neg\phi$ soit le cas ? Un moyen de maintenir la clause de vérité avec cette interprétation de $\Box$ en portée étroite consisterait à assimiler la connaissance à la connaissance logique ; si l'on suit cette voie, alors un contre-exemple à $K\phi \rightarrow L\phi$ est impossible à trouver dès lors que ϕ exprime une loi déductive. Cette interprétation est trop restrictive, puisqu'elle limiterait la validité de la clause de vérité aux seuls cas où ϕ est une vérité formelle. De façon générale, l'idée que toute connaissance soit d'ordre logique parce que connaissance et déduction expriment toutes deux des propriétés infaillibles est rejetée, de sorte que ni $K\phi \rightarrow L\phi$ ni $L\phi \rightarrow K\phi$ ne seront valides. Cela conduit Hintikka à rechercher une autre interprétation de $\Box$ que L afin de surmonter l'objection du sceptique : "Lorsque '$\Diamond$' est entendu comme exprimant la possibilité logique, (7) est un état de choses concevable".[251] Mais que dire de (6), et peut-on admettre qu'il soit logiquement possible de connaître quelque chose de faux ? (7) paraît bien moins contestable que (6) puisque l'on connaît un grand nombre de faits dont la fausseté est logiquement concevable. De deux choses l'une : soit la possibilité logique signifie la satisfiabilité dans un système formel, et (6) est indéfendable dans un système de logique épistémique où la clause de succès est associée à l'opérateur de connaissance ; soit la possibilité logique est autre chose qu'un synonyme de satisfiabilité dans un modèle, et nous devons rechercher le sens dans lequel (6) représente une affirmation inacceptable.

Supposons maintenant que $\Box$ désigne une autre sorte de modalité forte telle que la *nécessité physique*, ou *causale*. Si le sceptique dit vrai, alors il est possible qu'une connaissance soit réfutée par une loi de la nature ou, plutôt, que la vérité de ϕ ne soit garantie par aucune une loi naturelle. Nous ne serons pas plus loquaces que Hintikka sur cette interprétation particulière de la nécessité. Vu la tournure prise par la notion de connaissance jusqu'ici, pour quelle raison une étude de soi-disant 'lois de la nature' devrait-elle entrer en ligne de compte dans la signification de base du concept de savoir ? Ces lois suggèrent une relation de *causalité* et peuvent soit prêter à confusion, soit être sujets à caution. Y a-t-il un sens à s'interroger sur le lien d'une déclaration de connaissance avec des lois telles que la pesanteur ou la gravitation, lorsque l'énoncé ϕ de $K\phi \rightarrow \Box\phi$ porte par exemple sur un fait conceptuel abstrait, une opération mathématique ou une assertion métaphysique de type 'ma main existe' ? Sans parler des réticences de Quine sur la notion même de loi de la nature, Hintikka abandonne vite cette interprétation de la modalité forte :

[251] [Hintikka 1970a], p. 150.

> Le concept de possibilité causale ou naturelle impliquée ici est en manque si grave d'explication que la plupart des philosophes, y compris moi-même, douteraient qu'elle puisse aider ici à clarifier la situation.[252]

On pourrait reprocher à cet abandon de ne pas tenir compte des avancées réalisées entre-temps au sein de la logique philosophique : [Hendricks 2003] prolonge le plaidoyer de [Scott 1970] en faveur des logiques multimodales[253] et préconise l'introduction de modalités telles que la nécessité physique lorsqu'il s'agit de décrire les procédures de preuves scientifiques au sein de l'épistémologie formelle. Mais notre objection sera maintenue à l'égard d'une telle multiplication des modalités : nos objections portées contre la notion ambiguë de 'loi de la nature' ne s'évaporent pas sous le seul prétexte que des logiciens actuels feraient usage de multimodalités ou concevraient les énoncés de connaissance comme le résultat de conventions scientifiques. Certes, rien n'interdit *a priori* d'adopter cette réduction de la connaissance à des conventions scientifiques par la méthode du forçage [*forcing*] de [Hendricks 2003], en épistémologie formelle ; or si elle a au moins l'avantage de dissoudre les objections radicales du sceptique contre la validité de $(K\phi \rightarrow \phi)$, elle a néanmoins l'inconvénient de faire de la logique philosophique une logique *descriptive* sans significations de base particulières pour les opérateurs modaux. Il existe par ailleurs d'autres moyens plus simples de surmonter la validité de $\Box(K\phi \rightarrow \phi)$, conformément aux analyses précédentes de [Hintikka 1962] et au modèle explicatif de [Hintikka 1968].

Hintikka évoque notamment la notion de *nécessité conceptuelle* pour garantir la relation d'implication entre connaissance et vérité : la défense de la clause de vérité sous sa forme sceptique $K\phi \rightarrow \Box\phi$ ne reposerait ni sur une théorie de la vérité comme logique ou analytique, ni sur une définition de la vérité comme ensemble de conventions scientifiques, mais sur une acception particulière de la notion de connaissance. En d'autres termes, Hintikka prétend que c'est l'usage même du verbe 'savoir' qui présuppose (sans la justifier pour autant) la vérité. Il ne faut donc pas chercher une justification à $(K\phi \rightarrow \phi)$ ailleurs que dans l'usage purement linguistique de notre concept de connaissance. Cette approche rend justice à l'appellation de logique épistémique : il n'est pas question de rechercher une quelconque preuve indépendante ψ pour justifier la vérité de ϕ dans $K\phi \rightarrow \phi$. Tel était le risque encouru par une recherche des conditions matérielles suffisantes à la déclaration de vérité de ϕ dans $(K\phi \rightarrow \phi)$, jusqu'ici. Mais la référence à une connexion conceptuelle entre $K\phi$ et ϕ permet d'en rester à un niveau d'analyse strictement formel, sans tenir compte du contenu de ϕ et en pleine conformité avec la définition de la logique formelle comme indépendante du sujet (*topic-neutral*).[254] Par nécessité conceptuelle, Hintikka renvoie implicitement à l'analyse de la

---

[252] [Hintikka 1970a], p. 151. L'idée de nécessité physique ou causale a fait recette dans le cercle des philosophes analytiques australiens, parmi lesquels David Armstrong et Brian Ellis. Nous n'avons détaillé précédemment l'analyse des énoncés modaux par ce dernier, dans le cadre de sa sémantique des ensembles de croyance.

[253] Le discours de [Scott 1970] a valeur de manifeste en faveur des logiques multimodales : "Voici ce que je considère comme une des plus grandes erreurs parmi toutes en logique modale : la concentration sur un système doté d'un seul opérateur modal. La seule façon d'avoir des résultats philosophiquement significatifs, en logique déontique ou en logique épistémique, c'est de combiner ces opérateurs avec : des opérateurs de temps (sinon, comment pouvez-vous formuler des principes de changement ?) ; des opérateurs comme la nécessité historique ou physique (sinon, comment pouvez-vous relier l'agent à son environnement ?), et ainsi de suite" (p. 143).

[254] Pour preuve qu'un énoncé de connaissance peut être étudié logiquement sans tenir compte de son contenu, [Lenzen 1980b] propose notamment une analyse axiomatique de *De la Certitude* décrivant les relations

connaissance comme certitude que Wittgenstein proposa dans *De la Certitude* : si je suis certain de la vérité de certains énoncés, et suis autorisé à déclarer qu'ils constituent une connaissance absolue, c'est parce qu'ils correspondent à des règles grammaticales de mon langage. Il ne s'agit pas de justifier une déclaration de connaissance par une recherche de ce qui fonde sa vérité, mais de garantir cette vérité sur d'autres bases qu'une analyse de nos facultés de connaissance. Là où le sceptique veut justifier $K\phi$, Hintikka prétend que c'est l'opérateur K lui-même qui garantit la vérité de ϕ.[255] La certitude dont parle Wittgenstein est sans rapport avec la connotation psychologique qu'elle peut avoir chez Ayer, notamment, ou chez des internalistes tels que Prichard. Ici, la certitude caractérise un énoncé qui n'est ni vrai ni faux et constitue une sorte de règle grammaticale, en conséquence de quoi il ne *peut* faire l'objet d'aucune critique quant à sa valeur de vérité (il n'en a pas). Le moindre doute sur sa signification est considéré comme un non-sens, d'où la validation de la formule $\neg\Diamond(K\phi \wedge \neg\phi)$, contrairement à l'argumentation sceptique en vertu de laquelle une déclaration de connaissance $K\phi$ doit être justifiée par une évidence préalable ψ en faveur de ϕ. A proprement parler, Wittgenstein ne dirait donc pas de ces nécessités conceptuelles qu'elles expriment des énoncés 'vrais', dans la mesure où ne sont vrais et faux pour lui que les états de choses empiriques décrivant un fait du monde. A cet égard, la clause de vérité peut être défendue par une relation conceptuelle entre $K\phi$ et ϕ, mais $(K\phi \rightarrow \phi)$ ne signifie pas en tant que tel que toute connaissance porte sur un état de choses (un fait propositionnel) et ϕ ne représente pas à proprement parler un énoncé, si par énoncé on entend une phrase déclarative dont la fonction est de décrire un état du monde.

Une fois la parenthèse refermée sur le sens à donner aux valeurs de vérité (insister sur leur sens factuel ou *ontologique* constitue une des hypothèses de travail de cette thèse ; voir section 3.2.3), on s'aperçoit que Hintikka s'inspire clairement de la lecture wittgensteinienne de la connaissance lorsqu'il parle de la certitude de certains énoncés de connaissance de base : ce n'est pas au sens psychologique de Ayer qu'il entend la certitude, mais au sens conceptuel d'une connaissance non-contingente parce qu'indépendante de l'état du monde. Ses références à [Arner 1959] et [Malcolm 1961] nous donnent l'occasion furtive d'entrer dans un débat épistémologique sur la justification : quelques théories générales telles que le fondationalisme, le fiabilisme ou causalisme s'affrontent quant à ce qui permet de justifier la vérité d'un énoncé. Le fondationalisme fut défendu notamment par [Chisholm 1982], qui défendit la thèse d'introspection positive par sa thèse de l'auto-présentation [*self-presentation*] de certains faits à la conscience. Nous avons vu précédemment que Hintikka

---

d'inférence entre plusieurs concepts épistémiques tels que la croyance, la connaissance et la certitude. Par contraste, l'objection des sceptiques à la clause de vérité tente de montrer que $(K\phi \rightarrow \phi)$ n'est acceptable que pour une classe réduite d'énoncés ϕ. L'interprétation de $\Box$ comme nécessité conceptuelle a l'avantage de maintenir le caractère strictement formel de l'étude des axiomes épistémiques : $(K\phi \rightarrow \phi)$ est valide indépendamment du contenu de l'énoncé ϕ.

[255] La différence entre *justifier* et *garantir* est la clef pour comprendre la clause de vérité et surmonter le doute sceptique : "Le problème de la justification de notre connaissance est différent de celui de la garantie de notre connaissance. Le premier est typiquement issu du scepticisme moderne post-cartésien. On se demande ce qui *pourrait* justifier une connaissance *qui ne l'est pas*. A défaut de cette justification, notre connaissance n'aurait qu'une valeur cognitive limitée. Le problème de la garantie part du fait de la connaissance, et cherche à découvrir pourquoi notre connaissance est *garantie* par le processus même par lequel elle se met en place" (R. Pouivet, "Łukasiewicz : de l'aristotélisme autrichien à l'aristotélisme polonais", *Philosophiques* **26**(1999), p. 264 note 4). A ce titre, la connaissance n'est plus à proprement parler une croyance vraie *justifiée*, mais garantie.

rejette cette idée de transparence de la conscience pour garantir l'introspection positive, bien que son appellation ait le tort d'assimiler l'axiome **4** à une acception cartésienne ou psychologique de la connaissance. Le fiabilisme ou causalisme d'Alvin Goldman sera rappelé plus loin dans le cadre de l'argument de Gettier, puisque une connaissance est justifiée selon lui si elle est causalement liée par un processus cognitif suffisant. Hintikka ne retient aucune de ces théories de la justification et suggère une explication non-cognitive : l'usage de certains énoncés est garanti par le rôle qu'ils jouent dans l'organisation de notre connaissance, et aucun fait empirique n'est susceptible de les révoquer sous peine de non-sens. En résumé, la connaissance $K\phi$ de Hintikka et Malcolm ne présuppose pas la vérité absolue de ϕ et peut faire fi des arguments radicaux tels que l'argument cartésien du rêveur éveillé :

> personne ne peut raisonnablement se demander s'il rêve ou s'il est pris d'hallucination lorsque tout paraît normal. Cela revient à 'contester' les choses les plus évidentes. Qu'est-ce qui pourrait être considéré comme une preuve qu'il était en train de rêver ? Ou qu'il ne rêvait pas ?[256]

Dans [Hintikka 1970a], l'auteur donne l'exemple de la bouteille d'encre de [Malcolm 1961] dont je ne saurais douter de l'existence. La 'vérité' de l'énoncé 'il y a une bouteille d'encre sur la table devant moi' n'a pas un sens ontologique, ici : elle ne signifie pas la satisfaction d'un énoncé par un état de choses, car l'existence de l'objet perçu fait partie des présupposés que le Wittgenstein du *De la Certitude* préserve contre toute attaque du sceptique. Hintikka exploite cet argument issu du réalisme du sens commun de Moore pour rejeter tour à tour les objections empiriques citées dans [Dretske 1970] : si, lors d'une visite au zoo, je prétends savoir qu'il y a un zèbre devant moi, alors mon savoir implique que ce zèbre existe. Mais, répond Dretske, le gardien du zoo a très pu se jouer des visiteurs et déguiser un mouton en zèbre. Dretske en tire la conséquence que l'énoncé p : 'il y a un zèbre devant moi' peut très bien être faux à mon insu, auquel cas la contraposition $\neg p \rightarrow \neg Kp$ s'applique et réfute ma déclaration de connaissance. Or la connaissance ne signifie pas l'indubitabilité ontologique d'un fait déclaré, chez Hintikka, mais l'impossibilité de trouver une raison supplémentaire de douter d'une connaissance : $\neg\Diamond(K\phi \wedge \neg\phi)$.[257] L'analyse de la vérité par [Malcolm 1961] corrobore la définition qu'en donnait [Hintikka 1962], en ce qui

---

[256] [Arner 1959], p. 89. D'autres extraits confirment le caractère injustifiable de ϕ dans $(K\phi \rightarrow \phi)$: "En philosophie, l'existence de mes mains, de mon bureau et de ma bouteille d'encre ont tendance à devenir des sujets d'investigation. De telles exigences s'annulent tout simplement d'elles-mêmes : les fondements ostensibles sont si étendus que le fait de demander une preuve, quelque chose que l'on puisse admettre, témoigne que rien ne peut être autorisé. De telles exigences ne peuvent pas être satisfaites par une explication". ([Arner 1959], p. 89). "Aucune expression de doute n'est possible lorsque l'on sait, non pas parce que l'on a découvert en soi la confiance suprême mais parce que l'on dispose d'une évidence définitive, et que toute vérification et réserve ont perdu leur raison d'être. On ne peut pas savoir directement que l'on sait" ([Arner 1959], p. 91).

[257] La connaissance selon Hintikka n'a pas à satisfaire les exigences du doute radical des sceptiques : "Qu'est-ce qui est impliqué exactement dans l'exigence que les raisons de la connaissance au sens plein du terme soient *favorables* ? Pour ce que nous voulons faire, il suffit de signaler la conséquence évidente de cette exigence : si quelqu'un dit 'Je sais que p' au *sens fort* du savoir, il nie implicitement que toute information supplémentaire l'eût conduit à modifier son point de vue. Il s'engage dans l'idée qu'il persisterait à dire qu'il sait que ϕ est vraie –ou persisterait à dire tout au moins que ϕ est effectivement vraie –même s'il en savait davantage qu'il n'en sait maintenant" ([Hintikka 1962], pp. 20-21 ; nous soulignons). A cet égard, l'énoncé 'Il y a un zèbre devant moi' du visiteur du zoo n'est peut-être pas aussi concluant face à l'argument de l'illusion perceptive de Dretske que l'énoncé 'j'ai une main', puisque ce visiteur peut finir par découvrir le subterfuge du gardien et réviser sa pseudo-connaissance après coup. Tout dépend des conditions exigées pour la satisfaction de ϕ dans $K\phi \rightarrow \phi$.

concerne la défense de la thèse KK : savoir que ϕ se définit comme ce qui coupe court à toute discussion supplémentaire concernant la vérité de ϕ [*discussion-stopper*], et

> si 'possible' veut dire ici 'possible aussi loin que mon savoir puisse aller', nous avons précisément mon sens fort.[258]

On retrouve ici les raisons pour lesquelles les axiomes épistémiques ne semblent pas sensibles aux attaques des sceptiques, ni devoir entrer dans des complications formelles pour y parer : le rejet de l'axiome **T** peut être obtenu soit par le biais d'une modification de l'opérateur logique qui relie l'implication de $K\phi$ à ϕ, soit par la théorie du pistage [*keeping track*] et la sémantique des mondes proches de Nozick, dont la procédure est comparable à la sémantique de voisinage [*neighbourhood semantics*] de Scott et la sémantique des situations de [Barwise & Perry 1983]. Mais quelles que soient les controverses sur le sens de la clôture épistémique telle que l'interprètent les sceptiques (voir section 2.2.4.1), la littérature mêlant logique épistémique et scepticisme n'apparaît pas dans les écrits de Hintikka parce qu'elle n'affecte pas la signification de base de K et recourt inutilement à une sémantique des conditionnels contrefactuels.[259]

En conséquence, la relation logique entre $K\phi$ et ϕ est présentée comme une relation conceptuelle en vertu de laquelle le fait de dire 'je sais' doit entraîner l'impossibilité de contrarier cette déclaration au moment où elle est *prononcée*. L'argument du sceptique, s'il a le mérite de remettre sur le métier la question des conditions d'assignation de la vérité à un énoncé, ne remet pas plus en cause la validité de la clause de succès que des auteurs tels que Wittgenstein, Hintikka, Malcolm et Arner. Ce qui est connu est vrai par définition, diront-ils tous d'un commun accord : $(K\phi \Rightarrow \phi)$, mais les conclusions qu'ils tirent de ce principe sont mutuellement opposées. Le sceptique obtient par contraposition une conclusion inverse aux auteurs ci-dessus, car il déduit de notre absence de preuve de la vérité d'un énoncé que nous ne savons rien à son sujet : $[(K\phi \Rightarrow \phi) \wedge \neg\phi] \rightarrow \neg K\phi$. Quant aux auteurs précités, ils détachent de l'existence de certaines connaissances la garantie de certaines vérités conceptuelles : $[(K\phi \Rightarrow \phi) \wedge K\phi] \rightarrow \phi$. La différence réside dans leur interprétation de la notion de vérité : le sceptique parle uniquement de vérités de fait, au sens où un énoncé ϕ

---

[258] [Hintikka 1970a], p. 153. L'auteur caractérise la connaissance de ϕ comme la possession d'un argument décisif [*conclusive*] en faveur de ϕ : "Je ne suis pas en position de dire 'je sais' tant que mes raisons de le dire ne sont pas telles qu'elles me donnent le droit de négliger toute autre preuve ou information. Nous devons réaliser toutefois que le fait d'avoir ce droit ne signifie pas forcément que les raisons sont si fortes qu'elles impliquent *logiquement* que ce que l'on prétend savoir est vrai. Cela peut vouloir dire simplement que les raisons que l'on a sont telles que toute enquête supplémentaire serait inutile pour les *besoins normaux* des locuteurs du langage". ([Hintikka 1962], p. 20). Le rejet par les sceptiques de la clause de vérité s'explique par le degré de sévérité de ce qu'ils considèrent comme un argument 'décisif'. Voir par exemple la position sceptique de Peter Unger : "A defence of skepticism", *The Philosophical Review* **80**(1971), pp. 198-219, opposée au savoir contextualiste de David Lewis : "Elusive knowledge", *Australasian Journal of Philosophy* **74**(1996), pp. 549-567.

[259] Cette littérature se concentre sur l'utilisation sceptique des principes épistémiques **K** et **T**. Dans les deux cas, la stratégie du sceptique consiste à poser ces principes et en conclure l'impossibilité du savoir par contraposition. Sur la théorie contrefactuelle de la connaissance, voir notamment [Nozick 1981] ; sur les réponses à cette formalisation contrefactuelle de la théorie du pistage (savoir, c'est avoir une croyance qui suit la vérité 'à la trace', de sorte que l'on ne croi*rait* pas ϕ si ϕ *était* fausse), voir par exemple B. Garrett: "Keeping track of Nozick's trackers", *Ratio* (N.S) **1**(1992), ainsi que A. Brueckner : "Losing track of Nozick", *Ratio* (N.S.) **2**(1992), pp. 194-198.

correspond à un état de choses ; les autres auteurs parlent de vérités conceptuelles, au sens où l'énoncé en question n'entretient pas de relation d'ajustement avec le monde mais sert à caractériser la grammaire du concept de connaissance. La clause de succès se défend ainsi par des arguments de nature distincte puisque les anti-sceptiques s'intéressent ici à ce que le second Wittgenstein considérait comme des règles de langage, c'est-à-dire des énoncés non descriptifs et qui expriment l'usage approprié d'un concept.

Or si l'on admet cette interprétation de la clause de succès, le nombre des déclarations de connaissance s'en retrouve assez limité : si le droit de dire que l'on sait quelque chose repose sur la nature conceptuelle de la 'vérité' (non-ontologique) de ϕ, doit-on en conclure que tous les énoncés de type empirique ou simplement factuels ne seront jamais des contenus de connaissance à proprement parler ? [Arner 1959] précise bien qu'une déclaration de connaissance indubitable aujourd'hui peut être douteuse demain et se trouver réfutée après-demain, dans le cas où les conditions de justification disponibles seraient améliorées et permettraient d'augmenter les moyens d'expérimentation de la vérité en question. Mais cela n'interdit pas de parler de connaissance à part entière, toute provisoire soit-elle. Encore une fois, l'admission de la clause de succès ($K\phi \rightarrow \phi$) repose sur la fonction linguistique de verbe 'savoir', à partir duquel on peut conclure la vérité de ϕ sans en avoir la preuve manifeste. Pour corroborer cette interprétation de la clause de vérité, [Hintikka 1968] fait référence aux verbes dits 'parenthétiques' de [Urmson 1952], dont la tâche consisterait à garantir la vérité d'un énoncé par l'emploi préalable de certains concepts. On trouve dans cette catégorie verbale le concept de savoir :

> Urmson y déclare que la fonction de 'verbes parenthétiques' [*parenthetical*] tels que 'sait' et 'croit' est d'indiquer la situation de preuve dans laquelle une affirmation est faite (...) La condition essentielle (C.KK*) formule précisément cette caractéristique de la logique du savoir qui les stoppe. Sans elle, le concept de savoir ne servirait pas pleinement l'objectif de base qu'il est censé servir (...) dans le discours ordinaire.[260]

La même explication fondamentale est donc proposée pour les axiomes **T** et **4** : l'usage du concept 'savoir' et, à une moindre échelle, l'usage du concept 'croire' transmettent au locuteur une autorité sur le contenu de son énoncé, ce qui signifie que celui qui dit 'savoir' n'a pas à justifier la vérité de sa déclaration mais la *suggère* par son acte de discours. L'internalisme de Pritchard et le doute radical du sceptique ont été révoqués ici de la même façon, c'est-à-dire par une approche conceptuelle des formules $\Box(K\phi \rightarrow \phi)$ et une analyse pragmatique des effets de l'usage de certains concepts.

L'interprétation proposée ici a un prix : elle oriente l'étude des concepts épistémiques dans un sens clairement pragmatique des actes de discours, en témoigne le recours à la notion linguistique des verbes parenthétiques ; elle centre l'étude de ces concepts sur la relation instaurée au sein d'un dialogue entre un locuteur et son interlocuteur, puisque c'est l'utilisation de la locution 'savoir que ϕ' par un locuteur particulier qui garantit pour celui-ci le fait de *tenir* ϕ *pour* vrai. La différence entre l'analyse cognitive et l'analyse linguistique repose sur le point de vue duquel on examine l'opérateur K, selon qu'il caractérise un processus de justification ou un jeu de langage. Ce n'est pas l'analyse cognitive qui a été

[260] [Hintikka 1968], p. 13.

invoquée plus haut pour la défense de la clause de succès mais une analyse en termes d'actes de discours, laquelle indique une déclaration de connaissance de ϕ de la part d'un locuteur mais sans prétendre à la vérité absolue (dans le monde réel) de ϕ.[261]

Si l'on reprend le critère précédent de l'indéfendabilité épistémique, en vertu duquel le contenu d'un énoncé peut être vrai mais impossible à prononcer pour un agent, la nécessité conceptuelle symbolisée par $\Box$ dans $\Box(K\phi \rightarrow \phi)$ n'est rien d'autre qu'une *nécessité épistémique* de forme $K(K\phi \rightarrow \phi)$: ce que je sais à coup sûr, c'est qu'une connaissance est quelque chose de vrai. Quel rapport avec l'indéfendabilité épistémique, qui portait non pas sur l'axiome **T** mais sur **4** ? Il y a un lien entre la clause de vérité et l'introspection positive, comme le montrera plus bas l'interprétation pragmatique de $(K\phi \rightarrow \phi)$ par [Zemach 1969]. Pour montrer ce lien, notons qu'une des objections que [Hocutt 1972] avait soulevées contre la thèse KK plaidait tout au contraire en faveur de cette dernière : dans la note 159 de la section 2.2.2, Hocutt prétendait en effet que la stratégie de réduction sémantique de l'opérateur K via (C.K) permet d'inférer un résultat absurde de l'introspection positive : si le savoir implique le savoir que l'on sait, alors tout ce qui est vrai est connu parce que (C.K) permet de passer du théorème $K(\phi \rightarrow K\phi)$ à l'autre théorème $(\phi \rightarrow K\phi)$ après élimination de l'opérateur de portée large. Or ce résultat n'est plus absurde dès lors que l'on interprète $\phi \rightarrow K\phi$ dans une perspective pragmatique des actes de discours : si l'on part de l'idée que le fait de déclarer ϕ engage le locuteur sur la vérité de *sa* déclaration, alors ce n'est pas l'énoncé $\phi \rightarrow K\phi$ mais l'énonciation de $\phi \rightarrow K\phi$ qui est considérée comme valable, c'est-à-dire compatible avec les connaissances de l'agent locuteur. Pour distinguer l'énoncé de ϕ et son énonciation, on utilisera la formule entre crochets $\ulcorner\phi\urcorner$ pour symboliser l'acte d'énonciation de ϕ. Autrement dit : 'je dis que ϕ est vraie, donc je le sais' ne revêt pas la même signification que 'je dis que ϕ est vraie, donc tout le monde le sait' et, pour reprendre les termes de [Hintikka 1962], seule la première formule est épistémiquement auto-confirmante.[262]. On aperçoit mieux le lien logique entre $\ulcorner\phi \rightarrow K\phi\urcorner$ et le critère d'indéfendabilité épistémique (section 2.2.4.2.2) : s'il est objectivement possible de ne pas savoir quelque chose de vrai, $\Diamond(\phi \wedge \neg K\phi)$, en revanche il n'est pas subjectivement possible d'*asserter* que quelque chose est vrai sans qu'on le sache : $\neg\Diamond K_a(\phi \wedge \neg K_a\phi)$, quel que soit le locuteur *a*. La clause de vérité exprime ainsi en termes d'énoncés l'idée que toute énonciation présuppose une connaissance préalable : $\ulcorner\phi \rightarrow K\phi\urcorner$, où l'antécédent de la version locutoire devient le conséquent de la version illocutoire et vice versa. C'est en vertu de cette interprétation pragmatique de la clause de vérité que l'antécédent $K\phi$ et le conséquent ϕ peuvent être inversés et montrer l'affinité existant entre l'interprétation des axiomes **T** et **4**. Insistons aussi sur le fait que la connexion établie ici entre $(K\phi \rightarrow \phi)$ et $(\ulcorner\phi \rightarrow K\phi\urcorner)$ n'entraîne pas l'équivalence de $K\phi$ et ϕ et la trivialité de l'opérateur de connaissance, en vertu du raisonnement suivant : $(K\phi \rightarrow \phi) \wedge (\phi \rightarrow K\phi)$, donc $K\phi \equiv \phi$'. Nous sommes ici en présence de deux types d'expression distincts : les formules $(K\phi \rightarrow \phi)$ et $\ulcorner p \rightarrow Kp\urcorner$ ne représentent pas un

[261] On pourrait comparer en un sens la vérité de ϕ dans $(K\phi \rightarrow \phi)$ à l'expression d'un *engagement ontologique* de la part du locuteur : ϕ ne dit rien *sur* la réalité authentique du monde, il parle simplement *de* ce que le locuteur considère comme étant la réalité du monde.

[262] La formule assertorique ϕ est épistémiquement auto-confirmante si sa version modale $K\phi$ est auto-confirmante ; ainsi, $\ulcorner\phi \rightarrow K\phi\urcorner$ est épistémiquement auto-confirmante parce que $K(\phi \rightarrow K\phi)$ est auto-confirmante.

seul et même type d'expression (un énoncé pour la première, une énonciation pour la seconde).

Or si c'est une *déclaration* de vérité qui garantit le droit d'attribuer une déclaration de connaissance, cela veut dire que $\ulcorner\phi\urcorner$ implique (pragmatiquement) $\ulcorner K\phi\urcorner$ et que les deux autres conditions classiques du savoir, la croyance et la justification, ne jouent pas un rôle essentiel pour garantir la connaissance. [Zemach 1969] considère cette situation comme un paradoxe pragmatique pour la notion de connaissance, dans sa relation à la vérité : la vérité de $K\phi$ implique la vérité de ϕ, mais c'est le fait de dire 'ϕ' (symbolisé : $\ulcorner\phi\urcorner$) qui implique le fait de dire '$K\phi$' (symbolisé : $\ulcorner K\phi\urcorner$). L'analyse pragmatique de la clause de vérité, si elle permet d'esquiver le doute radical des sceptiques, ne crée-t-elle pas une forme de circularité en échange ? Bien que la relation établie ici ne soit pas une circularité logique, dans la mesure où nous n'avons pas ($K\phi \rightarrow \phi$) et ($\phi \rightarrow K\phi$) mais ($K\phi \rightarrow \phi$) et $\ulcorner\phi \rightarrow K\phi\urcorner$, voyons de plus près ce que Zemach entend par ce paradoxe ainsi que les conclusions qu'il faut en tirer sur les relations conceptuelles entre connaissance et vérité.

[Zemach 1969] compare la définition classique du savoir au cas d'un médecin qui consulterait son dictionnaire médical pour découvrir les conditions nécessaires et suffisantes du cancer : s'il est écrit qu'un malade est atteint du cancer si et seulement si : 1. il a une tumeur de type A, 2. il se comporte de la manière B, et 3. il a le cancer, cela peut signifier deux choses distinctes : soit la condition 3 est redondante et les conditions 1 et 2 suffisent pour diagnostiquer un cancer ; soit la condition 3 est nécessaire et le cancer est indéfinissable, de sorte que les conditions 1 et 2 sont inutiles et ne suffisent jamais à elles seules pour inférer la présence d'un cancer. La situation est parallèle avec la définition classique du savoir : une personne sait que ϕ si et seulement si : A. elle croit que ϕ, B. elle possède une justification adéquate en faveur de ϕ, et C. ϕ est vraie. Par analogie avec le cas du cancer, Zemach dit que soit la condition C est redondante et la croyance justifiée suffit pour spécifier un cas de connaissance, soit les conditions A et B sont insuffisantes pour garantir un cas de connaissance et seul C y parvient. Mais s'il s'agit d'une *circularité logique*, c'est-à-dire d'un truisme dans le cas du cancer, puisque la condition 3 (avoir le cancer) répète ce qui est à définir (le cancer), il s'agit d'une *circularité pragmatique* dans le cas de la connaissance : la condition C (être vrai) ne répète pas dans les termes ce qui est à définir (la connaissance), mais la notion de vérité présuppose toutefois la connaissance dans la mesure où une *déclaration* de vérité de (i.e. $\ulcorner\phi\urcorner$) présuppose une connaissance acquise. Que faut-il entendre par l'idée de présupposition pragmatique ?

De même que dans les paragraphes précédents, [Zemach 1969] suggère par la notion de 'pragmatique' le cas d'attribution de connaissance à soi-même exigée par le sceptique. Il est possible qu'une personne sache quelque chose sans savoir elle-même qu'elle le sait, et le sceptique ne le conteste pas à son interlocuteur. Ce que le sceptique conteste, en revanche, c'est que le sujet soit lui-même en mesure de justifier la vérité de sa propre affirmation. Dans ce cas, il est impossible pour le sujet d'affirmer ϕ parce que cette déclaration de vérité suggérerait (impliquerait pragmatiquement) qu'il possède effectivement ce qu'en réalité il ne

possède peut-être pas : une connaissance.[263] Le lien conceptuel est confirmé de cette façon entre **T** et **4**, c'est-à-dire entre la clause de vérité et le paradoxe de Moore :

> je ne prétends pas que la définition admise du savoir est logiquement circulaire. Mais je prétends qu'elle est *pragmatiquement* circulaire, *i.e.* qu'elle devient nécessairement circulaire à l'opposé du but recherché [*self-defeatingly*], dans son application à elle-même. Autrement dit, sa relation au cas de circularité ci-dessus est pareille à la relation de 'p, mais je ne crois pas que p' ou 'je ne sais pas faire d'affirmation en [français]' à 'p et non p'.[264]

Le paradoxe pragmatique porte sur les *actes* de langage qu'un sujet doit effectuer afin de garantir une déclaration de connaissance : pour savoir effectivement que ϕ, il faut que ϕ soit cru, que l'on ait une preuve adéquate de ϕ et que ϕ soit vrai ; or il est impossible de découvrir que ϕ est vrai sans savoir par là-même que ϕ est vrai, dans la mesure où une déclaration de vérité d'un énoncé va de pair avec sa connaissance :

> Mais dès que je découvre que la première condition est satisfaite, i.e. que p est le cas, je sais que p : il est impossible pour moi d'établir le fait que p sans parvenir à savoir que p.[265]

Le dernier critère de la connaissance est indissociable de la connaissance elle-même : être en mesure de dire si ϕ est vrai nécessite que l'on sache si ϕ est vrai, donc il faut déjà savoir que ϕ est vrai avant de pouvoir déclarer que ϕ est vrai. Bien que ce que l'on affirme puisse être vrai ou faux indépendamment de notre déclaration, nous ne pouvons parler *en pratique* de sa vérité sans supposer qu'on le sait, et c'est en cela que le paradoxe de la connaissance est d'ordre pragmatique.

L'ambiguïté centrale du paradoxe repose sur les notions d'*implication* et de *présupposition*. Nous avons précisé plus tôt qu'il n'y a pas une relation de circularité entre ϕ et $K\phi$, car si $K\phi$ implique ϕ en revanche ϕ n'implique pas $K\phi$: $(K\phi \rightarrow \phi)$ est valide, mais pas $(\phi \rightarrow K\phi)$. On pourrait exprimer cette différence en disant que la circularité pragmatique met aux prises deux types d'implication distincts : une implication logique et une implicature, ou présupposition. Si l'on exprime le fait que ϕ présuppose ψ sous la forme $(\phi \Rightarrow \psi)$, le paradoxe pragmatique signifie que la vérité de ϕ présuppose la connaissance de ϕ par le locuteur : $\phi \Rightarrow K\phi$, tandis que la connaissance de ϕ implique la vérité de ϕ : $(K\phi \rightarrow \phi)$. La circularité pragmatique entre ϕ et $K\phi$ se manifeste ainsi dans l'usage de deux 'implications' distinctes :

---

[263] "A première vue, il semble que le sceptique puisse défendre sa position en disant –'oui, si "p" est effectivement vraie, i.e. si p est le cas, alors nous savons bel et bien que p. Mais nous ne pouvons pas *dire* si 'p' est effectivement vraie. Donc bien qu'*à notre insu* il puisse être vrai que nous sachions que p, nous ne pouvons pas savoir que nous savons que p". ([Zemach 1969], p. 285 ; nous soulignons). L'auteur en conclut que le problème du sceptique ne porte pas tant sur la vérité possible d'un énoncé que sur les moyens de la reconnaître : c'est un problème de logique auto-épistémique, pour ainsi dire, porté sur les conditions d'*attribution de* connaissance à soi-même.

[264] [Zemach 1969], p. 283.

[265] [Zemach 1969], p. 284. On retrouve cette connexion entre ϕ et $K\phi$ lorsque l'on combine la théorie déflationniste de la vérité et la théorie de la croyance linguistique de Ramsey (Chapitre Trois, section 39) : 'ϕ est vrai' = $\ulcorner\phi\urcorner$ et $\ulcorner\phi\urcorner$ = 'j'asserte que ϕ', donc 'ϕ est vrai' = 'j'asserte que ϕ'. Si l'on assimile l'assertion à une déclaration de savoir, on obtient ainsi l'équivalence $\ulcorner\phi\urcorner = \ulcorner K\phi\urcorner$. Rappelons que $\ulcorner\phi\urcorner = \ulcorner K\phi\urcorner$ ne donne pas $(\phi \rightarrow K\phi)$ mais $\ulcorner\phi \rightarrow K\phi\urcorner$ ou $\ulcorner\phi\urcorner \rightarrow \ulcorner K\phi\urcorner$, qui expriment des relations logiques entre des énonciations et non entre des énoncés.

($K\phi \rightarrow \phi$) et ($\phi \Rightarrow K\phi$), où la présupposition signifie $\ulcorner\phi \rightarrow K\phi\urcorner$ et peut être interprétée en termes d'implication épistémique sous la forme $K(\phi \rightarrow K\phi)$. Plutôt que de multiplier inutilement des relations conditionnelles distinctes, la notion de présupposition peut être rendue par l'idée de connaissance d'un conditionnel : si ϕ présuppose $K\phi$, cela signifie que la connaissance de ϕ implique la connaissance de $K\phi$ ou que ϕ ne peut être connue sans que $K\phi$ le soit également : $(\phi \Rightarrow K\phi) =_{df} K(\phi \rightarrow K\phi)$, donc $(\phi \Rightarrow K\phi) =_{df} (K\phi \rightarrow KK\phi)$.

La notion de présupposition est apparue auparavant dans la théorie de Strawson sur les énoncés singuliers, en vertu de laquelle la vérité ou fausseté d'un énoncé tel que '*a* est F' présuppose l'existence de *a* : si l'on admet cette relation de présupposition entre signifiance et existence, on constate comme dans le cas ci-dessus que la présupposition existentielle peut s'exprimer également sous la forme conditionnelle $[Fa \rightarrow (\exists x)Fx]$: si la vérité de Fa présuppose l'existence de *a*, alors on peut dire que Fa ne peut être vraie sans que $(\exists x)Fx$ le soit également. Présupposition et implication se retrouvent également dans la théorie de l'engagement ontologique de Quine, dans lequel un énoncé vrai ne dit pas ce qui *est* le cas mais ce que locuteur *considère* comme étant le cas, c'est-à-dire ce dont le locuteur présuppose l'existence dans son schème conceptuel. Bien que la logique épistémique de Hintikka ne comporte pas de symbole distinct pour l'implication pragmatique (relation de présupposition) et l'implication matérielle (relation conditionnelle), on peut considérer que les trois formules qui suivent expriment la même chose : $(\phi \Rightarrow K\phi)$, $\ulcorner\phi \rightarrow K\phi\urcorner$ et $K(\phi \rightarrow K\phi)$ signifient que l'énonciation $\ulcorner\phi\urcorner$ suggère, présuppose ou implique (pragmatiquement) la connaissance avérée de l'énoncé ϕ. L'avantage de ce symbolisme est qu'il rend compte de l'ordre de démonstration dans lequel les concepts de connaissance et de vérité apparaissent : il n'y a pas de confusion entre antécédent et conséquent, ici, mais une distinction entre une relation logique d'antécédent à conséquent ($K\phi \rightarrow \phi$) et une relation pragmatique de condition à conséquence ($\phi \Rightarrow K\phi$).

Etant donné que la vérité de ϕ est une condition préalable à la connaissance de ϕ, et puisque la connaissance de ϕ est une condition préalable à la *déclaration de* la vérité de ϕ, personne ne peut *déclarer* ϕ sans s'engager sur $K\phi$ dans le même temps. [Zemach 1969] voit là un paradoxe pragmatique qui revient à assimiler dans l'usage les concepts de vérité et de connaissance :

> il y a toujours un fossé logique entre "'p' est hautement corroboré" et "p est vrai". Mais nous savons que p seulement si nous savons que 'p' est vrai.[266]

Pour revenir sur la définition classique du savoir, celle-ci dit que croyance, justification adéquate et vérité sont les trois conditions nécessaires et suffisantes pour la connaissance : $K\phi \equiv (B\phi \wedge J\phi \wedge \phi)$. Chacun de ces membres conjoints est considéré comme une condition nécessaire : $(K\phi \rightarrow B\phi)$, $(K\phi \rightarrow J\phi)$ et $(K\phi \rightarrow \phi)$, mais aucun n'est supposé suffire à lui seul pour garantir une connaissance. Quant au paradoxe pragmatique de la connaissance, il a montré au contraire que 'ϕ' est suffisante pour conclure $K\phi$, mais à condition de rappeler que 'ϕ' représente ici une déclaration de vérité $\ulcorner\phi\urcorner$ et que la relation entre ϕ et $K\phi$ est une implication épistémique, ou pragmatique.

---

[266] [Zemach 1969], p. 285.

En conclusion, l'approche pragmatique des notions de connaissance et de vérité apporte à la fois un avantage et un inconvénient : elle permet de surmonter le doute radical du sceptique en réduisant le fait d'être vrai au fait de tenir pour vrai ; elle crée une certaine circularité entre connaissance et vérité, étant donnée la relation de présupposition qui a lieu entre tenir pour vrai et connaître. Deux étapes ont marqué ce résultat : pour défendre la plausibilité de ($K\phi \rightarrow \phi$) contre le sceptique, nous avons précisé que la vérité de ϕ n'est pas absolue mais exprime une sorte de vérité conceptuelle, c'est-à-dire une vérité consécutive à l'usage du concept de connaissance de ϕ. Donc si le fait d'être vrai équivaut plus ou moins au fait d'être en droit de déclarer la vérité, le conséquent ϕ de ($K\phi \rightarrow \phi$) n'exprime pas un état de choses objectif mais un acte de discours, l'énonciation ou assertion $\ulcorner\phi\urcorner$. Par analogie avec l'implication épistémique du paradoxe de Moore, le sceptique n'est pas réfuté par notre argumentation puisque l'affaiblissement des conditions de vérité de ϕ produit dans le même temps une déclaration de connaissance paradoxale : dire ϕ, c'est re*connaître* la vérité de ϕ et présupposer ce qui est encore à démontrer pour lui, à savoir $K\phi$. Le dilemme est donc le suivant : soit 'ϕ est vraie' exprime un fait objectif au sujet de la proposition ϕ (son 'être vrai'), auquel cas le sceptique affirme de droit qu'aucune forme de justification ne permet de justifier cette déclaration ; soit 'ϕ est vraie' exprime simplement un fait subjectif au sujet de la connaissance de ϕ (son 'être tenu pour vrai'), auquel cas ϕ ne symbolise plus un énoncé mais un acte d'énonciation qui présuppose ce qui est à démontrer. La forme logique ($K\phi \rightarrow \phi$) est donc plus ambiguë qu'il n'y paraît, dès lors qu'il s'agit de rechercher les conditions d'énonciation d'une vérité. Il va de soi qu'un tel problème ne se pose pas si, comme c'est généralement le cas en logique, on ne s'intéresse pas aux conditions de vérité matérielle et stipule simplement la vérité ou fausseté de ϕ. La logique n'est pas plus de l'ontologie que de l'épistémologie, encore une fois, et ce n'est pas son rôle de déterminer si un énoncé simple est vrai ou faux : le problème portait ici sur la relation conceptuelle entre les concepts de vérité et de connaissance (que signifie 'être vrai' et 'connaître' ?), et non sur la relation entre les contenus de vérité et de connaissance (qu'est-ce qui est vrai, qu'est-ce que je sais ?). La logique modale épistémique a introduit un autre angle d'analyse dont le logicien classique ne se soucie pas : la dimension pragmatique des actes de discours, et notamment la relation d'implication entre les déclarations de vérité et de connaissance. C'est cette dimension qui a créé la difficulté et mis en doute la formulation de la clause de vérité, uniquement. Si elle évite les complications liées à l'interprétation cognitive (externaliste ou internaliste) de l'opérateur de connaissance, une telle analyse montre en tout cas que le symbolisme logique peut nous induire en erreur (que veut dire le signe ϕ, dans la formule ($K\phi \rightarrow \phi$) ?) ou supposer des relations logiques qui ne sont pas aussi évidentes qu'il y paraît tout d'abord.

Au Chapitre 3, nous ferons allusion à une autre sorte de critique de l'axiome **T** : la critique anti-réaliste des relations entre vérité et connaissance, dans laquelle la vérité n'est plus indépendante mais synonyme de croyance justifiée. Puis nous comparerons l'interprétation pragmatique ci-dessus de la connaissance avec son interprétation anti-réaliste et en conclurons ce qui, après analyse, paraîtra le mieux disposé à servir de signification de base pour l'opérateur K. Quoi qu'il en soit, les remarques de cette section ont eu pour but de montrer que l'analyse classique de la connaissance comme croyance vraie justifiée produit non pas

une contradiction explicite (une circularité logique), mais une contradiction implicite (une circularité pragmatique) entre les quatre concepts épistémologiques de connaissance, de croyance, de justification et de vérité. Compte tenu de notre interprétation des axiomes **T** et **4**, une signification de base se dégage progressivement pour K et qui nous conduit vers une *logique de l'auto-épistémologie*, c'est-à-dire une logique des attributions de connaissance centrées sur la première personne.

Avant de revenir plus en détail sur cet aspect des attitudes épistémiques, un autre problème plus profond apparaît dans la théorie de la quantification : à l'image du phénomène d'opacité référentielle, les formules épistémiques posent des difficultés lorsqu'il s'agit d'identifier leurs objets et compromettent leurs conditions de signifiance. Si l'identité d'un objet d'attitude varie selon le sujet croyant, comment garantir la signification de ces inférences et la conséquence logique entre des attitudes irrégulières ? Nous allons voir non seulement que des conditions minimales d'identification existent pour ces objets, mais que ces conditions justifient à nouveau une interprétation illocutoire des attitudes épistémiques.

2.3 Logique et contexte

Une façon de rappeler au bon souvenir des aspects locutoires de la logique modale épistémique est sa théorie de la quantification, dans laquelle sont étudiées les relations sémantiques (et non plus pragmatiques) entre les attitudes épistémiques et les objets (c'est-à-dire les constantes d'individu) sur lesquels elles portent. L'analyse des modalités sera étendue ici aux formules du premier ordre. Mais la même question se posera que celle qui a déjà été posée en rapport à la clause de vérité : sous quelles conditions peut-on dire d'un énoncé de connaissance ou de croyance qu'il est vrai ? L'analyse quantifiée de ces énoncés nous déplace vers le problème contextuel de l'*individuation* des objets d'attitude, individuation dont dépend la signification des énoncés.

Pour aborder les méthodes d'analyse de la logique modale épistémique sous leur version quantifiée, nous reviendrons sur le cas de l'*opacité référentielle*, principal obstacle aux règles de la théorie de la quantification. Nous reviendrons également sur les objections posées à cet égard par l'adversaire majeur de la logique modale, W. V. Quine, puis nous insisterons sur les réponses proposées par Hintikka avant d'arriver sur un point plus fondamental de l'enquête : l'individuation des objets d'attitude est-elle du ressort de la logique, où dépend-elle de critères invariablement contextuels ? Après avoir montré les impasses auxquelles peut conduire la question de l'individuation, nous reviendrons finalement à une approche *illocutoire* et proposerons de traiter l'opacité référentielle en termes plus modestes (mais plus sûrs) d'énonciations. Nous en tirerons quelques enseignements généraux sur la portée des méthodes de logique en philosophie et aboutirons ainsi au troisième chapitre, consacré plus en détail à l'élaboration d'une autre logique des attitudes.

2.3.1 Le problème de l'opacité référentielle

Une version quantifiée de l'extensionalisme est la loi de substitutivité des identiques : deux variables d'individus peuvent être librement substituées *salva veritate* si elles ont la même référence, c'est-à-dire si elles prennent pour valeur le même individu. Cette règle se décline

en deux sens : une loi d'identité des indiscernables, en vertu de laquelle deux individus sont identiques s'ils partagent exactement les mêmes propriétés :

(I1) $(\forall x)(\forall y)[(\forall F)(Fx \equiv Fy) \rightarrow (x = y)]$

puis une loi converse d'indiscernabilité des identiques, en vertu de laquelle deux individus partagent exactement les mêmes propriétés s'ils sont identiques,

(I2) $(\forall x)(\forall y)[(\forall F)(x = y) \rightarrow (Fx \equiv Fy)]$

A strictement parler, la loi de Leibniz correspond à la version (I2), mais les deux sont souvent considérées comme interchangeables et nous entendrons par loi de Leibniz ce qui suit :

(LL) $(\forall x)(\forall y)[(\forall F)(Fx \equiv Fy) \equiv (x = y)]$

Or nous avons déjà signalé que [Frege 1892] avait perçu la difficulté d'application de cette règle d'extensionalité dans certains contextes de discours, parmi lesquels les contextes de croyance : un individu peut porter des croyances contradictoires sur un seul et même individu, à l'instar des croyances de Tom sur l'auteur de *De Senectute* (dans [Quine 1953]). Soit p l'énoncé 'Cicéron a écrit De Senectute' ; bien que les termes d'individus Cicéron et Tullius renvoient à la même personnalité romaine, Tom croit de l'individu Cicéron qu'il a écrit cet ouvrage mais ne le croit pas de Tullius ; or cela signifie que Tom croit et ne croit pas la même chose du même individu : $B_{Tom}p \wedge \neg B_{Tom}p$. Il peut soutenir également des croyances contraires s'il croit que Tullius n'a pas écrit *De Senectute*, auquel cas on obtient la version encore moins défendable $B_{Tom}p \wedge B_{Tom}\neg p$.[267]

Deux réactions sont envisageables face à ce cas de figure. Une attitude 'répressive', tout d'abord : les conséquences de l'opacité référentielle ci-dessus sont si graves que le coupable doit être puni, d'où le rejet de la loi de Leibniz. C'est la voie qui sera empruntée par Hintikka et que nous décrirons ci-dessous : $[(LL) \rightarrow \bot] \rightarrow \neg(LL)$. Une attitude 'préventive' : la loi de Leibniz est maintenue, mais à condition de faire en sorte que son application n'aboutisse pas aux conséquences graves décrites plus haut : $(LL) \wedge \neg\bot$. Nous chercherons à défendre cette attitude, après avoir examiné la première et observé l'impasse vers laquelle semble se diriger la logique philosophique.

Dans [Frege 1892], nous rencontrons un exemple d'attitude préventive à l'égard de (LL) : pour conserver cette loi, l'auteur proposa de distinguer un sens habituel et un sens inhabituel de la référence, c'est-à-dire un contexte direct et un contexte indirect (ou 'oblique'). Dans un contexte de croyance, qui représente un contexte de discours indirect, le sens habituel d'un énoncé (la pensée que l'énoncé exprime) devient sa référence et la substitution devient dès lors syntaxiquement incorrecte : la substitution évoquée dans (LL) s'effectue dans un contexte transparent entre des références directes, et non dans un contexte opaque entre des références indirectes (le sens des références directes). Or nous avons signalé plus tôt que l'on ne peut pas effectuer une substitution *salva veritate* entre des pensées ou références inhabituelles : l'explication de Frege était que, dans un contexte de croyance, les co-références deviennent des sens distincts et ne peuvent plus être substitués l'un à l'autre en raison de la différence de leur sens.

---

[267] En vertu de l'inférence logique $(B\neg\phi \Rightarrow \neg B\phi)$, une croyance contraire est *a fortiori* une croyance contradictoire. Nous verrons toutefois que [Zemach 1971] conteste l'inférence ci-dessus, dans le cas des attributions de croyance. Nous ne la contestons pas, pour notre part.

Pour [Hintikka 1957], en revanche, la loi (LL) doit être déchue parce qu'elle est réellement mise en défaut : il y a bel et bien un échec de la substitution. Le problème de l'opacité référentielle ne provient pas selon lui d'une absence totale de référence mais, au contraire, d'une abondance ou multiplicité de références.[268] La substitutivité est toujours permise entre les termes habituellement co-extensionnels puisque, contrairement à ce que dit Frege, ces termes conservent leur fonction référentielle dans un contexte inhabituel : les termes d'individus désignent toujours des individus, les énoncés désignent toujours des valeurs de vérité. Comment expliquer alors l'opacité référentielle ? Selon Hintikka, la différence avec le contexte de discours habituel (non-modal) peut s'expliquer dans les termes de la sémantique relationnelle : elle vient du fait que, dans un contexte de croyance, la fonction référentielle des termes est maintenue mais leurs références peuvent varier d'un monde possible à un autre. En logique classique, par contre, le domaine de référence est constant et assigne une référence unique pour chaque terme. Le comportement des termes référentiels dans un contexte de croyance a une répercussion sur la valeur de vérité des énoncés qui les contiennent. Soient par exemple p représentant l'énoncé 'Cicéron a écrit *De Senectute*' et q l'énoncé 'Tullius a écrit *De Senectute*'. Puisque les termes co-référentiels 'Cicéron' et 'Tullius' désignent le même individu dans le monde réel, toute prédication qui est vraie ou fausse de l'un est vraie ou fausse de l'autre dans ce monde réel : $v(p)=1 \Rightarrow v(q)=1$. En logique doxastique, le domaine de référence devient variable et assigne des références multiples pour chaque terme en plus de leurs références stables du monde réel. 'Tullius' ne désigne pas l'unique individu Cicéron pour Tom, soit parce qu'il hésite sur son identité soit parce qu'il croit que Tullius désigne quelqu'un d'autre. Dans les deux cas de figure, les expressions modales 'Tom croit que Cicéron a écrit *De Senectute*' et 'Tom croit que Tullius a écrit *De Senectute*' n'auront pas la même valeur de vérité : $v(B_{Tom}p) \neq v(B_{Tom}q)$. Soit Tom ignore l'identité exacte de Tullius et l'associe à plusieurs personnes possibles, auquel cas il ne croit rien de précis à son sujet parce que le nom de Tullius est compatible avec plusieurs individus dans l'ensemble des croyances de Tom : $v[w,\neg B_{Tom}(\text{Tullius} = \text{Cicéron})]=1$, donc $v[w^*,(\text{Tullius} \neq \text{Cicéron})]=1$ pour au moins un monde possible w* compatible avec les croyances de Tom. Soit Tom identifie le nom de Tullius à une autre personne que Cicéron, auquel cas il croit quelque chose de précis à son sujet parce que le nom de Tullius renvoie au même individu dans l'ensemble des croyances de Tom : $v[w,B_{Tom}(\text{Tullius} \neq \text{Cicéron}]=1$, donc $v[w^*,(\text{Tullius} \neq \text{Cicéron})]=1$ dans tous les mondes possibles w* compatibles avec les croyances de Tom. Les trois types de croyances décrits ici montrent que la substitution des termes co-référentiels échoue et condamne (LL) : selon l'instruction du locuteur, il est possible pour lui de croire que Cicéron et Tullius désignent le même individu, de ne pas croire qu'ils désignent le même individu ou, pire, de croire qu'ils désignent un individu différent.

Le problème de la signification des énoncés modaux est uniquement dû à un problème de détermination de la référence, donc, et non à un changement de fonction des termes selon le contexte de discours : la question de la signification ne doit pas être dissociée de la question

[268] "L'opacité référentielle n'est pas due ici à quelque chose d'étrange se produisant dans les façons dont nos termes singuliers réfèrent aux objets, ni à quelque chose d'inhabituel concernant les objets auxquels ils cherchent à référer. C'est simplement et uniquement dû au fait que nous devons considérer plus d'une façon dont ils peuvent référer (ou ne pas référer) aux objets. Ce que nous devons traiter n'est donc pas tant une absence de référentialité qu'un type de référentialité multiple" ([Hintikka 1962], pp. 139-140).

de la référence.[269] L'interprétation dans un modèle suffit pour expliquer les conditions de succès d'une nomination. Soit *f* la fonction de valuation assignant des valeurs d'individus dans un domaine des éléments d'un ensemble de mondes possibles w de W, où W désigne l'ensemble des mondes compatibles avec les croyances d'un sujet croyant. Puisque la théorie de la signification et la théorie de la référence ne font qu'une pour Hintikka, donner une signification à un énoncé c'est assigner des référents à ses termes, mais selon un processus plus complexe que dans la logique classique à modèle unique. La croyance à un énoncé p:F*a* est interprétée formellement comme la fonction *f* assignant terme d'individu *a* la même valeur dans tous les mondes possibles pour le sujet. Le lien entre signification et référence dépend donc de ce que les logiciens modaux appellent la *fonction individuante*, c'est-à-dire la fonction dont le rôle est d'interpréter ou assigner une même valeur à chaque terme dans chaque monde ; l'échec de la référence se manifestera tant que cette fonction n'est pas appliquée, c'est-à-dire tant que l'assignation des valeurs n'est pas constante d'un monde possible à un autre. Lorsqu'il y a échec de la substitution de deux termes co-référentiels *a* et *b*, explique [Hintikka 1957], cela veut dire que (a ≠ b) est vrai dans au moins une des mondes compatibles avec les croyances du sujet et que la fonction individuante ne fournit pas la même interprétation des termes à travers tous les mondes possibles pour l'agent ; on dira de l'agent qu'il *ne* croit *pas* que *a* est (le même individu que) *b*. S'il croit que *a* n'est pas *b*, en revanche, (a ≠ b) est vrai non seulement dans un mais tous les mondes doxastiquement possibles pour l'agent. Dans le cas précédent de Tom, si celui-ci ne croyait pas que Tullius a dénoncé Catilina c'est parce que Cicéron et Tullius n'étaient pas identiques dans tous les mondes et désignaient deux individus différents dans au moins un des scénarios w* compatibles avec l'ensemble de ses croyances.

Remarquons que nous ne parlons pour l'instant que du 'monde intérieur' de Tom, c'est-à-dire l'ensemble des croyances qu'il tient indépendamment de l'état du monde actuel. Le problème porte jusqu'ici sur la formulation logique adaptée pour décrire les modalités de la référence ; celle-ci revenait à dire que deux termes sont *référentiellement transparents* s'ils renvoient au *même* individu dans *tous* les mondes doxastiquement possibles. La transparence est évidemment assurée en logique classique, puisque le domaine est unique et ne peut pas varier ; elle peut assurée également en logique épistémique, lorsque les attitudes de l'agent sont déterminées et n'expriment pas une incertitude à l'égard de l'identité des individus dont il parle(c'est-à-dire lorsque la modalité qui exprime son attitude est de type □). Mais le problème de l'identité se complique en théorie de la quantification, lorsque la signification d'une formule dépend de l'identité des références dans le monde intérieur d'un sujet *et* dans le monde extérieur de tous les sujets (le monde 'actuel') : c'est le problème de la *transquantification* (ou 'quantification croisée') qui désigne le fait de quantifier au-delà des ensembles de mondes possibles d'un agent particulier. Pour évoquer ce problème, on applique la loi de Généralisation Existentielle dans un contexte modal :

(GE□) $\Box Fa \rightarrow (\exists x)(\Box Fx)$

(où □ désigne un contexte de discours tel que la croyance).

[269] L'intitulé du premier paragraphe de [Hintikka 1969b] le dit clairement : 'C'est se méprendre que d'opposer l'une à l'autre la théorie de la référence et la théorie de la signification' (p. 37). Une fois cela dit, il reste à formuler une théorie de la référence capable de rendre compte des cas d'opacité référentielle.

Deux thèmes interdépendants sont examinés en logique modale épistémique : la question de l'*identité* des individus, d'une part ; la question de l'*existence* des individus, d'autre part. Les deux problèmes sont liés puisque, lorsque l'on quantifie sur un individu, celui-ci reçoit des propriétés particulières qui permettent de l'identifier ;[270] on obtient ainsi la symbolisation $(\exists x)B_a(x = c)$ pour dire qu'il y a un individu quelconque x dont *a* croit qu'il la propriété d'être *c* (où *c* symbolise un nom propre ou une description définie). Comment préserver la loi (GE) si, pour reprendre un exemple de [Hintikka 1962], un agent tient une croyance sur un objet fictif ?[271] Faut-il admettre l'existence de cet objet dans le monde intérieur du sujet, où cette conséquence absurde suffit-elle à invalider (GE) ? C'est cette seconde voie ('répressive') que Hintikka empruntera ; la substitution *salva veritate* dans un contexte modal est invalide à moins d'ajouter à (LL) une hypothèse auxiliaire.

Or même si l'on restreint l'application de (LL) ou (GE) aux contextes extensionnels, d'autres difficultés apparaissent et qui sont propres à l'exercice de quantification dans un contexte modal. Un exemple est la Formule de Buridan (BF), dans laquelle sont examinées les relations conditionnelles entre quantification *de dicto* et quantification *de re*.[272] La loi d'exportation, tout d'abord (implication *de dicto-de re*) :

(BF_e) $\Box(\exists x)Fx \rightarrow (\exists x)\Box Fx$,

où la relation de conséquence va du monde de l'agent vers le monde actuel ; puis la loi d'importation (implication *de re-de dicto*)

(BF_i) $(\exists x)\Box Fx \rightarrow \Box(\exists x)Fx$,

où la relation de conséquence va du monde actuel vers le monde de l'agent.

Bien qu'elles paraissent intuitivement acceptables sous certaines interprétations de $\Box$, de telles formules semblent inacceptables dans un contexte doxastique. Voici un exemple pour la version de l'exportation, où l'inférence *de dicto-de re* est mise en échec. Paul croit *que* son père est un philanthrope et que son oncle inconnu, le frère jumeau de son père, est un affreux vendeur d'armes de destruction massive. Paul croit donc que son oncle n'est pas aussi bon

---

[270] L'analyse classique de la quantification existentielle dit en effet que 'x existe' signifie 'x est quelque chose' : $(\exists x)Fx$, pour une propriété F quelconque. Une question à venir consistera à demander si une analyse prédicative de ce genre est suffisante pour les énoncés de connaissance : toute expression d'un savoir est-elle un 'savoir que quelque chose' ?

[271] "Nous ne pouvons pas accepter une logique en vertu de laquelle '*a* croit qu'il est poursuivi par l'abominable homme des neiges' implique qu'il y a une chose telle que l'abominable homme des neiges" ([Hintikka 1962], p. 129). La solution consistera selon Hintikka à éliminer la présupposition existentielle des énoncés quantifiés et à restreindre les règles de quantification. Nous n'entrerons pas dans les détails de cette restriction du domaine de quantification des contextes épistémiques, comparable aux effets de la logique libre et qui a suscité des réactions critiques de la part de [Castañeda X], [Lenzen 1976] ou [Girle 1975], entre autres. D'autres détails sur les restrictions de la portée existentielle de $(\exists x)$ se trouvent également dans [Hintikka 1969].

[272] La Formule de Buridan correspond à la version existentielle de la Formule de Barcan, qui porte sur un contexte quantifié universellement : $\Box(\forall x)Fx \rightarrow (\forall x)\Box Fx$.
Dans un énoncé quantifié *de dicto*, la modalité porte sur les mots : elle a une portée large et gouverne l'énoncé total (le *dictum*), y compris le quantificateur : $\Box(\exists x)Fx$. Dans un énoncé quantifié *de re*, la modalité porte sur les choses (le terme sujet) : elle a une portée étroite et ne gouverne pas le quantificateur : $(\exists x)\Box Fx$. La distinction entre les modalités *de dicto* et modalités *de re* prolonge la distinction médiévale entre le sens divisé (*in sensu diviso*) et le sens composé (*in sensu composito*) des contextes opaques ou obliques (*oratio obliquae*). Cela dit, dans [Chisholm 1963] l'auteur prétend (p. 788) que les deux couples de notions ne sont pas tout à fait analogues : les médiévaux concevaient deux sens divisés, et l'un d'eux était *de re*. Dans le sens divisé *de re*, la modalité portait sur la copule de l'énoncé et non sur le terme sujet. Or étant donné que la logique moderne, fonctionnelle, ne contient plus de copules, le sens divisé *de re* n'est plus formalisable tel quel et le sens divisé *de dicto* est associé désormais à un cas de modalité *de re*.

que son père, bien au contraire. Or il s'avère que son père et son oncle ne font qu'un : le père de Paul est vendeur d'armes et, s'il a inventé l'existence de son frère jumeau, c'est pour ne pas décevoir son fils. Nous dirons alors que, *à l'insu de* Paul, celui-ci croit de son père que c'est un philanthrope et de son oncle, c'est-à-dire de son père (via (LL)), que c'est un affreux vendeur d'armes : Paul croit ainsi de celui qui est son père que c'est à la fois un philanthrope et un misanthrope, donc que ce n'est pas un philanthrope. Bien que la contradiction vienne de l'incapacité de Paul à identifier son oncle, admettre la validité de (BF_e) devrait nous inciter à admettre la présence chez Paul de croyances contradictoires et irrationnelles, même s'il n'en est rien.[273] L'inférence est donc mise en échec, et des contre-exemples existent également pour la converse (BF_i), c'est-à-dire l'inférence *de re-de dicto*. En voici un : si un jeune militant de l'UMP croit *de* son Président de la République favori, Jacques Chirac, qu'il a toujours été un européen convaincu, alors il croit *que* l'auteur de l'appel de Cochin est un pro-européen de la première date. Cette conclusion peut être considérée comme absurde parce que contradictoire dans les termes : l'appel de Cochin fut un message anti-européen lancé par Chirac en 1976, ce qui revient à faire dire au jeune militant que Chirac est un anti-européen et un pro-européen à la fois. Certes, un tel ensemble de croyances n'aura rien de contradictoire pour le militant si que celui-ci est trop jeune pour avoir connu pas l'appel de Cochin, ou s'il a conscience de sa teneur anti-européenne mais ne sait pas que Chirac en fut l'auteur. Mais cet ensemble devient contradictoire pour lui s'il connaît bel et bien l'appel de Cochin *et* l'identité de son auteur. Sur la base de ces deux hypothèses plausibles, le jeune militant de l'UMP est supposé penser que quelqu'un peut être pro-européen et anti-européen à la fois. C'est ce que (BF_i) valide, en vertu duquel ce qui est vrai d'un individu l'est sous toutes ses désignations possibles.[274] Pour résumer les deux versions de (BF), celles-ci posent un problème commun d'identification du sujet de l'énoncé par l'agent épistémique.

Ce cas de figure met en évidence une difficulté propre à la logique modale épistémique quantifiée, dans la mesure où (BF) révèle une inconsistance fréquente entre deux ontologies distinctes : une ontologie générale, qui comprend l'ensemble des objets réels du monde extérieur ; une ontologie locale, qui comprend l'ensemble des objets réels ou fictifs du monde intérieur des agents.[275] La question centrale est celle-ci : si l'identité des individus est la condition préalable à la signifiance d'un énoncé (pas d'identité, pas de sens), comment doit s'articuler la relation entre les deux ontologies pour garantir cette identification, et en quoi cette relation affecte-t-elle la logique modale épistémique ?

---

[273] La langue naturelle contient certaines indices grammaticaux qui permettent de caractériser les formules *de dicto* et *de re* : tandis que 'croire que' est une modalité *de dicto*, parce qu'elle porte sur un énoncé (la complétive qui suit le pronom relatif 'que'), 'croire de' est une modalité *de re* parce qu'elle porte sur un objet. 'Savoir que' et 'savoir qui' se distinguent pour les mêmes raisons.

[274] (BF_i) semble plus acceptable en général que (BF_e), :mais voir [Lenzen 1978] pour d'autres contre-exemples à cette dernière (pp. 84-6). Dans l'exemple ci-desssus, le militant UMP n'aurait pas eu de croyances incohérentes s'il ne pensait pas que Chirac a *toujours* été un européen convaincu : rien ne l'empêche de changer d'opinion politique ou de faire preuve d'opportunisme, auquel cas on est en présence d'un cas de révision de l'ensemble des croyances de Chirac entre 1976 et aujourd'hui.

[275] L'ontologie locale est introduite par les domaines de référence des mondes possibles : "Lorsque nous faisons de la logique modale, nous faisons plus d'une chose à la fois. Nous utilisons certains symboles –constantes et variables –pour faire référence aux objets *réellement existants* de notre domaine de discours. Mais nous les employons pour référer également à certains états de choses qui ne se réalisent pas forcément". (Hintikka [1957], pp. 63-4).

2.3.1.1 Une 'solution' : la transquantification

Un opérateur modal est considéré comme une fonction intensionnelle, c'est-à-dire une règle de signification qui va d'un ensemble de mondes possibles vers une valeur de vérité ; en conséquence, la signification d'une formule modale exige le parcours de plusieurs domaines de référence distincts. Ces mondes peuvent inclure le monde actuel, indépendant des mondes possibles relatifs à un sujet, mais cette inclusion n'est pas essentielle lorsqu'il s'agit de transquantification. La méthode utilisée par Hintikka pour déterminer le sens des énoncés quantifiés consiste ainsi à quantifier à travers plusieurs mondes possibles, dans le cadre d'une sémantique relationnelle dont le domaine de référence se compose de valeurs de prédicats et d'individus. Tandis que la description abstraite des règles d'assignation de ces valeurs est l'affaire du logicien, l'étude des conditions effectives de l'assignation d'une valeur est celle du philosophe ; mais puisque sa logique épistémique a vocation à servir de logique philosophique, Hintikka endosse les deux habits au sein de sa logique modale épistémique et va tenter d'examiner les deux examens précédents : construire une sémantique pour la logique épistémique quantifiée, et justifier le choix des règles de signification.

Reprenons le cas de la loi d'exportation (BF_e), et montrons son rapport avec la loi de Leibniz (LL). Les analyses de Hintikka montrent que (BF_e) n'est pas valide telle quelle dans un contexte de croyance, mais qu'une condition supplémentaire peut être introduire pour valider l'inférence *de dicto-de re*. Cette condition convent aussi bien pour le logicien (épistémique) que pour le philosophe (épistémologue). Un exemple de [Hintikka 1962] porte sur l'identification du docteur Salazar, un ancien dictateur du Portugal des années 1960-70. Si Tom sait (avant 1974)[276] que le Portugal est une dictature, alors en vertu de (BF_e) Tom doit être capable d'identifier le dictateur du Portugal sous toutes ses désignations. Cette exigence est la conséquence de la quantification *de re*, qui porte sur un individu et non plus sur un *dictum* ou énoncé. Une quantification *de dicto* ne porte pas sur un individu en particulier mais sur une certaine description d'un individu quelconque : *a* sait qu'un individu gouvernait le Portugal avant 1974 en tant que dictateur, mais peu importe qui il est. En revanche, une quantification *de re* porte sur un individu précis que *a* doit connaître en tant que tel : Tom sait *qui* est le dictateur du Portugal. Les constructions grammaticales en 'savoir que' et 'savoir qui' expriment ainsi deux types de quantification distincts, et l'inférence de l'une à l'autre montre clairement l'absurdité d'une application de (BF_e) au sein d'un contexte épistémique. Si l'on part en effet de la formule *de dicto*

(8) $$K_a(\exists x)(x = b),$$

où le terme d'individu *b* symbolise la description définie 'le dictateur du Portugal', (BF_e) suppose que l'on puisse en inférer

(9) $$(\exists x)K_a(x = b),$$

[276] L'exemple donné est de 1962 et, à l'époque, Salazar était encore le dictateur du Portugal. Si nous précisons la date de l'énonciation (antérieure à la révolution des œillets de 1974), c'est parce qu'une ambiguïté pourrait être soulevée : est-il pourvu de sens de quantifier aujourd'hui sur un individu mort, qui n'existe donc plus ? L'élimination des présuppositions existentielles chez Hintikka apporte une forme de réponse au problème ; mais ce problème disparaît complètement si l'on traite les attitudes épistémiques en termes d'énonciations indexicales plutôt que d'énoncés atemporels, c'est-à-dire si la signification d'un énoncé dépend du moment auquel il est prononcé par son locuteur. L'ambiguïté soulevée ici et sa résolution annoncent l'approche illocutoire que nous proposerons plus loin, dans le cadre du problème de l'individuation des objets de croyance.

mais une telle inférence est invalide. Dans le conséquent, le quantificateur se situe *en dehors* de la portée de l'opérateur de connaissance et indique que la variable liée *x* appartient au domaine de référence de l'ontologie générale, c'est-à-dire au monde actuel. Or valider (BF_e) revient à prétendre qu'un acte de nomination est toujours appliqué avec succès par un sujet locuteur. Si (x = x) et (x = y) sont traités comme deux modes d'identification interchangeables, cela veut dire que le nom propre et la description définie se valent : on n'en apprend pas plus sur l'identité d'un individu lorsqu'il est présenté dans les termes d'une description définie ou d'un nom propre, puisque l'agent est toujours capable d'identifier correctement l'individu auquel il pense ; en d'autres termes, la validation de (BF_e) signifie en substance qu'un agent maîtrise tout aussi bien les identités *a priori* et triviales que les identités *a posteriori* et informatives. Or cela est absurde : si le dictateur du Portugal se trouve être le docteur Salazar (si b = c), alors il suffit à *a* de savoir que le Portugal est un régime dictatorial pour connaître l'identité de son dictateur,

(10) $K_a(\exists x)(x = c)$.

La raison de l'inférence (8)-(10) à partir de (9) repose sur la présupposition d'une transparence référentielle des termes d'individus *dans n'importe quel contexte de discours*, ce qui permet d'affirmer que ce qui est vrai d'un individu (disons *a*) l'est sous toute désignation de cette chose (que ce soit un nom propre, le docteur Salazar, ou une description définie, le dictateur du Portugal). Elle permet d'autre part une inférence entre les constructions verbales de type 'savoir que' et 'savoir qui', dans la mesure où le fait de savoir que quelqu'un a une certaine propriété implique le fait de savoir qui il est.

Formellement, cela signifie que pour toute modalité $\Box$, si $\Box F\phi$ et si $(\phi = \psi)$ alors $\Box F\psi$. La réponse de Hintikka à ce résultat contre-intuitif consiste à renforcer la condition d'inférence de $\Box F\phi$ à $\Box F\psi$ par le biais de $\Box(\phi = \psi)$: pour que ϕ et ψ soient F dans tous les mondes possibles, ils doivent être identiques dans chaque monde. On obtient ainsi une version renforcée et révisée de (LL) où l'identité extensionnelle devient l'identité intensionnelle, c'est-à-dire extensionnelle dans tous les mondes possibles :

(LL$\Box$) $[\Box Fa \wedge \Box(a = b)] \rightarrow \Box Fb$

Cette solution est évidente et s'applique également au cas ci-dessus : pour que *a* puisse inférer de sa connaissance du régime politique portugais (une dictature) qu'il est gouverné par Salazar, il doit savoir au préalable que le dictateur du Portugal est le docteur Salazar :

(11) $K_a(b = c)$

avant de pouvoir inférer (10) de (8). Une fois cet individu identifié, la condition (11) est satisfaite et Tom peut en inférer (10). Cette hypothèse auxiliaire permet d'éviter également les conséquences paradoxales auxquelles conduirait une interprétation transparente des modalités, par opposition à une interprétation opaque : admettre un sens transparent des modalités épistémiques ainsi que dans [Quine 1953] (voir plus bas), c'est aboutir à la conclusion que si Tom sait que Chirac est Chirac alors il sait *ipso facto* qui est Chirac même s'il ne l'a jamais vu ni entendu. D'autre part, admettre ce sens transparent des attitudes revient à fixer le nombre des prédications possibles d'un individu selon ses propriétés réelles : si un agent sait qui est un individu dès qu'il sait quoi que ce soit à son sujet, tout ce qu'il pourra prédiquer de lui devra être consistant avec les propriétés réelles de cet individu. Pour reprendre un exemple

précédent, le jeune militant de l'UMP ne pourrait pas prédiquer la propriété 'être un pro-européen de première date' de l'individu Chirac en vertu de (BF_e).[277] La transparence référentielle permettrait d'inférer indûment une connaissance *a posteriori* de la connaissance *a priori* de l'identité de soi à soi, mais l'hypothèse auxiliaire de Hintikka l'interdit et aboutit au rejet de la transparence dans un contexte épistémique. Le remède (LL□) de Hintikka est efficace et élégant, et la validation de (BF_e) à partir de cette condition renforcée semble permettre de réduire les expressions 'savoir si' en termes de 'savoir que' : je sais qui est *b* si je sais au préalable que *b* possède également telle et telle propriété, faisant de la connaissance propositionnelle une notion de base pour l'analyse formelle.[278]

Par ailleurs, la distinction entre la condition transparente de Leibniz (a = b) et la condition renforcée de Hintikka □(a = b) suppose le rejet de la *désignation rigide* de Kripke, Marcus et Føllesdal : d'après cette théorie, appelée également Nouvelle Théorie de la Référence, un nom propre n'a aucune signification particulière (aucun sens, ou mode de désignation) et sert simplement d'étiquette pour désigner un individu. Si tel est le cas et si un nom propre instaure une relation directe entre un terme et son porteur, alors il lié à sa référence dans tous les mondes possibles en vertu de sa nature purement référentielle. [Kripke 1971] exprima logiquement cette idée par le principe d'identité nécessaires des noms propres :

(=□) $(\forall x)(\forall y)[(x = y) \rightarrow \Box(x = y)]$

Les conséquences de ce principe sont moins graves que la présupposition de transparence référentielle sous toute désignation. Si on l'applique aux modalités épistémiques, on ne met certes plus le nom propre et la description définie sur un pied d'égalité puisque (=□) ne vaut que pour les cas d'identification par nom propre. En revanche, l'application de ce principe implique que l'agent épistémique ne se trompe jamais dans l'identification d'un nom propre : il peut pas ne pas savoir que Tullius est l'auteur de *De Senectute*, mais il ne peut pas ne pas savoir que Tullius désigne le même individu que Cicéron. Le contre-exemple de Tom conduit [Hintikka 1967] à rejeter la validité de (=□) par l'absurde, tout au moins lorsque □ s'interprète par K ou B :

> il n'y a pas de catégories *syntaxiques* dont les membres satisfont toujours [(11) $(\exists x)[(b = x) \wedge K_a(b = x)]$] ou [(13) $(\exists x)K_a(b = x)$]. C'est le cas par exemple de la catégorie des noms propres car on pourrait ne pas savoir qui est *b*, bien quand même le terme *b* serait un nom propre. La raison élémentaire de ceci est que le problème de savoir si un terme singulier b satisfait (11) ou (13) est un problème *contingent*.[279]

L'identification contingente de *b* et *c* par un agent *a* s'oppose donc à leur identité métaphysiquement nécessaire, pour reprendre la distinction de Kripke entre nécessité épistémique et nécessité métaphysique.

---

[277] A moins d'admettre

[278] D'autres constructions grammaticales en termes de savoir sont moins facilement analysables, toutefois, parmi lesquelles l'expression 'savoir comment' ; la distinction entre savoir propositionnel et savoir-faire a été discutée ailleurs par Ryle ainsi que [Lemmon 1967], mais elle n'entre pas dans le débat à suivre portant sur l'individuation des objets d'attitude. [Hintikka 1962] évoque rapidement les cas de 'savoir quand' et 'savoir où', proposant une quantification sur des variables de temps et de lieu.

[279] [Hintikka 1967], p. 45.

Il est donc permis de modérer cette critique de la Nouvelle Théorie de la Référence, pour au moins deux raisons. D'une part, l'identité de deux individus quelconques *b* et *c* dans tous les mondes possibles n'indique pas de soi que ces mondes soient de nature épistémique ; ils peuvent désigner pour Kripke des mondes de nature métaphysique, et [Smith 1983] le rappelle.[280] D'autre part, la théorie de la désignation rigide est conçue par [Føllesdal 1986] comme un 'idéal régulateur', c'est-à-dire comme décrivant la propension des noms propres à désigner un seul et unique individu pour tous les sujets locuteurs ; mais cet idéal n'est pas atteint par définition et n'implique pas qu'un nom propre est référentiellement transparent dans les faits.[281] Non seulement les contextes épistémiques ne sont pas extensionnels, puisqu'ils ne respectent pas les lois extensionnelles classiques et celle de Leibniz en particulier, mais ils ne sont pas intensionnels non plus : contre la règle de nécessitation du système **K**, Hintikka avait concédé que les vérités logiques ou mathématiques pouvaient ne pas être connues même si on les considère comme nécessairement vraies (dans tous les mondes *logiquement* possibles), dans la mesure où la compétence linguistique ou logique est un fait contingent relatif au sujet lui-même plutôt qu'à la nature syntaxique de la preuve. Deux énoncés logiquement équivalents ont la même valeur de vérité dans tous les mondes logiquement possibles, donc ils sont substituables dans tous ces types de mondes ; mais ils ne sont pas substituables dans tous les mondes épistémiquement possibles, comme nous l'avions exprimé tout à l'heure en termes de 'mondes possibles impossibles', ce qui signifiait que des mondes épistémiquement possibles peuvent être logiquement impossibles. Même chose pour le cas de la désignation rigide : la distinction que Kripke avait introduite dans sa *Logique des Noms Propres* entre nécessité métaphysique et nécessité épistémique, si elle permet d'affirmer que deux noms propres co-référentiels sont substituables *salva veritate* dans tous les mondes métaphysiquement possibles, ne permet pas d'en dire autant pour les mondes

[280] "Les motivations sous-jacentes à une approche de type Kripke sembleraient être métaphysiques, ou ontiques, tandis que les motivations sous-jacentes à une approche hintikkienne sont de toute évidence épistémiques. En effet, Kripke et d'autres ont considéré la question de l'identité transmondaine en rapport aux modalités ontiques de la nécessité et de la possibilité, alors que Hintikka a abordé l'identité transmondaine principalement en rapport aux modalités épistémiques, ou plutôt intentionnelles, du savoir, de la croyance et de la perception" ([Smith 1983], p. 262). L'interprétation de □ par K ne serait donc pas celle à laquelle Kripke songe dans sa formulation de (=□). Smith assimile ici les expressions 'ontique' et 'ontologique', ce qui est incorrect : la première renvoie à l'étant, c'est-à-dire l'étude de l'être (de ce qu'est un objet quelconque) en tant que perçu, jugé, senti, etc. ; la seconde renvoie à l'être, c'est-à-dire l'étude de l'objet en soi ou indépendamment des propriétés par lesquelles un sujet peut l'envisager. La confusion entre être et étant est courante, illustrée notamment par la notion de nécessité ontique (vrai 'en soi') ; nous considérerons au contraire que la modalité épistémique est une modalité ontique : elle indique l'identité d'un objet 'pour soi', c'est-à-dire du point de vue d'un sujet particulier et non en soi. Nous reviendrons sur ce lien entre 'épistémique' et 'ontique' (section 3.3).

[281] [Føllesdal 1986] admet qu'un désignateur rigide puisse ne pas être utilisé pour référer au même objet dans tous les mondes épistémiquement possibles, mais il n'en devient pas incohérent pour autant parce qu'il suggère une interprétation 'souple' de la notion de rigidité : "La rigidité ou pureté, telle que je la conçois, n'est pas incompatible avec un déplacement de référence de ce genre. A la place, je considère la rigidité comme un *idéal*, comme quelque chose de comparable à une idée kantienne, qui régule la manière dont nous utilisons le langage pour parler du monde. Il y a une *tendance à la normativité* dans notre usage du nom et d'autres termes singuliers purs, à l'idée de faire toujours de notre mieux pour garder le contact de la référence et continuer d'y référer. Parfois nous nous trompons, et l'on n'est pas sûr à la fois de ce que nous croyons et de ce que sur quoi nos croyances portent tant que nous n'avons pas mis de l'ordre ou qu'un nouvel usage n'a pas été établi. Tout notre discours sur le changement, la causalité, l'éthique, le savoir et la croyance ainsi que sur d'autres modalités présuppose que nous pouvons conserver nos termes singuliers comme référant aux mêmes objets. Dans la mesure où nous échouons, ces notions deviennent incohérentes" (p. 179).

épistémiquement possibles. Il existe à nouveau des 'mondes possibles impossibles' dans le cadre de la désignation rigide, si l'on tient compte de cas d'opacité référentielle qui sont épistémiquement possibles mais métaphysiquement impossibles.

Nécessité logique, nécessité métaphysique et nécessité épistémique sont donc trois catégories modales distinctes. Il existe certes des systèmes logiques susceptibles de combiner ces différentes sortes de mondes possibles, mais la présence d'une nécessité métaphysique ne paraît pas indispensable pour les besoins explicatifs des seules modalités épistémiques.[282] L'échec de la déduction et de la désignation entraînent ainsi deux cas d'hyperintensionalité, c'est-à-dire deux situations dans lesquelles le principe d'intensionalité n'est pas respecté. Dans la mesure où un agent peut afficher son incompétence à l'égard d'une preuve déductive ou d'un cas de désignation rigide, son incompétence logique et linguistique font du contexte épistémique un contexte hyperintensionnel par excellence.[283]

Malgré la simplicité de la solution (LL□) apportée ci-dessus à l'opacité référentielle, l'identification dans des mondes possibles pose d'autres problèmes plus profonds. L'objection faite dans [Quine 1961a] se reproduit ici : que signifie que deux gros hommes possibles ne fassent qu'un, et comment garantir l'identité des individus à travers tous les mondes possibles, en particulier épistémiques ? Notons que cette question semble porter davantage sur un problème de vérité matérielle que de vérité formelle : si l'on ne demande pas à un logicien classique de justifier la vérité d'une prémisse, alors pourquoi demander à Hintikka de justifier l'hypothèse de l'identité des référents dans tous les mondes possibles ? Cette hypothèse n'est-elle pas une simple hypothèse conditionnelle qui suffit pour les besoins de la *logique* épistémique ?

---

[282] "Cette motivation *pragmatique* de la logique épistémique fait appel à plusieurs commentaires. Tout d'abord, pour qu'elles soient considérées sérieusement afin d'agir (ou que l'on agisse sur nous), les possibilités que j'ai appelées scénarios doivent en un sens des possibilités réelles (concrètes). En conséquence, il n'est pas réaliste (n'en déplaise aux consorts de Chalmers) d'introduire une dimension séparée de possibilités épistémiques différentes de celles réelles ou métaphysiques". [Hintikka 2003], pp. 37-8. L'auteur fait référence ici à la sémantique bidimensionnelle de [Chalmers 2004], dans laquelle est reproduite la distinction de Kripke entre deux ensembles distincts de mondes métaphysiquement et épistémiquement possibles. Un monde peut être épistémiquement possible et métaphysiquement impossible, selon Chalmers, mais Hintikka reproche à cette dichotomie d'être 'irréaliste' ; un tel débat renvoie au langage comme fait social et au lien indissociable entre les aspects épistémiques et institutionnels de l'acte de référence. Hintikka adopte probablement une approche constructive et révisable de la référence, par opposition à l'aspect métaphysique qui suggère l'institution de noms propres hors de toute communauté d'agents.

[283] On ne devrait pas parler d'intensionalité et d'hyperintensionalité mais de non-vérifonctionnalité et de non-intensionalité, en vertu des deux règles suivantes :

(PS1) Si deux énoncés Q et Q' ont la même valeur de vérité dans un monde (modèle ou valuation) donné, alors l'énoncé P contenant Q comme sous-énoncé a la même valeur de vérité (dans ce monde) que l'énoncé P' que nous obtenons à partir de P en substituant Q' à Q.

Un système est intensionnel *stricto sensu*, c'est-à-dire non-extensionnel, s'il n'obéit pas au théorème de remplacement qui suit, une seconde version du principe de substitution :

(PS2) Si deux énoncés Q et Q' sont logiquement équivalents, i.e. s'ils sont vrais exactement dans les mêmes mondes, alors l'énoncé P contenant Q comme sous-énoncé est logiquement équivalent, i.e. est vrai exactement dans les mêmes mondes, à l'énoncé P' que nous obtenons à partir de P en substituant Q' à Q.

La logique épistémique de [Hintikka 1962] est non-vérifonctionnelle : deux énoncés factuels vrais peuvent recevoir des valeurs de vérité différentes dans au moins un monde épistémique ; elle est encore intensionnelle *stricto sensu*, dans la mesure où deux énoncés logiquement équivalents y conservent la même valeur de vérité dans tous les mondes possibles.

Le problème de l'identité des individus dans des mondes possibles est un problème de mise en pratique des conditions de vérité : tout comme la clause de vérité ou axiome **T**, la difficulté soulevée par [Quine 1961a] consiste à déterminer le sens dans lequel deux individus possibles désignent une seule et même valeur. Une explication de cette identité des possibles, si elle n'est pas exigée pour un logicien pur, est exigée en revanche pour un logicien philosophe ; [Kripke 1963] proposait au départ des valuations arbitraires pour les termes d'individus, dans un domaine constant et où chaque variable recevait une et une seule valeur. Mais une fois que la logique modale fut exploitée avec des ambitions philosophiques, la désignation rigide fut présentée par ce même Kripke comme une théorie de la référence capable d'illustrer ses résultats formels et de les démarquer du caractère arbitraire des modèles constants de Carnap et Marcus.[284]

Cette mise en l'épreuve de la logique épistémique quantifiée se fera en deux temps : après avoir étudié les arguments de Quine contre le principe d'identification dans des mondes possibles, ainsi les solutions que celui-ci proposa en échange, nous reviendrons sur l'approche des logiciens modaux et la justifierons sans que ce résultat ne paraisse incompatible avec le discours de Quine : bien que ce dernier se soit présenté comme un farouche adversaire de la logique modale sous tous ses états, quantifiés ou non, nous montrerons que la notion modale épistémique de 'fonction individuante' concorde avec la théorie de la connaissance de Quine et ne fait que formaliser son épistémologie naturalisée.

2.3.1.2 Les obstacles selon Quine

Pour Quine, (LL) et (GE) apparaissent comme deux lois 'inviolables' de la logique classique que les modalités contrarient ; conformément à sa stratégie de mutilation minimum, le maintien de ces deux lois est préférable à tout ce qui peut les contester, donc la logique modale doit être abandonnée dans son ensemble en raison de son coût théorique inacceptable : non seulement son calcul des propositions est inutile, mais son calcul des prédicats est inutilisable sans le recours à un essentialisme qui, du point de vue de Quine, conduit le logicien modal de Charybde en Scylla. Mais pourquoi considérer ce rejet comme inacceptable ? Parce que la substitution *salva veritate* des termes co-référentiels, pierre de touche de l'analyse logique classique, est mise en échec dans les énoncés modaux, qu'ils soient aléthiques ou épistémiques. Quine considère ces échecs comme le symptôme du phénomène d'*opacité référentielle* : les termes référentiels n'ont plus leur référence habituelle dans ces contextes modaux, selon Frege ; selon Quine, ils n'ont plus de référence du tout ; nous savons qu'ils en ont plusieurs, pour Hintikka, mais Quine voit dans cette multiplication des références une autre raison de refuser la logique modale : elle entraîne une inflation ontologique, une 'jungle aristotélicienne' peuplée d'habitants au statut ontologique particulier (réels sans être actuels, inexistants sans être absents) mais avec laquelle l'austérité théorique de Quine ne peut s'accommoder.

Quelques exemples d'opacité référentielle reviennent de façon récurrente dans les articles de Quine sur la logique modale. L'extension du terme singulier '9' est un objet

[284] Sur la théorie de la désignation rigide de Kripke, voir son ouvrage de référence *La logique des noms propres* (version originale: *Naming and Necessity*), Les Editions de Minuit, Paris (1982) ; sur les références de Carnap et Barcan, voir section 1.1.1 note 13.

nécessairement supérieur au nombre 7, mais l'extension du terme singulier 'le nombre des planètes', bien que désignant le même objet que le terme '9', est supérieure à 9 mais pas nécessairement. L'introduction de l'adverbe 'nécessairement' produit ainsi contre un contre-exemple à la (LL). Bien que 'Cicéron' et 'Tullius' désignent le même individu, celui qui écrivit *De Senectute*, Tom peut croire que l'énoncé 'Cicéron a écrit *De Senectute*' et ne pas croire l'énoncé 'Tullius a écrit *De Senectute*', dont le terme singulier est pourtant co-référentiel à celui du premier énoncé.

Faut-il en conclure que (LL) n'est pas valide pour tous les types d'énoncés et que, par conséquent, elle n'est pas une loi logique universellement acceptable, c'est-à-dire applicable à n'importe quel phrase déclarative pourvue d'une valeur de vérité et dans n'importe quel contexte d'énonciation ? Contrairement à [Hintikka 1962], qui opte pour cette conclusion négative, l'attitude de Quine consiste au contraire à maintenir à tout prix l'intégralité des principes de la logique classique en l'état. Pour ce faire, il propose plusieurs méthodes de paraphrase des énoncés modaux. [Quine 1953] distingue trois degrés d'engagement modal : la modalité fonctionne soit comme un prédicat attaché à un terme singulier ('2 + 2 = 4' est nécessaire), soit comme un opérateur attaché à un énoncé clos (il est nécessaire que 2 + 2 = 4), soit comme un opérateur attaché à un énoncé ouvert (quelque chose est nécessairement égal à 4).

Les deux premiers types d'engagement modal sont inoffensifs aux yeux de Quine pour sa théorie des fonctions de vérité : que l'expression de la nécessité soit présentée comme un prédicat ou comme un adverbe ne constitue pas de différence notoire pour la structure logique de la formule, et une réduction de la nécessité au rôle de prédicat peut être opérée dans les deux cas. Un exemple du premier degré d'engagement modal est 'Il est nécessaire que 2 + 2 = 4', de forme logique 'Nec(⌜p⌝)'. Si l'on ne symbolise pas ici le verbe 'être nécessaire' par un opérateur modal de type □, c'est parce qu'il n'en est pas un. Une analyse logico-modale inciterait à distinguer ainsi les termes de la formule ci-dessus : '2 + 2' est une description définie et un argument, '= 4' est une propriété et une fonction à une place, et 'être nécessaire' est un opérateur appliqué à l'énoncé de structure fonction-argument '2 + 2 = 4'.[285] Mais l'analyse de Quine propose une autre forme logique sans introduire d'opérateur modal. Dans l'énoncé modal ci-dessus, en effet, '2 + 2 = 4' ne représente pas un énoncé à part entière mais une partie d'un énoncé dont la modalité est le prédicat, donc Nec(⌜p⌝) équivaut à une simple prédication de type Fx où *x* symbolise le nom d'énoncé '2 + 2 = 4' et F le prédicat sémantique 'être nécessaire'. Si la citation d'énoncé ⌜p⌝ n'a pas ici la fonction syntaxique de l'énoncé p, c'est parce que le symbole 'p' n'est pas *utilisé* mais *mentionné* : il joue le rôle d'un terme singulier (un nom d'énoncé) dans la formule modale d'ensemble et, de même qu'un terme singulier n'est ni vrai ni faux en tant que tel, un nom d'énoncé n'est ni vrai ni faux mais contribue à la vérité ou fausseté de l'énoncé dont il fait partie. La forme logique 'Nec(⌜p⌝)' ne déroge donc pas aux règles de valuations de la logique extensionnelle, malgré des apparences symboliques contraires. Un exemple du second type d'engagement modal de type

[285] Il serait plus correct d'interpréter l'énoncé *d'identité* '2 + 2 = 4' comme une fonction à deux places Fxy (plutôt que comme une fonction à une place Fx) : si 'être égal' est la fonction de l'énoncé, alors '4' ne fait plus partie de la fonction mais devient un argument indépendant, et la fonction relie alors les deux arguments '4' et '2 + 2'. Peu importe que la fonction soit monadique ou polyadique ici, l'essentiel ayant été de transformer l'opérateur modal en un prédicat et d'éliminer la non-vérifonctionnalité.

'Nécessairement 2 + 2 = 4', où l'expression modale n'est plus un simple prédicat mais redevient un opérateur à une place appliqué à un énoncé. On retrouve ici la forme symbolique de la logique modale □p, où '2 + 2 = 4' retrouve sa fonction syntaxique d'énoncé p. Pour éviter le problème de la non-vérifonctionnalité tout en admettant la présence ici d'un opérateur modal, Quine propose la même paraphrase que pour le premier type d'engagement modal, à savoir la technique de citation ou de paraphrase métalinguistique pour donner à ce second engagement modal la forme inoffensive du premier : 'Nécessairement' est réinterprété comme un prédicat attaché au nom d'énoncé '2 + 2 = 4' ; la forme logique de 'Nécessairement 2 + 2 = 4' ne donne donc plus la fonction intensionnelle Lp, où l'adverbe 'nécessairement' est symbolisé comme un opérateur appliqué à un énoncé, elle est désormais la fonction non-modale extensionnelle N('2 + 2 = 4') de forme N*x*, où *x* est le nom d'énoncé '2 + 2 = 4' et N le prédicat sémantique 'être nécessaire' plutôt qu'un opérateur propositionnel. Désormais, l'argument *x* a remplacé l'énoncé p du logicien modal et la fonction N son opérateur intensionnel L. Puisque cette paraphrase permet à Quine d'éliminer un grand nombre de cas de non-vérifonctionnalité des énoncés naturels, revenons sur son principe et ses adeptes.

La paraphrase métalinguistique consiste à modifier la structure grammaticale d'un énoncé, mais à condition qu'une telle transformation puisse préserver sa signification. Lorsqu'une modalité s'applique à un énoncé entier (le *dictum* des logiciens médiévaux), elle fonctionne *de dicto* et peut être soumise à la réduction citationnelle de Quine. Avant d'admettre la sémantique modale et les opérateurs intensionnels, [Carnap 1937] proposait le même type d'analyse en exprimant la modalité nécessaire L par le prédicat 'est analytique' (nécessairement vrai, ou L-vrai). Quine interprète l'adverbe de nécessité comme un prédicat sémantique comparable à l'expression 'être vrai', et la comparaison faite par [Quine 1953a] entre sa paraphrase citationnelle et la méthode de concaténation de [Tarski 1944] illustre bien le but de la manœuvre : en vertu du critère d'adéquation matérielle de la définition de la vérité, l'énoncé 'la neige est blanche' est vrai si et seulement la neige est blanche, cette définition ne devant pas être interprétée comme une relation entre des mots et des choses mais comme une relation entre une expression du métalangage et une expression du langage-objet, c'est-à-dire entre un nom d'énoncé (entre guillemets) et un énoncé (sans guillemets). Pour mettre en évidence la différence entre les deux expressions, Tarski propose la méthode de concaténation qui consiste à épeler chacune des lettres des mots de l'énoncé ; dans ce cas précis, le nom d'énoncé 'la neige est blanche' correspond aux séquence de lettres suivantes : l-a, n-e-i-g-e, e-s-t, et b-l-a-n-c-h-e. La définition de la vérité consiste à dire alors de cette séquence ordonnée de lettres qu'elle exprime dans la langue française (le langage-objet) le nom d'un énoncé qui est vrai si et seulement si la neige est blanche. Or il n'est pas permis de dissocier les séquences de lettres ci-dessus sans perdre la définition de la vérité : c'est la séquence ordonnée totale qui reçoit le prédicat sémantique de vérité car ni 'la neige', ni 'est blanche' ne sont vrais ou faux séparément. Quine traite de la même manière l'énoncé 'Nécessairement 2 + 2 = 4' : l'énoncé vrai dont il s'agit ici n'est pas '2 + 2 = 4' mais 'nécessairement 2 + 2 = 4', c'est-à-dire 'Il est nécessaire que 2 + 2 = 4' ; '2 + 2 = 4' est un nom d'énoncé et joue le rôle d'un terme singulier dont on prédique la nécessité. L'adverbe de nécessité joue donc le rôle de prédicat à l'intérieur de cet énoncé modal, tout comme 'être

blanc' à l'intérieur du nom d'énoncé 'la neige est blanche'. Quine conclut de cette comparaison que la modalité de nécessité ne doit pas être dissociée ici de la séquence de mots qui compose l'énoncé : lorsque, dans un énoncé modal de forme L(F*a*), la citation transforme l'énoncé F*a* en un nom d'énoncé 'F*a*', l'ancien opérateur L attaché à un énoncé devient un prédicat attaché désormais à un nom d'énoncé, c'est-à-dire un terme singulier, et l'ancien terme singulier *a* devient une partie de terme singulier comparable au phonème 'n' dans le terme lexical 'n-e-u-f' ; or puisque la substitution n'est permise qu'entre des énoncés ayant la même signification (la même valeur de vérité) et non entre des parties d'énoncés dépourvus de sens, le résultat de la paraphrase citationnelle est le même que chez [Frege 1892] car il n'y a plus de référence à substituer dans ce cas de figure : on ne substitue pas plus un nom d'énoncé à un autre ('2 + 2 = 4' à 'Paris est la capitale de la France', dans l'énoncé modal 'Nécessairement 2 + = 4') qu'on ne substitue un phonème à un autre ('H_20' à 'eau', dans le terme singulier 'château'). Quine compare également un nom d'énoncé dans une citation à la relation entre le terme singulier 'château' et le phonème 'eau' qui en est une partie phonétique (et non sémantique) : 'eau' ne symbolise pas un terme singulier substituable lorsqu'il est utilisé comme un simple phonème dans le mot 'château', bien qu'il puisse avoir un sens par ailleurs et désigner le liquide H_20 lorsqu'il est utilisé comme un terme singulier.

La différence entre un phonème (la partie sonore d'un mot : 'eau' dans 'château') et un terme lexical (le mot complet : 'château') est comparée ici à la différence entre une pensée incomplète ('que 2 + 2 = 4') et une pensée complète ('Il est nécessaire que 2 + 2 = 4') : il y a une relation de partie à tout dans les deux cas, et la notion de nécessité doit être traitée comme une partie d'énoncé plutôt que comme un opérateur appliqué à un énoncé. Si les logiciens modaux ont traité l'adverbe de nécessité comme un opérateur indépendant de l'énoncé, pense Quine, c'est parce que, contrairement au phonème qui est dépourvu de sens dans la plupart des cas, une pensée incomplète est exprimée par une proposition subordonnée qui, elle, peut avoir un sens indépendamment de la principale à laquelle elle est subordonnée. En effet, substituer 'H_20' à 'eau' fait du terme singulier 'château' un terme *dé*pourvu de sens ('châtH$_2$0'), tandis que la substitution de 'Paris est la capitale de la France' à '2 + 2 = 4' fait de l'énoncé 'Nécessairement 2 + 2 = 4' non pas dépourvu de sens mais simplement faux ('Nécessairement Paris est la capitale de la France'). La confusion sémantique entre une pensée incomplète et une pensée complète est moins grossière que la confusion sémiotique entre un phonème et un terme lexical, mais les deux cas de figure représentent un même malentendu sur la fonction d'un signe : dans le terme lexical 'château', le phonème 'eau' ne doit pas être confondu avec l'autre terme lexical qui désigne le liquide H_20, de même que le prédicat de nécessité ne doit pas être confondu avec un opérateur d'énoncé. Quine a trouvé dans cette méthode de citation un moyen efficace pour distinguer les deux prédicats sémantiques de vérité et de nécessité : seul le premier s'applique à un énoncé, et le second s'applique seulement à des noms d'énoncés. D'autre part, cette paraphrase lui permet d'aligner les contextes intensionnels sur la notation des contextes extensionnels sans modifier le sens de l'énoncé de départ. L'occurrence de la notion de nécessité n'est donc pas essentielle dans ces deux premiers types d'engagement modal : il n'y joue pas le rôle d'un opérateur mais d'un prédicat, donc la nécessité n'y figure pas comme une constante logique.

Il semble en aller autrement dans le troisième type d'engagement modal, lorsque la variable d'un énoncé ouvert est liée à un quantificateur. Tant qu'elle portait sur un énoncé clos de type F*a*, la paraphrase métalinguistique était permise ; mais elle ne l'est plus dès qu'elle porte sur un énoncé ouvert de type F*x*, comme nous allons l'expliquer ci-dessous. Or l'analyse des fonctions propositionnelles est la seule qui importe vraiment pour Quine, dans la mesure où tout énoncé clos porte contient des constantes d'individus qui doivent être réduites après analyse à des variables liées. Dans une théorie référentielle de la signification, la valeur sémantique (de vérité) d'un énoncé suppose l'existence de son terme singulier et, donc, l'assignation d'une référence pour ce terme ("être, c'est être la valeur d'une variable liée").[286] En raison de cet engagement ontologique dont dépend la valeur de vérité de chaque énoncé, le véritable problème de l'engagement modal concerne les modalités *de re*, lorsque la variable liée se situe *à l'extérieur de* la portée de l'opérateur modal : $(\forall x/\exists x)\Box Fx$. Dans ce troisième type d'engagement modal, la nécessité est un opérateur attaché à une fonction propositionnelle telle que $x > 5$ dans $(\exists x)\Box(x > 5)$. Or si l'on applique maintenant la même paraphrase citationnelle que précédemment, on obtient une formule $(\exists x)N('x > 5')$ qui est dépourvue de sens parce que syntaxiquement incorrecte : la variable *x* liée au quantificateur existentiel n'a aucun rapport sémantique avec le *signe* 'x' du nom d'énoncé 'x > 5'. La méthode de paraphrase trouve donc ses limites d'application face à cette troisième forme d'engagement modal, la seule qui importe vraiment dans une interprétation référentielle de la quantification. La paraphrase citationnelle dont use Quine pour extensionaliser les formules modales peut être appliquée avec des modalités aussi bien aléthiques qu'épistémiques : si l'on remplace 'être nécessaire' par le prédicat 'être cru par Tom', l'analyse de 'Tom croit que Cicéron a dénoncé Catilina' donne ainsi 'Tom croit l'énoncé "Cicéron a dénoncé Catilina"'. A noter que la relation directe évoquée par certains (Russell, notamment) entre le sujet croyant et le ou les objets de son attitude (Cicéron et Catilina, ici) disparaît avec cette paraphrase nominale, puisque les deux célèbres individus sont réduits à des noms d'individu au sein d'un nom d'énoncé. Cette situation tend à confirmer l'opacité référentielle et l'idée qu'un 'objet' de croyance n'apparaît pas en toute transparence ou n'est pas en relation directe avec le sujet, comme le prétendait notamment la théorie réaliste du jugement de Russell (voir la note du §5.542 du *Tractatus*, où cette théorie de la relation directe est critiquée).

La difficulté des modalités *de re* réside dans les conditions d'identification des objets quantifiés, et cette difficulté n'apparaissait pas dans le cas des modalités *de dicto* (le second type d'engagement modal ci-dessus). [Quine 1956] distingue les deux types de modalités en termes de sens *notionnels* et de sens *relationnels* de la modalité : la modalité *de dicto* est notionnelle, elle correspond à ce que [Donnellan 1966] appellera l'usage 'attributif' des descriptions définies ; la modalité *de re* est relationnelle, elle correspond à ce que [Donnellan 1966] appellera l'usage 'référentiel' des descriptions définies.[287] Un exemple de [Quine 1956]

[286] "Le phénomène d'opacité référentielle vient d'être expliqué en faisant appel au comportement des termes singuliers. Mais les termes singuliers sont éliminables, nous le savons (...) par la paraphrase (...) si l'opacité est une infirmité dont il vaut la peine de se soucier, elle doit montrer des symptômes aussi bien en rapport à la quantification qu'aux termes singuliers." ([Quine 1953a], pp. 144-5).

[287] Voir [Donnellan 1966]. La distinction de Donnellan est pertinente dans un contexte modal, puisqu'elle permet d'évaluer l'énoncé en fonction du type d'information fourni par le terme singulier : soit pour indiquer une propriété d'un individu, soit pour faire référence à cet individu lui-même. Mais une telle distinction n'est pas pertinente dans une analyse à la Quine, dans la mesure où les descriptions définies finissent tôt ou tard par être

illustre une attitude propositionnelle dans laquelle (LL) et (GE□) sont mises en échec par les modalités épistémiques :

(a) Ralph croit que l'homme au chapeau brun est un espion

donc

(b) il y a quelqu'un dont Ralph croit que c'est est un espion,

car (1) implique (2) en vertu de (GE□). Dans le même temps,

(c) Ralph ne croit pas qu'un certain homme aux cheveux gris qu'il a aperçu sur la plage est un espion.

Or à son insu, il se trouve que

(d) l'homme au chapeau brun et l'homme qu'il a aperçu sur la plage sont un seul et même individu.

Contrairement à une interprétation référentielle de la quantification, dans laquelle la valeur de vérité des énoncés quantifiés est préservée lorsque l'on substitue deux termes co-référentiels, l'énoncé vrai 'Ralph croit que l'homme au chapeau brun est un espion' devient faux si le terme singulier 'l'homme aperçu sur la plage' est substitué au terme singulier co-référentiel 'l'homme au chapeau brun'. En vertu de (LL), la vérité de (a) et (d) implique la fausseté de (c), c'est-à-dire la vérité de :

(c') Ralph croit que l'homme aperçu sur la plage est un espion.

Par conséquent, Ralph tient les croyances contradictoires (c) et (c') si l'interprétation référentielle de la quantification est maintenue.[288]

Un exemple similaire est donné pour les modalités aléthiques dans [Quine 1953] : si le nombre 9 est nécessairement supérieur à 5, alors il y a quelque chose qui est nécessairement supérieur à 5 ; mais ce quelque chose, désigné également par le terme singulier 'le nombre des planètes', donne l'énoncé faux 'le nombre des planètes est nécessairement supérieur à 5' après substitution de '9' par 'le nombre des planètes'. La solution fregéenne du contexte indirect de discours est inutile ici puisque l'objet, la *res* nécessairement supérieure à 5 est une dénotation habituelle dans un énoncé modal *de re*. Dans les deux types de modalité, la valeur de vérité de l'énoncé n'est donc pas préservée après substitution des termes co-référentiels. [Quine 1956] constate ainsi que, sauf exception,[289] les modalités n'entrent pas dans le cadre de sa logique extensionnelle. De deux choses l'une : soit un énoncé impossible à enrégimenter dans sa notation canonique est considéré comme dépourvu de sens ; soit cette notation est trop pauvre pour tenir compte de la richesse expressive des langues naturelles et doit être étendue.

éliminées. Dans sa logique extensionnelle, toute information sur une propriété de l'individu est fournie par les prédicats, et le choix du terme singulier ne compte pas.

[288] L'interprétation substitutionnelle de la quantification représente une alternative concevable pour maintenir la loi de généralisation existentielle dans un contexte modal. Elle a été proposée notamment par Kripke et des logiciens libre pour évaluer les énoncés portant sur des objets fictifs. Dans "Reply to Professor Marcus", [Quine 1997] explique son refus de la quantification substitutionnelle. Le lien que Hintikka établit entre théorie de la signification et théorie de la référence explique son attachement à la quantification référentielle.

[289] [Quine 1956] suppose dans le cas des modalités épistémiques la validité du principe d'exportation (inférence *de dicto-de re*) : "Le type d'exportation qui conduit de [la modalité *de dicto*] à [la modalité *de re*] devrait sans doute être conçu en général comme une implication" (p. 188). Cette implication correspond à un sens des formules de Barcan/de Buridan (équivalence *de dicto-de re*), respectivement : $\Box(\forall/\exists x)Fx \equiv (\forall/\exists x)\Box Fx$. Pour une critique des deux sens de l'implication en logique épistémique, voir [Lenzen 1978], pp. 84-6.

Face à ces difficultés d'analyse, la réponse de Quine n'est pas la même selon qu'il traite la modalité aléthique ou la modalité épistémique : les premières peuvent être éliminées selon lui du discours scientifique, et Quine reproduit de la sorte l'attitude précédente de Frege et Russell face aux modalités aléthiques qu'ils considéraient comme des attitudes subjectives superflues pour la sémantique. Si Quine propose d'éliminer l'expression de nécessité de sa notation canonique, c'est aussi parce qu'elle conduit selon lui à un retour rétrograde et superflu à l'essentialisme aristotélicien.[290] Qui plus est, la distinction entre des propriétés essentielles et des propriétés contingentes peut créer de nouvelles ambiguïtés. [Quine 1960] cite l'exemple d'un mathématicien amateur de cyclisme ; en tant que tel, cet individu satisfait à la fois les propriétés d'être rationnel et d'être bipède. Mais il ne les satisfait nécessairement qu'en tant que mathématicien, et non en tant que cycliste : un mathématicien est nécessairement rationnel, un cycliste est nécessairement bipède, ce sont là deux énoncés dont on pourrait admettre la vérité. Mais l'élimination des descriptions définies dans la notation canonique crée un problème pour ces énoncés, dans la mesure où les propriétés d'être rationnel et d'être bipède ne sont pas nécessairement vraies de l'individu lui-même mais de l'individu *en tant que* mathématicien ou *en tant* que cycliste : un cycliste n'est pas nécessairement rationnel, un mathématicien n'est pas nécessairement bipède. Cette distinction, si elle donne tout son intérêt à l'expression d'usage attributif de Donnellan, n'en a aucun dans la notation canonique de Quine puisque l'occurrence des individus y est purement référentielle et porte sur des variables liées, indépendamment de leurs propriétés. Des vérités nécessaires peuvent-elles être admises indépendamment de cette spécification ? C'est ce que pensent en tout cas les théoriciens de la Nouvelle Référence (Kripke, Føllesdal, Barcan), selon lesquels les individus ont des propriétés essentielles en tant que telles.[291] Nous reviendrons encore sur cette théorie de la référence pure des noms propres qui consiste, contrairement à Quine, à considérer ces noms propres comme des termes purement référentiels et similaires à ses variables liées.

Dans le cas des modalités épistémiques, l'attitude de Quine est plus réservée que pour les modalités aléthiques parce qu'il les considère comme indispensables à la pratique des sciences :

> On peut abandonner la logique modale sans grand dommage (...) Mais on voit mal comment se passer des expressions d'attitude propositionnelle. Elles ont un caractère intentionnel, comme on dit, et renvoient à un acte projeté ou à un contenu de pensée dont on ne peut pas donner une analyse en termes naturalistes ordinaires. Telle était en tout cas la position de Brentano il y a deux générations, et telle est celle de Bergmann et de Chisholm aujourd'hui ; et je suis en peine d'argument satisfaisant à leur proposer.[292]

---

[290] "S'agissant de l'utilité, il y a moins à dire pour la modalité que pour les attitudes propositionnelles. L'expression remplit un but dans le discours quotidien, mais d'un genre superficiel (...) Une telle utilité est locale, passagère et sans problème, tout comme l'utilité des expressions indexicales. La sublimité de la vérité nécessaire ne devient pas tout à fait poussière, mais argile passablement banale" ([Quine 1993], p. 108).

[291] "L'essentialisme aristotélicien est précisément ce à quoi se ramène la substitutivité adoptée sans restriction dans les contextes modaux. Si une propriété appartient nécessairement à un objet, elle lui appartient nécessairement, sans égard pour la manière particulière dont on renvoie à l'objet". (D. Føllesdal : *Referential Opacity*, Oslo (1966), p. 120).

[292] [Quine 1964], p. 197.

En tant qu'adversaire d'une théorie scientifique de l'intentionnalité, Quine va tenter de trouver un moyen d'extensionaliser les modalités épistémiques, c'est-à-dire de leur donner une interprétation valable dans les limites imposées par la "notation canonique de la langue scientifique" ; ce langage n'est pas un langage purement formel et comporte un lexique restreint à la science privilégiée par Quine, la physique, et tout l'enjeu posé par la logique modale est de savoir si les énoncés épistémiques peuvent être analysés en termes psychophysiques, donc purement référentiels et extensionnels. En plus d'être non triviales, les modalités épistémiques sont des types d'intensions plus difficiles encore à traiter que les modalités aléthiques :

> D'un certain point de vue le problème des attitudes propositionnelles est plus redoutable que celui de la logique modale [aléthique], car l'essentialisme n'en viendra pas à bout. La division entre les attributs du célèbre Romain qui incitent Tom à croire qu'il a dénoncé Catilina et ceux qui l'en font douter, n'est même pas déterminée, fût-ce obscurément, par l'essence de ce personnage antique ; elle dépend aussi de Tom.[293]

Les solutions utilisées pour les modalités aléthiques échouent donc dans ce cas de figure ; la raison à cela est que la vérité d'un énoncé modal épistémique, comme Quine vient de le souligner, dépend en partie des croyances personnelles de Tom sur l'objet de son attitude et pas simplement des propriétés générales (impersonnelles) de cet objet. Or si de telles attitudes sont propres à un individu, il en ressort des difficultés pour déterminer les conditions de vérité générales d'un énoncé de croyance.

Nous avons vu précédemment que la sémantique relationnelle de Hintikka propose une solution en termes de mondes possibles, c'est-à-dire d'états de choses imaginés par un agent particulier. Mais la notation canonique de Quine ne peut admettre ces mondes s'ils ne sont pas explicables en termes physicalistes. Qu'est-ce qu'un monde possible pour Quine ? Un ensemble abstrait dont les conditions d'individuation de ses membres (et de lui-même) ne sont pas justifiées et qui, pour cette raison, ne spécifie que dans l'abstrait (formellement) la référence des termes d'individu. Quine avait déjà formulé ses doutes sur la notion de monde possible, en ce qui concerne l'identité de ses individus. Si ces individus ont leur référence fixée d'un monde à un autre, alors il devient possible de déterminer si la croyance de Tom porte effectivement sur un même individu lorsqu'il croit quelque chose des termes singuliers 'Cicéron' et 'Tullius'. Mais comment justifier cette fixation d'une référence à un terme à travers différents mondes possibles ? Quine doute qu'une telle individuation puisse être faite dans le cadre d'une langue scientifique ; pour reprendre l'exemple de [Quine 1961a],[294] l'auteur prétend que l'identification des individus entre deux mondes possibles distincts est

---

[293] [Quine 1964], pp. 196-7.

[294] "Prenez par exemple ce gros homme possible dans cette embrasure de la porte, et puis cet homme chauve possible dans la même embrasure. Sont-ils le même homme possible, ou deux hommes possibles ? Comment en décidons-nous ? Combien d'hommes possibles dans l'embrasure ? Parmi les possibles, les maigres sont-ils plus nombreux que les gros ? Combien sont semblables ? Ou le fait d'être semblable les rendrait-il un ? N'y a-t-il pas *deux* possibles semblables ? Est-ce la même chose que de dire qu'il est impossible pour deux choses d'être semblables ? Ou, en fin de compte, le concept d'identité est-il simplement inapplicable aux possibles non réalisés ? Mais quel sens peut-on trouver à parler d'entités dont on ne peut pas dire de façon significative qu'elles sont identiques entre elles et distinctes des autres ? Ces éléments sont presque incorrigibles" ([Quine 1961a], p. 4.

impossible à effectuer sans décréter l'utilisation d'une fonction d'interprétation arbitraire, dont le rôle serait d'assigner une même valeur entre plusieurs mondes. Avant d'introduire sa théorie des désignateurs rigides, [Kripke 1963] avait concédé le fait que l'individuation des objets dans un monde possible s'apparentait à une affaire de décision arbitraire de la part du logicien.[295] Si une sémantique relationnelle se résume à la construction *ad hoc* d'un formalisme dont l'interprétation ne répond d'aucun motif précis, il en résulte que le discours des mondes possibles est inacceptable dans le cadre de la logique selon Quine.[296] Il existe plusieurs modes d'individuation dans une sémantique des mondes possibles : la théorie des contreparties de Lewis, l'eccéisme de Plantinga ou les contrefactuels proches de Nozick représentent trois types d'interprétation –parmi d'autres– des règles d'interprétation relatives aux modalités aléthiques, mais rien ne semble les justifier par rapport à la décision leibnizienne de refuser toute individuation *à travers* des mondes possibles. Quine exige un motif pour la transquantification ; comment décider lorsque Tom identifie le célèbre orateur Romain entre plusieurs états de choses possibles, et comment y parvenir sans prendre en compte plusieurs modes d'individuation variables selon le contexte de discours ? Parce que Quine estime que la logique épistémique quantifiée ne peut évaluer ses énoncés sans tenir compte des conditions d'usage des termes singuliers, il en conclut que, tout comme les termes indexicaux,

> la logique et la sémantique de la nécessité et des attitudes propositionnelles n'a d'intérêt qu'à titre d'étude en linguistique et en psychologie, et non à titre de reflet des structures générales de la réalité.[297]

La logique modale est concevable, preuve en étant faite par les diverses constructions formelles proposées depuis la fin des années 1950 ; mais elle n'a pas droit de cité dans la philosophie de la logique de Quine, basée sur l'enrégimentation de tous les idiomes ordinaires susceptibles d'intéresser le discours scientifique.[298]

---

[295] "Intuitivement, supposons que P(x) signifie 'x est chauve' –devons-nous assigner une valeur de vérité à l'instance de substitution 'Sherlock Holmes est chauve' ? Holmes n'existe pas, mais dans d'autres d'états de choses il aurait existé. Devrions-nous assigner une valeur de vérité définie à l'affirmation (disant) qu'il est chauve, ou non ? Frege et Strawson n'assigneraient pas de valeur de vérité à l'affirmation ; Russell le ferait. Pour les besoins de la logique modale, nous soutenons que les différentes réponses à cette question représentent des *conventions* alternatives. Toutes sont défendables" ([Kripke 1963], pp. 85-6).

[296] La difficulté pour la logique modale n'est pas tant de définir des règles sémantiques que de justifier leur fonctionnement : "Bref, l'instrument créé est purement abstrait et son usage est pernicieux dès qu'on lui confère la mission de résoudre d'authentiques problèmes ontologiques, alors qu'il n'est que la caution scientifique d'une mythologie sans fondement qui juxtapose à ce qui existe actuellement des individus purement virtuels, renvoie à la 'jungle de l'essentialisme aristotélicien' et, immergeant l'univers réel dans une pluralité de mondes possibles, incite à confondre l'imaginaire avec le réel, détruisant ainsi les fondements de la pensée objective." (M. Boudot : *Dictionnaire des philosophes*, "Hintikka", pp. 1215-1220, en part. p. 1218).

[297] [Quine 1986], p. 227. [Quine 1976] répète cet argument indexical pour exclure les attitudes propositionnelles de la liste des constantes logiques : "La notion de savoir qui est quelqu'un, ou ce qu'est quelque chose, n'a de sens qu'à la lumière de la situation. Elle dépend totalement de la question plus spécifique que l'on a pu avoir en tête. En voyant le rôle central que joue cette notion dans la logique de la croyance de Hintikka, je me sens davantage conforté dans mon attitude récalcitrante" (p. 863).

[298] La logique reste liée à un souci d'universalité chez Quine, qui n'est pas un pluraliste : "L'idée de Quine, c'est que la logique doit être universelle, dans le sens de Ryle qui écrit : 'Une partie de ce qui caractérise les termes sur lesquels reposent les raisonnements (de la logique formelle), c'est que ces termes sont indifférents par rapport au sujet traité, c'est-à-dire *topic-neutral*' (Ryle, *Dilemmas*, Cambridge Univ. Press. (1954), p. 115). Si

L'idée de monde possibles n'est pas totalement exclue de l'œuvre de Quine, néanmoins;[299] on trouve dans [Quine 1968] une tentative de naturaliser les attitudes propositionnelles en termes de coupes de mondes possibles. Ces coupes de mondes, que Quine utilise pour redéfinir les objets de croyance en termes physiologiques de classes de stimulations sensorielles éprouvées par le sujet de l'attitude, ont pour but d'éliminer le discours phénoménologique hérité de Brentano et Meinong, d'une part, puis de définir des attributions de croyances au sein d'un état de choses centré, ce qui correspond à l'ensemble des capteurs sensoriels susceptibles d'être activés lors d'une croyance, d'un doute ou d'autres attitudes décrites d'habitude en termes de propositions préfabriquées.[300] Nous sommes ici dans la théorie de la signification-stimulus de Quine, où la proposition linguistique exprime en dernière analyse les réactions physiologiques d'un sujet aux informations sensorielles de son environnement –les animaux dépourvus d'entités linguistiques ont également des croyances, rappelle Quine et, en ce sens, les modalités *de dicto* sont des attitudes réservées aux seuls êtres humains. Ce faisant, la notation canonique choisie ici implique un rejet de la forme logique traditionnelle de l'attitude propositionnelle, considérée d'habitude comme une relation binaire entre un sujet et une proposition préétablie. Nous avions vu plus tôt que la proposition fregéenne ne pouvait être admise sans admettre dans le même temps des entités abstraites telles que les *Sinne*. Mais aucune hypostase de ce genre n'est concevable dans le cadre d'une épistémologie empiriste à la Quine, dont la thèse d'indétermination de la traduction interdit par ailleurs de postuler une identité analytique arbitraire entre les énoncés de langues naturelles et une Proposition majuscule, c'est-à-dire un *Satz an sich* à la Bolzano.

Au contraire, il s'agit de reconstruire les étapes par lesquelles les sujets construisent ces entités linguistiques publiques au sein d'une épistémologie empiriste ; c'est ce que Quine propose dans sa théorie de l'apprentissage du langage et de l'acquisition de la référence, dans laquelle les locuteurs passent d'énoncés holophrastiques à des énoncés observationnels puis à des énoncés généraux, ces énoncés théoriques que postule d'emblée la forme logique modale de l'attitude propositionnelle. A partir de cette procédure rigoureuse de déconstruction linguistique suivie d'une reconstruction physiologique, Quine tente de remonter à la source des objets de croyances afin de déterminer leurs conditions d'individuation : la solution à l'opacité référentielle doit se trouver dans une différence des classes de stimulations sensorielles observées lors d'une attitude du sujet, conformément aux méthodes mises en œuvre dans la langue scientifique. Nous n'entrerons pas davantage dans les détails de cette

---

l'on érige [$\Box\phi \rightarrow \phi$] en vérité de logique, on perd cette universalité, cette universalité que Ryle met en exergue. Dès lors, pour Quine, la 'logique épistémique' est improprement nommée, elle ressemble davantage à l'axiomatique d'une théorie particulière, telle que l'arithmétique, qu'à la logique des prédicats ou des propositions (...)" (in [Gochet 1978], pp. 146-7).

[299] On en trouve la trace dans [Quine 1968] et [Quine 1975] ; dans les deux cas, la question des mondes possibles a trait au statut des propositions et à la question attenante des conditions d'identité (synonymie) des énoncés.

[300] [Gochet 1978] décrit (p. 194) la naturalisation des états de choses propositionnels : "Les états de choses centrés font intervenir toutes les possibilités de vacuité et d'occupation de tous les points de l'espace-temps. Or il n'est pas nécessaire de prendre en considération ce territoire immense pour définir un état de chose limité comme la présence du chat A sur le toit B. Nous pouvons en négliger de vastes portions. Nous pourrions limiter notre attention à la surface du corps de notre animal autocentré et ne prendre en considération que les possibilités d'activation et d'inactivation de ses terminaisons nerveuses". On retrouve ici la réduction des ensembles maximaux consistants de Carnap aux portions d'espace épistémiques de Hintikka, moins les propositions qui servaient à les décrire.

procédure extensionnelle, sinon pour indiquer que Quine reconnaîtra le semi-échec de ce projet réductionniste.[301]

Pour résumer, l'identification transmondaine (à travers des mondes possibles) peut être contestée pour au moins deux raisons.

D'une part, Quine a reproché à cette solution de ne pas avoir de sens ou de reposer sur une stipulation artificielle, donc infondée.[302] L'identité de deux termes singuliers dans *le* domaine de référence de la logique classique, celui de l'ontologie générale, ne soulève aucune difficulté pour Quine puisqu'elle est censée être garantie par le discours scientifique, tandis que leur identité dans une ontologie locale demande en revanche une explication. Comment expliquer la transaction entre les deux sortes d'ontologie ? Hintikka va répondre à ce problème d'identification, au point de montrer que sa réponse est conforme aux affirmations de Quine lui-même sur la relativité de l'ontologie.[303] Les mondes possibles ne seraient rien d'autre que la mise en expression ou internalisation de la continuité spatio-temporelle des portions d'objets, notamment.

D'autre part, est-il toujours vrai que je sais *qui* est une personne parce que je sais *qu'*elle a telle et telle propriété ? Hintikka reconnaît lui-même que non, après avoir suggéré une réduction des propositions nominales de type 'savoir qui' en propositions complétives de type 'savoir que' : la condition décisive pour identifier un individu et savoir qui il est peut varier selon le contexte, et aucun ensemble de descriptions ne semble en mesure de répondre à ce problème pratique : de ce fait, l'affirmation précédente selon laquelle toute expression *de re* en 'savoir qui' serait réductible à une expression *de dicto* en 'savoir qui' est plus que contestable.[304] Mais en vertu de sa conception de la logique comme modèle explicatif, la

---

[301] "Il y a là une frontière blessante pour tous ceux d'entre nous qui partagent le sentiment que ce qui n'est pas dit en termes de réalités naturelles n'est pas expliqué. Peut-être touchons-nous ici au point décisif dans le problème de l'âme et du corps" ([Quine 1964], p. 197). L'auteur atténuera plus tard ce dualisme par un monisme heuristique, ou théorique : "J'acquiesce à ce que Davidson appelle le monisme anomal, en d'autres termes au physicalisme de principe : il n'y a pas de substance mentale, mais il y a des façons *irréductiblement mentales* de grouper les états et les événements physiques" [Quine 1993], p. 106.

[302] Voir la remarque ci-dessus de Boudot sur les stipulations d'identité, note 296. Une logique modale épistémique a-t-elle pour "mission de résoudre d'authentiques problèmes ontologiques", comme Boudot le prétend, ou doit-elle s'en tenir à formaliser des conditions de vérité générales d'énoncés ? La question se pose parce que la logique modale telle que nous l'envisageons se présente comme une 'logique philosophique'.

[303] "Hintikka ne cessera de dénoncer l'abord strictement formel dans lequel les mondes possibles sont supposés donnés *tout constitués*. Les méthodes alors mises en œuvre garantissent éventuellement qu'une théorie n'est pas contradictoire ; en revanche, loin de contribuer à une authentique élucidation conceptuelle, elles dissimulent souvent les problèmes philosophiques. Son attitude personnelle se caractérisera par le souci de distinguer dans les méthodes formelles ce qui est construction arbitraire et ce qui prolonge en les explicitant de véritables démarches de la pensée qui peuvent trouver leur expression dans le langage courant. De là une articulation originale et harmonieuse des techniques formelles sophistiquées et de l'analyse de la langue ordinaire" ([Boudot X], p. 1218). Un exemple de relation harmonieuse entre philosophie de la connaissance et langage de la connaissance est le parallèle constaté entre les modes de connaissance par accointance et les constructions en termes d'objet direct, opposé au parallèle entre les construction en 'que' et la connaissance par description.

[304] "Dans la pratique, il est fréquemment difficile de dire si un énoncé donné de forme 'a sait qui est b' est vrai ou non. Les critères pour savoir quand on peut dire qui est cet homme-ci ou cet homme-là sont extrêmement variables. Parfois, la connaissance du nom de la personne en question suffit ; parfois, non. Une 'accointance' d'une certaine sorte est souvent requise." ([Hintikka 1962], p. 149 note 9). Le terme 'accointance' ne renvoie pas ici à son sens usuel (avoir des accointances, c'est-à-dire des relations sociales haut placées), mais au sens technique de la connaissance par relation directe de Russell ; la théorie de la connaissance de [Russell 1911]

variabilité des processus d'identification ne constitue pas pour Hintikka une charge accablante contre la quantification modale. Ni la théorie descriptiviste de Frege, ni la théorie des faisceaux de descriptions de Searle et Wittgenstein, ni la théorie de la référence directe de Kripke et Marcus n'apportent la solution clef à l'individuation des objets d'attitude, laquelle solution ne réside pas d'après Hintikka dans le choix linguistique d'une expression particulière mais dans un processus non-linguistique de récognition. De là son allusion à deux modes de connaissance distincts et empruntés à [Russell 1911] : un mode *perspectival*, un mode *descriptif*. Dans le mode perspectival propre aux modalités *de re*, nous pouvons identifier un individu par accointance, c'est-à-dire par une relation directe avec lui dans le champ de perception. Dans le mode descriptif propre aux modalités *de dicto*, l'individu est reconnu non pas physiquement mais à partir d'une description de ses propriétés. Le recours nécessaire à l'un de ces modes d'identification peut varier selon le contexte de discours et le stock de connaissance de l'agent.[305] Le résultat est une distinction formelle entre deux paires de quantificateurs pour chaque mode de connaissance : une paire (∀/∃) pour la connaissance descriptive, une paire (U/E) pour la connaissance perspectivale.[306]

Cela étant, une distinction symbolique ne montre pas que la solution au problème de l'identité transmondaine est de nature logique, puisqu'elle ne le résout pas par le biais d'un calcul. Pire, elle alourdit le symbolisme de la logique modale sans en garantir l'impérieuse nécessité. La solution au problème posé par [Quine 1961a] appartient-elle encore à la logique épistémique, ou renvoie-t-elle plutôt à un strict problème d'épistémologie ? Si tel est le cas, la solution est-elle indiscutable et préserve-t-elle encore le critère de *neutralité* du discours logique? C'est parce que nous en doutons que nous chercherons finalement un moyen plus simple de régler après ce qui suit la question de l'opacité référentielle.

2.3.2 Individuation et identification

La théorie de la référence multiple de Hintikka a l'avantage d'expliquer l'opacité référentielle sans recourir à des entités abstraites telles que les pensées hypostasiées de Frege ; elle nécessite une sémantique relationnelle dans laquelle plusieurs mondes, modèles ou domaines de référence sont admis en plus du monde actuel de référence. Chaque énoncé trouve donc son interprétation au sein d'un ou plusieurs domaines de référence et ce principe d'interprétation multiple, inconcevable pour les partisans de la logique comme langage

---

supposait une accointance à la base de toute forme de connaissance, y compris pour les 'objets' d'énoncés abstraits ou généraux.

[305] Le mode perspectival sera mis en avant s'il s'agit pour l'agent de savoir qui est une personne figurant dans son entourage ou champ visuel ; le mode descriptif sera préféré s'il s'agit de reconnaître un personnage historique lors d'un quizz. Ce dernier mode est plus intellectuel que perceptif et renvoie à une question de compétence linguistique : plus l'agent associe de propriétés à un individu, plus il est en mesure de savoir qui il est par description.

[306] La logique de la perception de [Hintikka 1969] sera décrite par cette seconde paire de quantificateurs. D'autres verbes intentionnels seront traduits dans le même sens : se souvenir, reconnaître, percevoir, etc. (voir [Aho 1994]) renvoient à des connaissances non-propositionnelles, d'où la distinction entre des constructions en termes de complétive (que) ou en termes d'objet direct, selon le mode de connaissance en question. La distinction dans la langue française entre 'savoir' et 'connaître' distingue ces deux modes d'individuation : "Le fait que '$(Ex)K_a(b = x)$' soit vraie pourrait être exprimé de façon quelque peu imprécise en disant que 'a connaît b'. J'ai cependant fini par réaliser qu'il y a certaines différences logiques extrêmement importantes entre le fait que *a* sache qui est b et le fait que a connaisse *b*." [Hintikka 1967], note 6, p. 37

universel, permet de résoudre dans le même temps le problème des ontologies abstraites : les *Sinne* de Frege ne sont autres que

> des habitants de notre monde réel à nous, quoique des résidents d'une région platonique abstraite de celui-là. Puisque de telles entités abstraites sont inacceptables pour Quine, il n'est pas étonnant qu'il ne soit pas satisfait du concept même de signification, sauf à un niveau purement comportemental. Ce qui a été signalé montre que cette insatisfaction est le produit de deux présuppositions différentes, non seulement celle de l'aspect abstrait et du caractère non-empirique des entités de signification comme les *Sinne*, mais aussi, et surtout, celle de la présupposition du monde unique.[307]

Mais la solution des domaines d'interprétation multiples, proposée par les partisans de la logique comme calcul, crée une autre difficulté en retour : quand bien même les individus des mondes possibles seraient des habitants du monde réel, comment s'effectue et que signifie leur transition d'un monde à un autre ?[308]

La difficulté dépasse le seul cadre de la logique modale épistémique quantifiée, puisque le problème de l'identification à travers des mondes possibles remonte au temps de la scolastique médiévale : contre les solutions essentialistes, en vertu desquelles deux individus renvoient au même s'ils partagent la même essence (identité sortale), Leibniz n'admettait pas une telle procédure et supposait au contraire que tout individu ne peut appartenir qu'à un seul monde ; dans le même registre, David Lewis estimera que chaque individu d'un monde a son analogue ou contrepartie la plus ressemblante dans un autre mais que les deux ne sont jamais exactement identiques. Le lien entre ce discours métaphysique et Hintikka est assez ténu, dans la mesure où il s'agit dans les deux cas de deux approches distinctes de l'identité : la métaphysique s'intéresse à l'*individuation* des objets considérés eux-mêmes, c'est-à-dire ce qui constitue leur eccéité (ce qui fait qu'ils sont ce qu'ils sont et pas autre chose) ; l'épistémologie, en revanche, s'intéresse au processus d'identification ou de reconnaissance des individus par un sujet, c'est-à-dire aux facultés par lesquelles celui-ci est capable d'identifier deux entités de mondes possibles comme ne faisant qu'un. La question de l'identité dans les mondes possibles ne se pose donc pas de la même façon selon l'interprétation donnée à l'opérateur modal : le problème est d'ordre ontologique pour les modalités métaphysiques, il est d'ordre ontique pour les modalités épistémiques. Dans le cas de Hintikka, le problème est de savoir dans quel sens on peut dire de deux valeurs de mondes possibles distincts qu'elles sont identifiées comme ne faisant qu'une ou ne renvoient qu'à un seul individu commun.

La distinction entre *ontologie* et *idéologie* peut aider à comprendre ce qui distingue l'approche de Quine et de Hintikka. Le premier reproche au second d'enfler l'ontologie sans raison suffisante, dans la mesure où chaque introduction d'une constante liée exige que l'on reconnaisse l'existence d'une valeur correspondante. En d'autres termes, Quine reproche à la sémantique relationnelle de multiplier les engagements en multipliant les domaines de

---

[307] [Hintikka 1997b], pp. 466-7.

[308] "Mais qu'est-ce que cela *veut dire* qu'une chose dans un monde possible est identique à une chose dans un autre monde possible ? Ou si nous ne savons pas ce que cela veut dire, comment pouvons-nous *décider* si un individu d'un monde possible est identique à un individu d'un autre ?" ([Chisholm 1963], pp. 793-4 ; nous soulignons).

référence, selon [Quine 1961a], et force le sémanticien des mondes possibles à admettre l'existence d'être irréels tels que le gros homme dans l'embrasure de la porte. Mais la théorie des lignes de mondes de Hintikka montre qu'aucun engagement ontologique de ce genre n'accompagne *a priori* la sémantique des mondes possibles. Il n'y a d'engagement ontologique véritable que lorsque la valeur d'une variable est la même à travers tous les mondes possibles pour un agent, autrement dit, lorsqu'il a parfaitement identifié son objet (de pensée), et lorsque cette valeur appartient au domaine du monde *actuel* : on peut identifier avec succès un personnage fictif tel que le père Noël sans lui prêter jamais existence. La quantification dans un monde ne suffit donc pas à déclencher un engagement existentiel : elle revient simplement à imaginer des ontologies variables, conformément à la thèse de la relativité ontologique de Quine lui-même. Encore une fois, le sens que Hintikka attribue à la notion d'idéologie diffère de celle de [Quine 1951a] dans la mesure où les deux logiciens ne partagent pas les mêmes présupposés sémantiques ; tandis que, pour Quine, l'idéologie désigne l'organisation des objets selon leurs relations mutuelles au sein d'un schème conceptuel (le tableau global que l'on se fait du monde), pour Hintikka elle désigne les outils conceptuels au moyen desquels le sujet se constitue ou créée ces objets :

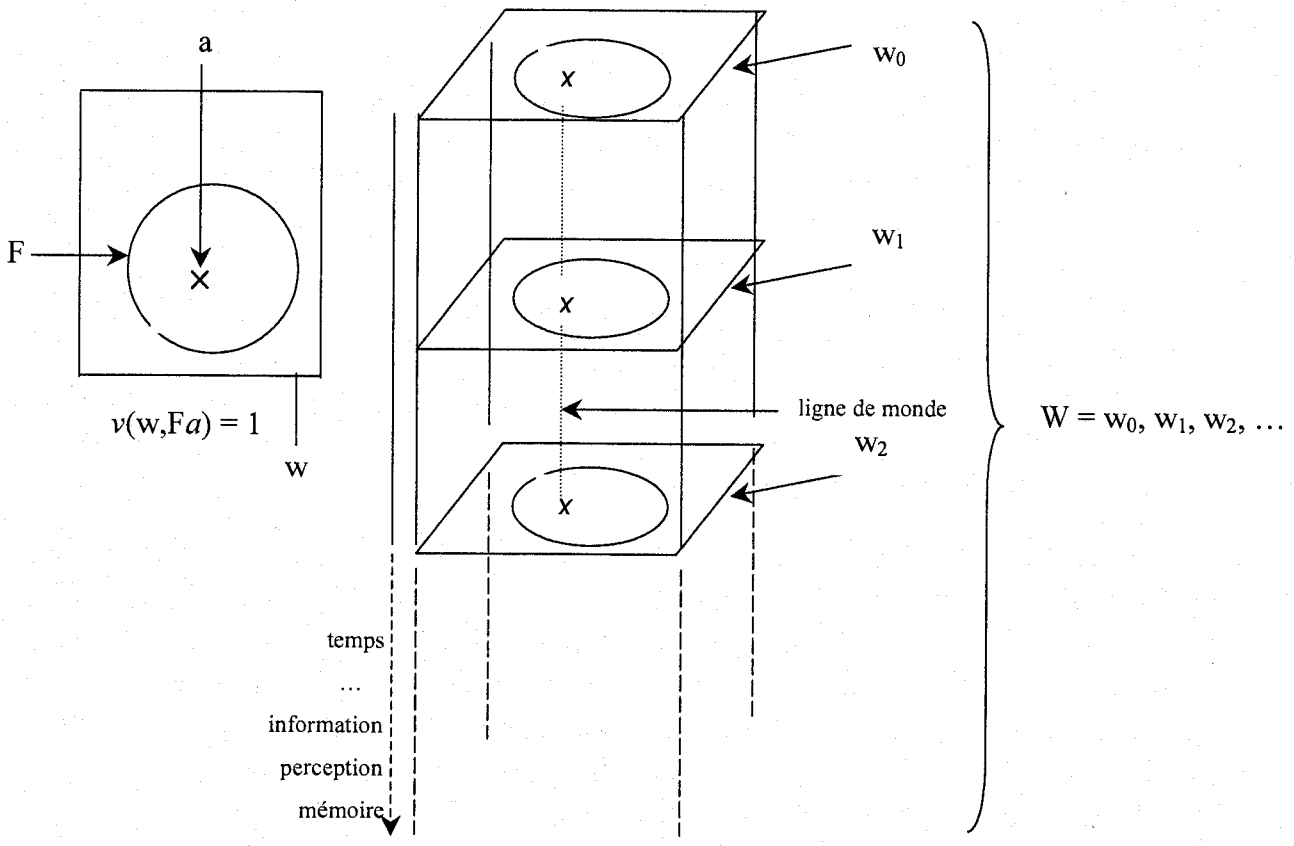
> Il nous faut distinguer (1) ce à quoi nous sommes engagés –au sens où nous croyons que cela existe, dans le monde réel ou dans un autre monde possible – et (2) ce à quoi nous sommes engagés – au titre de moyen conceptuel de traiter du monde, au titre d'élément de notre système conceptuel. Notre ontologie est définie par (1) – et notre 'idéologie' par (2). Je pense pouvoir dire que les mondes dans lesquels il nous faut quantifier appartiennent à notre idéologie, non à notre ontologie.[309]

Nous n'avons pas affaire à d'*autres* entités dans les mondes possibles qu'à celles de notre monde actuel, celui que Quine décrit en termes de schème conceptuel ; ce sont au contraire ces entités possibles qui constituent point par point ou monde par monde les entités réelles de notre schème quotidien. Pour Hintikka, c'est donc l'idéologie qui prime sur l'ontologie en participant à la constitution des objets, justifiant qui plus est l'idée quinienne d'ontologie relative à une sélection d'objets de pensée. Cette notion d'idéologie, si elle suggère un processus intellectuel de constitution des objets du monde, n'est pas purement conceptuelle et peut renvoyer à d'autres formes d'activité que la connaissance descriptive par le biais de concepts, comme nous le verrons par la suite dans l'examen des différents modes d'individuation.

Formellement, la distinction entre ontologie et idéologie explique le problème de l'assignation de référence en termes de *ligne de monde*, de *clivage* ou de *fusion* d'individus dans des mondes possibles. Le propre de l'idéologie est de fixer un domaine de référence stable chez le sujet pensant : une ligne de monde est parcourue d'un bout à l'autre par une fonction individuante, laquelle associe à une constante d'individu la même valeur dans chaque monde possible. En cas de clivage, la fonction individuante 'perd le fil' de la ligne et la variable son unicité à travers les mondes, soit parce que la fonction lui associe une autre valeur dans un monde soit parce qu'elle ne lui en associe aucun (la fonction individuante n'est

[309] [Hintikka 1969b], p. 45.

donc plus totale et devient partielle)[310]. Pour illustrer un cas de transparence référentielle où la référence du terme d'individu apparaît clairement à l'agent, on peut utiliser une représentation graphique que Vuillemin avait proposée pour le processus de prédication.[311] L'analogie est la suivante : de même qu'une prédication substantielle associe une propriété constante à un terme singulier à travers le temps, l'identification d'un objet par un agent peut être considérée comme la persistance des propriétés de cet objet à travers le temps. Ces propriétés seront des prédications ou descriptions définies, dans le cas d'une connaissance descriptive ; mais elles peuvent être d'une nature différente selon le type de modalité en vue et la nature des mondes possibles. La métaphore des lignes de monde donne le principe d'identité modale suivant :



Pour garantir la persistance d'une valeur à travers tous les mondes, Hintikka invoque sa thèse de l'*intentionnalité comme intensionalité* : toute construction en termes de verbes intentionnels trouve son explication sémantique en termes d'intensions ou de mondes possibles.[312] Une précision est à faire toutefois sur la catégorie des verbes intentionnels : ceux-ci sont indiqués principalement par les constructions complétives en 'que' (croire que,

[310] La partialité d'une fonction individuante peut se manifester lorsque l'on dit par exemple que l'on n'a 'aucune idée' de ce que signifie un nom commun : aucun objet particulier n'est assigné à ce nom. En ce sens, la notion d'idée renvoie à l'assignation d'une valeur dans un monde possible.
[311] J. Vuillemin : "Essai sur les formes fondamentales de la prédication", XXX (1984), p. 16.
[312] Le lien établi entre intension et intention est expliqué historiquement par [Lemmon 1977] : "Au dix-neuvième siècle, Sir William Hamilton remplaça 'compréhension' par 'intension', en épelant fautivement le mot avec un 's' par analogie avec 'extension'. Depuis lors, le terme 'intention' a pris une direction, via Brentano et Chisholm, et le mot 'intensionalité' en a pris une autre via Carnap et Quine. (...) il ne fait pas de doute que [ce dernier terme] est intimement lié au premier" (p. 8).

penser que, vouloir que, savoir que), suggérant que le sujet vise un objet de pensée (une 'proposition', ou 'Objektiv') ; cette approche phénoménologique des verbes intentionnels va inspirer l'interprétation que donne Hintikka de l'identification dans les mondes possibles.

A la différence de ces verbes suivis d'une complétive, les constructions verbales en objet direct n'ont pas de complétive et suggèrent que le sujet ne vise pas un objet de pensée abstrait mais un objet concret, dont l'existence ne dépend pas de la seule conscience qu'il en a dans un monde possible (croire (à), voir, percevoir, sentir, toucher ou connaître font partie de ces verbes).[313] L'ambiguïté concernant le double sens de l'opérateur 'K' s'expliquerait ainsi par les deux modes d'individuation distinct qui le caractérisent, selon que cet opérateur épistémique s'exprime dans un énoncé en termes de 'savoir que' ou de 'connaître' :

> Il y a plusieurs autres phénomènes importants que l'analyse en mondes possibles peut révéler mais qui sont apparemment impossibles à saisir directement à partir des phénomènes de la conscience. Un cas du genre est le fait que nous transquantifiions (dans une pratique conceptuelle courante, codifiée dans la sémantique des langues naturelles) entre différentes situations possibles de deux manières totalement différentes. Leur différence est parallèle à la différence entre la sémantique de la construction d'un verbe épistémique + une clause interrogative et celle de la construction d'un verbe épistémique + un objet direct. L'impossibilité de recouvrir cette distinction importante du point de vue phénoménologique est illustrée par l'incapacité à la faire chez pratiquement tous les analystes antérieurs, ainsi que par la difficulté à l'expliquer chez les philosophes, en dépit du fait que, du point de vue des mondes possibles, elle constitue une caractéristique assez évidente de la situation conceptuelle.[314]

Les conditions dans lesquelles deux termes d'individus renverraient au même objet dépendraient ainsi de lois relatives à la structure de la conscience, par analogie avec les études phénoménologiques de l'intentionnalité entreprises par Brentano puis poursuivies par Chisholm. A ce titre, l'objection de Quine contre des individuations purement arbitraires en logique modale s'évanouit ; les conditions d'identification sont contingentes mais soumises à des lois susceptibles d'être décrites, et le choix de l'analyse sémantique s'appuie par ailleurs sur des indices de construction verbale (complétive : identification descriptive ; objet direct : identification perspectivale). Le lien entre intensionalité, intentionnalité et phénoménologie a motivé notamment les études de Hintikka sur Husserl, puisque leur domaine d'étude est

---

[313] La connaissance par perception est liée à la question de l'accointance et renvoie ainsi à une construction grammaticale non complétive, ce qui exige une analyse particulière des énoncés de connaissance. Sur l'analyse de ces énoncés, voir notamment [Hintikka 1989a] et [Niiniluoto 1979].

[314] [Hintikka 1981a], p. 118 note 15. Cette distinction élimine une objection portée par [Morick 1975] à la thèse d'intentionnalité comme intensionalité, puisque l'auteur ne semble pas prendre acte de la différence entre les deux types de constructions verbales : Morick signale que l'énoncé 'le Roi a cherché le Saint Graal' contient un verbe intentionnel mais que la substitution s'y applique *salve veritate* ; un énoncé de nature intentionnelle pourrait ainsi être vérifonctionnel, c'est-à-dire non-intensionnel. Mais la thèse de Hintikka porte plutôt sur les constructions en 'que' : croire que, penser que, se rappeler que, etc. Quant aux constructions en objet direct (rechercher, croire (à), connaître), les conditions d'individuation de leurs objets ne sont pas tout à fait les mêmes : un athée peut dire *salva veritate* de quelqu'un qui croit *au* bon dieu qu'il croit à ce qui n'existe pas, donc de tels verbes ne sont intentionnels qu'en apparence, au sens où Hintikka l'entend. Ce problème de délimitation renvoie plus profondément au problème de la nature intentionnelle ou non des verbes tels que 'percevoir', 'sentir' : sentons-nous des objets directement, c'est-à-dire sommes-nous réellement en accointance avec des objets qui existent indépendamment de nous, où la sensation produite par ces objets est-elle due à notre structure intentionnelle ?

analogue mais abordé par des moyens d'analyse distincts : la phénoménologie pour Husserl, la sémantique formelle pour Hintikka.[315]

A l'analyse des modalités épistémiques et intentionnelles en général, [Hintikka 1982] oppose le cas des modalités logiques, incapables selon lui de recevoir des garanties d'individuation clairement définies pour ses formules. Cela confirme le caractère non arbitraire et contextuel des modes d'individuation pour les contextes épistémiques, dont le comportement des fonctions individuantes ne repose pas sur un fiat :

> Nous devons reconnaître par ailleurs que les 'lignes de monde' de l'identification croisée (les lignes notionnelles connectant chacune les incarnations ou rôles d'un seul et même individu dans différents mondes possibles) ne sont pas déterminées par Dieu, la Nature ou la Logique, qu'elles sont tracées en principe par *nous-mêmes*. Elles ne sont pas tracées de façon arbitraire, c'est vrai, mais au moyen de diverses considérations objectives telles que la continuité dans l'espace et le temps, la continuité mémorielle et la localisation dans l'espace visuel d'une personne. Par ailleurs, elles peuvent être retracées objectivement une fois qu'elles ont été tracées, indépendamment des faits et gestes d'un utilisateur du langage individuel.[316]

Au total, le problème de l'opacité référentielle trouve sa solution dans une sémantique des mondes possibles (relationnelle), et ceux-ci sont analysables par analogie avec la théorie de l'intentionnalité des phénoménologues : les objets des mondes sont des objets de pensée produit par la structure de notre conscience, pour autant que Hintikka semble le prétendre en réponse à ses détracteurs.[317] Les problèmes posés par les fonctions individuantes en logique épistémique correspondent ainsi à un problème épistémologique d'*identification* ou de constitution des objets, et non à un problème métaphysique d'*individuation* ou de nature des objets. Le discours spéculatif de l'eccéité ou de l'essentialisme n'a donc pas lieu d'être dans une sémantique des attitudes épistémiques.

Or bien que l'analyse des modalités épistémiques trouve un sens dans la phénoménologie des objets de pensée, et bien que le discours des mondes possibles n'entraîne pas un engagement ontologique injustifié, celui-ci semble buter toutefois sur un problème évoqué par [Chisholm 1967] et [Purtill 1970] : sous quelles conditions suffisantes (par quelle idéologie

---

[315] "Ma déclaration que le champ de la totalité des mondes possibles ne forme pas une totalité bien développée est éminemment en phase avec la déclaration de Husserl citée par [Mohanty 1981], selon laquelle 'l'univers des libres possibilités en général est un champ de discontinuité : il manque d'unité de contexte. Nous devons restreindre plutôt le domaine des mondes possibles à des possibilités adéquatement 'motivées', ce qui, encore une fois, est conforme à Husserl, je crois" ([Hintikka 1981a], pp. 115-6)

[316] [Hintikka 1986], pp. 209-10 ; nous soulignons. Ce passage semble constituer une réplique aussi bien à [Quine 1961a] qu'à [Quine 1976], concernant la relativité arbitraire de l'identité dans des mondes possibles. Rappelons toutefois que, à l'époque où 'On what there is' fut écrit, les mondes possibles étaient en large partie ceux de la nécessité logique, dont Hintikka vient de contester la possibilité d'individuation des objets. La réplique vaut en revanche pour [Quine 1976], puisque celui-ci déduisait de l'aspect 'invariablement contextuel' des conditions de connaissance d'un individu que leur analyse ne saurait être de nature logique.

[317] L'analyse de Hintikka est considérée tantôt à celle de Husserl, tantôt à celle de Kant, témoin le jeu de mots dans le titre de [Smith 1983] et les remarques suivantes : "Il n'y a pas de chose en soi qu'on puisse décrire et individualiser sans se fonder sur un cadre conceptuel particulier qui s'explicite dans la délimitation d'un ensemble de mondes possibles et dans le choix des fonctions individuantes. Hintikka en appelle à un néo-kantisme, à vrai dire teinté de phénoménologie ; il n'y a d'objets que de pensée ; seuls les noèmes sont des objets" ([Boudot X], p. 1357). La dernière section de [Hintikka 1969a] annonce clairement son orientation 'Vers un néo-kantisme sémantique', en effet (p. 59). La préfacière de [Hintikka 1989] dit de même, mais [Nef 1991] lui fait remarquer que, si tel est le cas, Hintikka ne défend pas en vérité un nominalisme mais un conceptualisme modal (voir section 1.3.3).

ou quels outils conceptuels) l'esprit (la noèse, diraient les phénoménologues) est-il en mesure d'identifier un objet et de le distinguer de tous les autres ? Le parcours de la ligne de monde donne un aperçu formel du processus global, mais certains cas de clivage de la ligne semblent inexplicables et, donc, ne pas trouver de sens dans une analyse en termes de mondes possibles. Pour expliquer la continuité des lignes de mondes, c'est-à-dire l'identité persistante des objets dans notre cadre conceptuel, l'hypothèse de la continuité spatio-temporelle est souvent évoquée par Hintikka et confortée, par ailleurs, par l'opposant farouche des modalités qu'est Quine : deux segments temporels appartiennent à la carrière (parcourent la ligne de monde) d'un seul et même objet s'ils sont reliés par une trajectoire spatio-temporellement continue. Ce critère est évoqué pour expliquer la ré-identification possible des personnages d'un environnement, notamment, et donne corps à la théorie des lignes de mondes. Mais un tel critère ne répond pas à tous les cas d'identification dans un contexte. Chisholm et Purtill citent des exemples de paradoxes sorites tels que l'argument du chauve ou l'arche de Noé. Supposons que deux bateaux voient leurs planches être progressivement échangées l'une avec l'autre ; à quel moment le sujet pensant ne reconnaîtra-il plus l'identité de départ des deux bateaux et prendra l'un pour l'autre ? Par ailleurs, si le critère de continuité spatio-temporelle prévaut, comment expliquer que chaque bateau demeure le même pendant un temps indéterminé malgré les trajectoires spatio-temporellement distinctes de certains de ses segments (les planches) ?[318]

Il n'est pas sûr que l'analyse des modes d'individuation par accointance trouve une réponse à ce problème du sorite, analogue au cas du tas de sable : dans tous les cas de prédication vague, le sujet perçoit l'identité de l'objet à chaque instant présumé (dans chaque monde possible compatible avec ses perceptions) mais est incapable de déterminer à quel instant (dans quel monde) ce qu'il connaît d'abord comme des grains de sable se transforme en tas de sable. Une telle connaissance ne passe apparemment pas par des critères descriptifs ou conceptuels, mais elle peut peut-être trouver une réponse de nature psychosensorielle. Quant à l'analyse sémantique de ces situations ambiguës, elle paraît tout aussi difficile puisqu'elle exige de trouver des conditions de vérité pour des énoncés tels que 'Je perçois des grains de sable' (équivalant à 'je sais que je vois des grains de sable') et 'Je perçois un tas de sable' (équivalant à 'je sais que je vois un tas de sable').[319]

L'avantage de cette étude des modes d'identification des objets d'attitude, malgré les failles qui viennent d'être évoquées, est qu'elle considère les mondes possibles des sortes de sous-ensembles du monde actuel : le monde est l'image résultant de nos outils conceptuels par lesquels nous l'élaborons, que ce soit par des facultés innées ou des voies conceptuelles

---

[318] Cette objection est considérée par [Ferret 1998] comme une raison de rejeter le critère d'identification par continuité spatio-temporelle. Il existe d'autres critères tels que l'unicité compositionnelle, pour garantir l'unicité d'un objet, mais nous n'irons pas plus loin dans ce débat.

[319] Une solution a été proposée par [Egré 2004] en termes de mondes possibles, dans la lignée des sémantiques de voisinage ou de situation : une connaissance perceptive K^p se distingue d'une connaissance réflexive K^r par des conditions de vérité moins drastiques, qui n'exigent pas que l'objet de connaissance appartienne à *tous* les mondes épistémiquement possibles mais à des paires de mondes 'proches' les unes des autres. Cette redéfinition de K, qui supprime par ailleurs la dualité entre $\Box$ et $\forall$, a pour but d'apporter une sémantique adaptée aux paradoxes des sorites et d'identifier les objets perçus dans des espaces de perception plus restreints : les nuances de couleurs, les grains de sable et toutes les prédications vagues sont concernées. Mais le problème demeure en amont : sur quelle explication baser cet appariement des mondes et, si aucune explication n'apparaît, ne doit-on pas en conclure que la modélisation logico-mathématique demeure *ad hoc*, arbitraire ou injustifiée ?

acquises (culturelles). La parenté entre le sens des mondes selon Hintikka et la noétique de Husserl est clairement revendiquée par le logicien, par ailleurs. L'inconvénient est qu'il semble substituer à l'engagement ontologique un engagement théorique, ou philosophique : faut-il s'engager dans une explication phénoménologique pour donner un véritable sens aux modalités épistémiques relationnelle, et la valeur du modèle explicatif ci-dessus dépend-elle de présupposés phénoménologiques, notamment ? Si un tel engagement est demandé pour mettre en évidence l'éclairage philosophique apporté par la logique épistémique, que pensera un matérialiste ou un naturaliste tel que Quine de ce genre de 'logique philosophique' et son traitement de l'opacité référentielle ? Dans la section qui suit, nous montrerons qu'un changement d'interprétation des opérateurs épistémiques peut changer les conditions de vérité des énoncés de connaissance et régler le problème de l'opacité référentielle à moindre frais (spéculatif).

Aussi instructives soient-elles, les analyses intentionnelles de Hintikka semblent plus tenir de l'exégèse spéculative que de l'analyse logique. Bien qu'elle respecte le critère de neutralité par rapport au sujet (les fonctions individuantes portent sur l'opérateur modal, uniquement, et leur mode d'opération ne varie pas selon la matière de l'énoncé), la sémantique relationnelle pose un problème concernant sa légitimité : a-t-elle besoin de l'explication phénoménologique pour être préférée à la méthode des matrices, ou est-ce au contraire la phénoménologie qui a besoin d'être confirmée par le discours des mondes possibles ?[320]

Quel que soit le rôle théorique de la section présente, nous allons proposer une autre lecture de l'opacité référentielle en général ; une lecture qui maintient (LL), contrairement à Hintikka, donc maintient la vérifonctionnalité des modalités épistémiques mais à condition de restreindre les attributions de croyance et de connaissance à un individu. C'est de cette attribution opaque que provient la difficulté de la logique modale épistémique quantifiée : comment déterminer lorsque deux attitudes propositionnelles en général sont les mêmes et, par là, ont les mêmes conditions de vérité ? Le critère d'identité dans tous les mondes possibles sera maintenu dans ce qui suit, mais l'interprétation des mondes sera changée : en termes illocutoires d'assertion, ainsi que dans la section précédente sur les modalités itérées. Pourquoi ce changement de perspective, et les conséquences ne sont-elles pas trop radicales pour une logique des modalités épistémiques ?

2.3.3 Une autre solution : attitude et assertion

Notre issue au problème de l'opacité référentielle est formellement analogue à la solution précédente de Hintikka, mais elle repose sur une autre 'logique philosophique', c'est-à-dire sur une autre interprétation des modalités : $\Box F\mathit{a}$ implique $\Box F\mathit{b}$ si et seulement $\Box(a = b)$, certes, mais notre approche illocutoire précédemment suggérée pour les modalités épistémiques redéfinit entre-temps le sens de $\Box$. La modalité forte ne signifie plus simplement la connaissance mais l'acte explicite d'énonciation, ou *assertion* : si un locuteur

[320] On trouve dans [Hintikka 1989] : "Les intentions de l'intentionnalité" une discussion des modes de conscience selon Husserl en termes de sémantique des mondes possibles, avec quelques corrections apportées par Hintikka sur les arguments noétiques. Est-ce à dire que l'analyse logique corrige et améliore le discours philosophique, ou bien l'inverse ? Ne faut-il pas plutôt séparer les deux et considérer l'analogie comme une combinaison de la logique et de la philosophie ? C'est le statut de la notion de 'logique philosophique' qui est en question, ici.

sait ou croit que F*a*, alors on est en droit d'en inférer qu'il sait ou croit que F*b* si et seulement si il a annoncé au préalable '*a* et *b* sont identiques'. Cette condition est nécessaire *et suffisante* selon nous pour garantir le sens d'un énoncé de connaissance ou de croyance. Avec [Quine 1961a], nous avons accordé que la solution formelle à l'opacité référentielle devait s'accompagner d'une explication sur le principe d'identification et le comportement des fonctions individuantes. Or celles-ci nous ont emmené vers des explications de type phénoménologique qui s'accordent avec le philosophe Hintikka, mais ne devraient pas s'imposer à n'importe quel logicien des modalités épistémiques.[321] Le but de ce qui suit est de trouver des conditions de vérité nécessaires et minimales pour les énoncés épistémiques, telles que l'on puisse spécifier les cas dans lesquels quelles (BF_e) et (BF_i) sont vérifiées.

L'explication est essentiellement la même que pour la thèse de l'introspection positive, défendue trompeusement en termes psychologiques d'introspection et de conscience de soi : de même que le 'sophisme du psychologue' (section 2.2.4.2) consistait à traiter les états mentaux comme des faits objectifs, le sophisme de la transparence référentielle exprimé par (LL□) réside dans le passage illégitime d'un acte personnel d'énonciation à un énoncé impersonnel. Autrement dit, le présupposé de la transparence référentielle des énoncés consiste à traiter nos propres énoncés comme des informations compréhensibles par tous et à attribuer aux locuteurs une compétence linguistique maximale ; or c'est cette attribution de nos propres croyances aux autres sur la base de nos propres compétences linguistiques qui conduit à l'opacité référentielle : les formes complétives en 'que' suggèrent que l'attitude épistémique de l'agent porte sur une 'proposition', d'après l'explication courante ; et puisque, d'après Frege, ces propositions sont des entités objectives (des *Sätze an Sich*) que chaque locuteur est capable de saisir par la pensée, l'objectivation des actes d'énonciation par le biais d'énoncés publics ou propositions conduit à traiter l'assertion 'p' comme une expression indexicale de la proposition *que* p. Or une croyance n'est jamais exprimée dans les faits par des entités objectives de ce genre. En adaptant son exemple au nôtre, on constate que [Zemach 1971] met également l'accent sur cette interprétation illégitime des croyances des autres dans nos propres termes, fait comparable au sophisme du psychologue où l'on attribue nos propres actes de conscience aux autres :

> Nous devons nous souvenir que [mon père est un vendeur d'armes de destruction massive] n'est pas *asserté* par [Paul]. Il est asserté par *nous*, et le terme ['vendeur d'armes de destruction massive'] utilisé dans celui-ci est un terme référentiel dans *notre* langage, faisant référence à

[321] L'explication des fonctions individuantes concerne davantage l'épistémologie que la logique épistémique, comme le déclarent plusieurs auteurs : "La question de savoir si quelque chose est réellement tel qu'il semble être ou n'est pas réellement tel qu'il semble être est une question qui mérite sans aucun doute que l'on médite à son sujet. Mais celle-ci appartient à l'épistémologie plutôt qu'à la logique épistémique" ([Lenzen 1978], p. 52) ; "Les premiers écrits de Russell sur la 'connaissance par accointance et la connaissance par description' devraient être compris à la lumière de ce contexte. Ce n'est pas critiquer Hintikka que dire qu'il ne *résout* pas le problème, si c'est là un problème véritable, car son ouvrage porte sur la *logique épistémique* et non sur l'épistémologie." ([Chisholm 1963], pp. 789-90). La distinction entre les champs de compétence de la logique épistémique et de l'épistémologie est reconnue par [Hintikka 1967] lui-même : "La description précise de ce sur quoi ces méthodes reposent n'est pas non plus un problème pour un logicien, car on s'aperçoit que ces critères d'identification croisée reposent sur tous les types de faits contingents (contingents, c'est-à-dire du point de vue logique). (Quelle est la base de l'identité personnelle, par exemple ? Il est évident que cela est un problème philosophique profond ; à l'évidence, ce n'en est pas un que nous pouvons ou devons aborder à ce stade de notre étude)". ([Hintikka 1967], p. 42).

> tout ce qui peut *en fait* être [un vendeur d'armes de destruction massive], autrement dit à [son père]. En utilisant le terme ['vendeur d'armes de destruction massive'], nous n'impliquons pas que [Paul] comprend lui-même ce terme ou qu'il serait disposé à l'utiliser dans le but de faire référence à [son père] (...) L'utilisation du terme est seulement *notre* façon de faire référence à cette entité.[322]

Si l'on rejette à la fois l'idée d'objectivation des énonciations en énoncés et la thèse de la saisie de propositions communes comme entités objectives, le problème de l'opacité référentielle peut être évité ou réduit dans les mêmes proportions que la thèse de la transparence de la pensée : un problème apparaît parce que les termes de ce problème sont mal posés.

En gros, la transition entre un acte de discours et un discours indirect rapportant les pensées d'un autre n'est légitime que si le rapporteur s'est assuré que le sujet croyant a publiquement prononcé les mots rapportés. Avec cette condition d'attribution des croyances sous assertion explicite, le caractère non-vérifonctionnel des énoncés épistémiques s'évanouit. Pour reprendre l'exemple du début de section, si Paul asserte ou déclare ouvertement les mots suivants : 'mon père est un philanthrope', l'expression de son énonciation au discours indirect[323] permet de dire légitimement de Paul qu'il croit que son père est un philanthrope ; mais s'il n'a pas affirmé lui-même 'mon père est un vendeur d'armes de destruction massive', le rapporteur ne sera pas en droit de pouvoir dire de Paul qu'il croit qu'un vendeur d'armes de destruction massive est un philanthrope. Notre conclusion formelle est donc la même que celle de Hintikka, mais elle est décrite en termes illocutoires d'assertions et non en termes locutoires de mondes épistémiquement possibles ; de plus, elle interdit d'attribuer des attitudes épistémiques à une personne tierce sans l'appui d'une déclaration manifeste.

En d'autres termes, toute attitude épistémique se définit sur la base de la notion primitive d'assertion, et la distinction entre connaissance et croyance repose encore une fois sur le rapport de ces concepts à la notion de vérité : soit le locuteur a l'intention d'affirmer la vérité d'un énoncé à ses interlocuteurs, et il exprime son assertion par une déclaration de connaissance ; soit il n'est pas certain mais suppose simplement la vérité de son énoncé, et il exprime son sentiment par une déclaration de croyance. Nous verrons dans la prochaine section qu'il existe un moyen de distinguer ces deux concepts fondamentaux de la logique épistémique : la connaissance exprime un degré de *force illocutoire* supérieur à la croyance, mais les critères distinctifs que propose la logique illocutoire de [Searle & Vanderveken 1985] montrent que la relation entre les déclarations de connaissance et de croyance est plus complexe qu'il n'y paraît ici : leur but illocutoire peut être considéré comme de nature différente (dire la vérité pour le premier, décrire un état mental pour le second), auquel cas leur différence ne sera pas simplement une différence de degré entre deux forces illocutoires de même nature. Nous illustrerons cette relation complexe par la suite.

---

[322] [Zemach 1971], p. 805.

[323] Le principe de décitation justifie la comparaison faite par Frege et Quine entre contextes de croyance et contextes de discours indirect. Il explique aussi que la référence dans un discours indirect puisse être 'inhabituelle', lorsqu'elle évoque une dénotation que le sujet croyant ne comprend pas ou n'est pas capable d'identifier. Mais cela n'implique pas une absence de référence dans le discours indirect, contrairement à l'interprétation de [Frege 1892].

Si l'on interprète la modalité □ par la notion d'assertion, le critère suggéré par Hintikka pour autoriser la substitution est maintenue : la condition □(a = b) pour passer de □F*a* à □F*b* entraîne toujours un rejet de (LL) et des autres règles d'extensionalité, mais le contexte épistémique est rebaptisé sous l'appellation de contexte *assertorique*. La substitution des termes co-référentiels dans un contexte assertorique n'est permise que dans la mesure où les attitudes en question portent sur des énonciations de forme logique ⊙ϕ et non plus sur des énoncés de forme ϕ, où ⊙ symbolise un marqueur de force illocutoire tel que 'savoir' et 'croire' mais aussi 'supposer', 'certifier', 'être certain', 'conjecturer', etc. L'hypothèse auxiliaire exigée par Hintikka pour garantir la transparence référentielle peut être exprimée ainsi : pour tout locuteur A, si A asserte '*a* est F' *et* si A asserte également '*a* est identique à *b*', alors la substitution des termes co-référentiels *a* et *b* est permise et l'on peut en inférer que A est disposé à asserter '*b* est F'. Si l'on ajoute la condition de sincérité, en vertu de laquelle une personne croit ce qu'elle dit, le fait que A dise '*b* est F' autorise les autres à inférer de ses déclarations que A croit que *b* est F. C'est sur cette base pragmatique que l'on peut attribuer des croyances à une personne et reformuler le contenu de ses pensées. Elle présente au moins deux avantages par rapport à une logique des modalités épistémiques. D'une part, l'attribution d'attitudes propositionnelles à une personne tierce peut éviter l'opacité référentielle, mais à condition qu'elle s'appuie sur les énonciations personnelles de cette personne et non sur des compte-rendus impersonnels de ses croyances. D'autre part, le problème de l'omniscience logique est évité également puisque l'on ne peut attribuer d'attitude à un sujet locuteur que si ce dernier a clairement énoncé le contenu de cette attitude : même si ψ est la conséquence logique de ϕ, l'énonciation par le sujet A de l'énoncé ϕ indique qu'il croit que ϕ est vrai mais ne suffit pas pour lui attribuer l'autre croyance que ψ est vrai. La procédure, aussi triviale soit-elle, permet de restreindre les attributions de croyances par le biais d'un critère finitiste de croyance comme assertion : il se peut que le locuteur A soit disposé à asserter la conclusion que ψ après avoir asserté sa prémisse ϕ, mais rien ne le garantit sans une énonciation explicite de ψ à l'appui. Une compétence linguistique et logique minimale accordée à chaque locuteur devrait nous permettre d'étendre le critère d'assertion à un critère plus souple d'assertabilité : A devrait asserter ψ s'il a asserté ϕ. Mais le résultat ne sera pas garanti et se heurtera de nouveau à des cas d'incompétence logique (non-omniscience) ou linguistique (opacité référentielle). Nous optons ainsi en faveur d'une interprétation assertorique des modalités, et posons le critère d'énonciation comme une condition nécessaire et suffisante à l'attribution d'une croyance : le rôle d'une énonciation de type déclaratif[324] est de communiquer une assertion, c'est-à-dire une croyance à un certain état de choses par le biais d'un énoncé déclaratif ; cette condition suffit à attribuer une croyance si la condition de sincérité du discours est respectée.

On pourrait objecter à cette analyse des modalités épistémiques en termes minimaux d'assertions ou actes d'énonciation correspondants qu'elle est trop restrictive, qu'elle interdit l'attribution d'un grand nombre de croyances implicites à autrui tant que ce dernier n'a pas

[324] Nous précisons 'de type déclaratif', parce que les énoncés de type interrogatif ou exclamatif n'expriment pas des assertions : ils expriment des demandes ou des souhaits, respectivement. Or comme la logique épistémique fait porter ses opérateurs modaux sur des énoncés de type déclaratif, il va de soi qu'une formule telle que Kϕ ou Bϕ représente une attitude portée sur un énoncé *déclaratif*.

explicitement énoncé la pensée qu'on lui rapporte. Mais notre démarche est la même que dans le cadre de l'analyse de l'axiome **4** : la condition d'introspection a été considérée par Hintikka comme reposant sur une condition préalable d'assertion manifeste ; il en va de même ici, où les règles extensionnelles de la logique classique ne posent plus problème à condition qu'elles portent sur des unités minimales de signification plus fines (plus sensibles au contexte de discours), c'est-à-dire les assertions ou énonciations, de préférence à des énoncés ou propositions dont le rôle dans la pratique du langage est aussi vide que l'hypothèse des 'faits mentaux' objectifs du psychologue.[325] Si cette condition de signification minimale est respectée, elle suppose donc une unité de temps et d'action pour les attitudes propositionnelles : chaque assertion est effectuée à un moment particulier et par un locuteur particulier, donc la signification de ses mots dépend des deux critères indexicaux que sont l'*intention*[326] communicative du locuteur et le moment de l'énonciation.[327] Mais cette indexicalité de la signification n'entraîne pas la conclusion absurde selon laquelle chaque acte d'énoncé constituerait un acte idiosyncrasique unique, doté de ses propres règles de signification et basé sur une sorte de langage privé : le sens des mots assertés dépend d'un apprentissage préalable du langage tel que [Quine 1974] l'a reconstruit empiriquement (d'abord des énoncés holophrastiques, puis des énoncés observationnels et, enfin, des énoncés généraux ou permanents), mais ils doivent avoir été assimilés par le locuteur avant d'être considérés comme crus ou connus de lui (voir section 3.2.1). Si cette condition de pensée comme assertabilité garantie d'un énoncé est respectée, le problème de l'opacité référentielle disparaît et les règles extensionnelles peuvent être restaurées.

Notre traitement illocutoire des modalités illocutoires est donc comparable à l'analyse proposée par [Zemach 1971], qui consiste à montrer que le maintien de (LL) dans un contexte de croyance ne produit aucune absurdité tant que ce contexte est défini en termes explicites d'assertion :

> Compte tenu de tout ceci, je crois pouvoir offrir une solution à l'énigme à un prix modique. Je conserverai la version *la plus forte* du principe de Leibniz –la substituabilité *salva veritate* des termes co-référentiels dans tous les contextes, je nierai le fait qu'il y ait des contextes 'opaques',

[325] Notre réduction des énoncés à des énonciations supprime ainsi le recours aux entités platoniciennes de Frege : ce ne sont pas des propositions objectives issues d'un troisième monde que nous saisissons (les *Gedanken*, ou pensées), lorsque nous produisons un énoncé. L'énoncé effectué n'est pas simplement l''habit sensible' ou expression d'une soi-disant proposition objective et indépendante: il doit être prononcé pour *produire* la pensée correspondante.

[326] L'intention communicative ne renvoie pas à la théorie de l'intentionnalité de Brentano, comme visée d'un objet abstrait : elle signifie simplement 'vouloir dire' [*to mean*], c'est-à-dire 'être disposé à dire quelque chose'. La phénoménologie n'entre absolument pas en ligne de compte, ici.

[327] La croyance est donc une relation triadique entre un acte d'énonciation, un contenu d'énoncé et un locuteur. Les critiques de [Church 1954], [Fischer 1963] et [Mates 1954] contre l'argument de l'isomorphisme intensionnel de Carnap étaient justifiées à cet égard : l'identité de structure entre deux énoncés ne suffit pas pour identifier les croyances qu'ils expriment ; mais ces critiques n'impliquent pas un rejet des énoncés de langue naturelle dans l'analyse des croyances. Au contraire : toute croyance est liée à l'emploi d'une certain énoncé dans une certaine langue naturelle, celle du locuteur. Une analyse logique des croyances dans un langage formel doit prendre en compte ces facteurs indexicaux.

qu'il y ait des sens 'composés' ou qui 'ne sont pas en position référentielle' remplaçant des termes référentiels, et j'éviterai cependant la contradiction.[328]

Croire que ϕ, c'est donc être disposé à produire l'énonciation 'ϕ' à l'appui, mais au sens d'une disposition qui n'est satisfaite que si le locuteur pense effectivement que ϕ est vrai : même s'il est possible que ce locuteur découvre par la suite la vérité de ϕ, il ne le croit pas tant qu'il n'a pas des raisons de le déclarer. La disposition à croire que ϕ n'est donc pas une pure potentialité et nécessite la justification préalable de son énoncé, de sorte que croire ce que l'on dit et dire ce que l'on croit sont deux actions interdépendantes. Sans cette équivalence entre pensée et assertion ou annonce publique, toute attribution d'attitude à une tierce personne est susceptible de provoquer l'opacité référentielle et de faire des contextes épistémiques des contextes irréductiblement intensionnels. Mais notre redéfinition des modalités épistémiques en termes d'actes illocutoires évite cette conséquence tout en justifiant le renforcement par Hintikka de (LL) (elle ne doit pas porter sur des énoncés mais des énonciations, et le passage d'une entité à l'autre s'exprime par l'occurrence de $\Box$ devant (a = b)). La condition de décitation repose donc sur un échange ouvert préalable entre l'interprétant et l'interprété, dans lequel le sujet croyant annonce à son interlocuteur les croyances qui pourront lui être attribuées.

La solution illocutoire est économique : elle évite une explication de l'identification en termes implicites de structures de la conscience ou de visée d'un objet abstrait ; elle s'inscrit dans la sémantique des mondes possibles de Hintikka tout en réduisant les lignes de monde aux informations conservées (donc constantes) à chaque étape d'un échange entre des interlocuteurs. Chacune des étapes d'une ligne du monde, c'est-à-dire chaque monde possible, ne représente rien d'autre qu'une déclaration susceptible d'être considérée comme vraie par le locuteur et considérée comme acquise. On pourrait certes reprocher à cette interprétation d'être trop indexicale et de réduire les attributions de croyance à un régime drastique, celui de l'assertion. Mais de même que pour l'analyse des modalités itérées, l'approche illocutoire ou assertive des modalités épistémiques a l'avantage de confirmer la majorité de nos intuitions sur les attitudes et explique facilement le problème de l'opacité référentielle, sans s'embarrasser d'explications en termes d'idéologie, de néo-kantisme sémantique ou d'outils conceptuels. Notre analyse 'dialogique' des croyances et connaissances en termes d'actes de discours élimine en outre le problème de la vérité des énoncés de connaissance : la clause de vérité devient une clause de *véracité*, dans la mesure où attribuer une connaissance à

[328] [Zemach 1971], pp. 804-5. Nous suivons l'auteur sur le lien entre attitude épistémique et assertion, mais nous ne le suivrons pas dans toutes les conséquences qu'il en tire : Zemach prétend par exemple que $B\neg p$ n'implique pas forcément $\neg Bp$ et que des assertions peuvent être manifestement contradictoires, contrairement à nous. Il base en outre ce rejet sur la comparaison de B à $\Diamond$: $\Diamond\neg p$ n'implique pas $\neg\Diamond p$, donc la croyance de non-p n'implique pas la non-croyance de p. Or cela revient à traiter la croyance comme une modalité faible et à perdre l'analogie structurelle entre B, $\Box$ et $\forall$ défendue jusqu'ici. Pour justifier son rejet de l'inférence ($B\neg p \rightarrow \neg Bp$), enfin, Zemach propose également de remplacer la conséquence paradoxale ($Bp \wedge \neg Bp$) de l'opacité référentielle par la forme non-contradictoire ($Bp \wedge B\neg Bp$). Mais nous refuserons ici de juger ce qu'une personne pense croire à tort ($B\neg Bp$) indépendamment de ses assertions explicites. Nous maintiendrons donc l'inférence ($B\neg p \rightarrow \neg Bp$) mais rejetterons l'antécédent $B\neg p$: tant que Paul n'a pas *dit* au préalable 'mon père est un vendeur d'armes de destruction massive', on ne peut pas dire de lui qu'il croit $\neg p$ (mon père n'est pas un philanthrope), donc ses croyances ne seront pas contradictoires pour lui mais seulement *pour nous*.

quelqu'un se fait sur la base d'un acte d'assertion ou de force illocutoire supérieur à celui d'une simple supposition.

Le cadre dialogique (ou 'interactionnel', selon les termes de [Vernant 1997]) évite ainsi ou dissout plutôt le problème de l'opacité référentielle : les 'ontologies locales' du locuteur et de l'interlocuteur (ce qu'ils considèrent comme vrai, de leur point de vue sur le monde) contiennent des 'propositions' indexicales, c'est-à-dire des énonciations, informations ou annonces mutuelles : le succès de l'attribution de croyance ou de connaissance est garantie, sous cette condition. Mais cette attribution suppose qu'à la logique propositionnelle locutoire soit substituée une logique des actes propositionnels, ou illocutoire.[329] Pourquoi cette solution n'est-elle pas suggérée plus souvent pour la logique modale épistémique ? Sans doute parce qu'une analyse des concepts épistémologiques souhaite conserver la différence qualitative entre connaissance et croyance : le critère de vérité qui les sépare n'est pas seulement d'ordre discursif et renvoie à des états de choses, donc à une ontologie générale où un objet de connaissance est vrai également dans le monde *actuel*.

Dans le chapitre qui suit, nous allons tenter de montrer au contraire que, malgré la définition classique du savoir, une approche illocutoire des relations conceptuelles entre vérité et connaissance est préférable et permet de fusionner ces deux dernières notions : une connaissance n'est autre qu'une *déclaration de vérité*. Pour ce faire, nous examinerons tout d'abord les analyses générales proposées en épistémologie, dominées par le débat entre réalisme et anti-réalisme ; puis nous justifierons l'orientation illocutoire donnée au discours épistémologique au sein de la logique modale épistémique. Nous clôturerons ce travail par une fusion des systèmes de logique classique et de logique modale épistémique au sein d'un langage formel commun : la logique modale illocutoire.

[329] On décrit couramment la logique classique ou non-modale en termes de logique 'assertorique' ou 'déclarative', mais la formule est trompeuse : cette logique des propositions ne contient pas d'assertions ou déclarations, seulement des propositions susceptibles d'être assertées ou déclarées. C'est l'aspect interactionnel des énoncés ou 'propositions' que nous avons mis en valeur ici, afin de traiter de l'opacité référentielle.

Chapitre 3

Quelle logique des attitudes ?

Résumé : Une réorientation pragmatique est proposée pour la logique modale épistémique, afin de régler les ambiguïtés couramment produites par cette logique philosophique. Il n'y a pas de distinction fondamentale entre logique classique et logique épistémique, dans la mesure où les énoncés représentent un mode d'expression illocutoire parmi d'autres.

Après avoir montré que les théories rivales de la vérité conduisent à des impasses conceptuelles (section 3.1), nous évoquerons le Principe de Charité pour illustrer la relation indissociable existant entre les concepts de vérité, de signification et d'assertion (3.2). Les attitudes épistémiques servent d'intermédiaire pour saisir la signification d'un énoncé à partir de ses conditions de *déclaration de* vérité. Les actes d'assentiment et de dissentiment priment donc sur les attributions de vérité et de fausseté dans la compréhension d'un énoncé.

Le résultat de cette approche sera une logique d'assertion et de supposition, dont le but est de formaliser les notions de vérité et de connaissance avec un opérateur modal primitif. Nous présenterons la motivation et les applications de ce système (section 3.3), dont l'issue est une logique assertorique générale.

Ce qui est irrationnel, en fait, c'est le comportement d'un homme qui persisterait à souscrire à une affirmation indéfendable après que son indéfendabilité ait été portée à sa connaissance.
(Jaakko Hintikka)

Oui ou non, as-tu perdu tes cornes ? – Non ? tu les as donc encore ! – Oui ? tu conviens de les avoir portées ! Ainsi, quoi que tu répondes, cornu tu es ou cornu tu fus
(Sophisme du cornu)

Si la logique est un bien un discours, elle constitue un usage purement rationnel de la discursivité. L'engagement assertif requis par la règle de détachement est bien acte, mais c'est un acte qui engage, en un jeu conventionnellement et algorithmiquement déterminé, non la croyance idiosyncrasique d'une individualité dont la rationalité est limitée, mais le savoir d'un sujet épistémique, anonyme et universel, finalement réductible à l'ensemble des règles de calcul.
(Denis Vernant)

3.1 Vérité et connaissance

La logique épistémique représente un système charnière entre deux disciplines : la logique et sa notion centrale de vérité, d'une part ; l'épistémologie et sa notion centrale de connaissance, d'autre part. La relation entre les notions de vérité et de connaissance a été examinée notamment dans le cadre de la clause de vérité (section 2.2.4.5), où nous avons relevé une ambiguïté en ce qui concerne leurs relations conceptuelles. C'est l'expression 'être vrai' qui pose fondamentalement problème, lorsque la logique épistémique prétend servir de logique philosophique et ne pas être un simple langage formel. Que faut-il entendre par 'la vérité de l'énoncé ϕ', dans la clause de vérité $K\phi \rightarrow \phi$: que ϕ est vrai dans un monde possible qui représente l'espace épistémique d'un agent ; que ϕ est vrai dans le monde réel ; ni l'un ni l'autre en particulier ? En vertu de la réflexivité de la relation d'accessibilité R, chaque monde w est accessible à lui-même : la vérité de $\Box\phi$ dans w implique la vérité de ϕ dans tous les mondes accessibles à w, y compris w lui-même. Mais le problème concerne la signification de la vérité : la clause de vérité suppose qu'une formule ϕ est vraie dans l'espace épistémique propre à un agent, mais elle suppose également sa vérité indépendamment de cet espace épistémique : la vérité d'une connaissance est supposée valable pour tous, et c'est ce qui la distingue du concept de croyance. La différence entre les mondes de connaissance et les mondes de croyance réside dans l'idée selon laquelle ces derniers ne seraient pas liés au monde réel, contrairement aux premiers. La réflexivité de R symbolise ainsi l'objectivité de principe de la connaissance ; mais un débat est susceptible de mettre cette objectivité en doute, plus profond encore que le cas pragmatique de [Zemach 1969] puisqu'il porte sur les conditions de vérité des énoncés eux-mêmes et non sur leur énonciation par un sujet.

Vérité et connaissance sont-elles deux notions indépendantes, ou peut-on réduire l'une à l'autre ?[1] Cette question prend pour point de mire le débat sémantique entre réalisme et anti-réalisme. Pour le *réalisme sémantique*, les énoncés sont vrais ou faux indépendamment de nous, mais pas pour l'anti-réaliste. Faut-il en conclure que l'assignation de valeurs sémantiques dépend de considérations philosophiques de départ, y compris en logique ? C'est la relation complexe entre vérité ontologique, vérité épistémologique et vérité sémantique (neutre) qui se trouvera au cœur de ce débat, nous conduisant de l'une à l'autre conception de la vérité. Pour montrer l'impact du débat ci-dessus sur la logique épistémique, nous verrons que les deux doctrines philosophiques en question conduisent chacune à des impasses conceptuelles : le réalisme sémantique se heurte au problème de Gettier, d'une part, et semble favoriser le choix d'une sémantique anti-réaliste ; mais l'anti-réalisme se heurte pour sa part au paradoxe de Fitch.

Dans un premier temps, nous nous arrêterons sur le fameux problème de Gettier et ses conséquences sur la notion de connaissance : elles susciteront notre rejet de la définition classique de la connaissance et une réduction de celle-ci à la croyance justifiée. Les relations conceptuelles entre vérité, connaissance, croyance et justification seront réarrangées en conséquence, dans une perspective anti-réaliste.

---

[1] Elles sont indépendantes dans la logique modale épistémique considérée jusqu'ici, puisque ϕ peut être vrai sans que je le sache : $\phi \wedge \neg K\phi$. Cette formulation présuppose une interprétation réaliste de la vérité, comme nous allons le voir.

Dans un second temps, nous rappellerons que la traduction de la vérité en termes anti-réalistes de connaissabilité crée un autre paradoxe, le paradoxe de Fitch. Des stratégies existent, nombreuses, pour résoudre ce paradoxe sur la base d'une redéfinition des opérations logiques. Mais la question portera sur le symbolisme choisi pour représenter la relation conceptuelle entre vérité et connaissance : exprime-t-on cette relation comme il se doit au sein de la logique épistémique, ou ne faudrait-il pas penser autrement la relation entre vérité et connaissance dans le discours philosophique ? C'est ce que nous tenterons de faire par la suite (section 3.2).

3.1.1 Réalisme et anti-réalisme

Par *réalisme sémantique*, on entendra l'idée selon laquelle les conditions de vérité des énoncés transcendent nos conditions de reconnaissance de leur vérité : 'le soleil brillera à Naples le 30 août 2009' est un énoncé potentiel qui, bien qu'invérifiable aujourd'hui, peut être considéré comme significatif pour un réaliste sémantique, puisque nous savons ce qui doit être satisfait pour en faire un énoncé vrai.

Par *anti-réalisme sémantique*, on entend au contraire l'idée selon laquelle la vérité d'un énoncé ne transcende pas les conditions de sa reconnaissance et dépend de la capacité du locuteur à le justifier. L'énoncé ci-dessus n'a pas de sens dans la mesure où personne n'est capable à ce jour de vérifier si le soleil brillera à Naples le 30 août 2009. En somme, la différence entres le deux doctrines portent sur le rôle du sujet dans l'assignation de vérité à un énoncé. Il existe toutefois une version modérée concernant les conditions d'assertabilité d'un énoncé : un énoncé peut être considéré comme vrai par un anti-réaliste s'il est vérifiable en principe bien qu'invérifié dans les faits. Le cas de la bataille navale d'Aristote devient ainsi un exemple d'énoncé significatif pour l'anti-réaliste modéré.

Si l'on considère que le moyen d'exprimer la vérité est l'*énoncé* et non la *proposition*, nous constatons que le différend entre réalistes et anti-réalistes en est sérieusement affaibli. Pourquoi ? Le réalisme sémantique suppose le principe de *bivalence*, selon lequel toute proposition est soit vraie soit fausse, et l'ambiguïté vient de la nature du porteur de vérité : le débat à suivre entre réalisme et anti-réalisme porte-t-il sur les conditions de vérité de *propositions*, considérées comme le contenu commun à plusieurs énoncés synonymes, ou sur les *énoncés* qui les expriment ? Nous n'entrerons pas dans les détails de ce débat, mais chercherons uniquement à observer son impact sur les formulations de la logique épistémique. La difficulté créée par l'anti-réalisme provient de sa lecture *épistémique* de la vérité, par opposition au réaliste sémantique : une valeur de vérité telle que l'anti-réaliste l'envisage est portée par un énoncé et non par la proposition, laquelle correspond plus ou moins au contenu de l'énoncé ; or l'énoncé, lorsqu'il est prononcé par un locuteur, signifie pour ce dernier l'affirmation qu'il existe un fait vérifiant le contenu de l'énoncé, dans la mesure où la valeur sémantique ne peut pas être assignée à un énoncé sans qu'une condition de preuve ne lui soit apportée au préalable. Par opposition, la conception réaliste de la vérité porte plutôt sur la notion de proposition et sur la relation générale entre un énoncé et le fait qui le rend vrai ou faux, indépendamment des circonstances d'énonciation. Il devient alors possible d'assigner une valeur à un énoncé par simple hypothèse, sans justification demandée à l'appui. Le rejet intuitionniste du tiers exclu devient par ailleurs le résultat d'un malentendu sur la nature des

porteurs de vérité : la bivalence peut signifier à juste titre que toute proposition est vraie ou fausse si, par proposition, on entend un énoncé susceptible d'être vérifié, et l'objection de l'anti-réaliste n'est convaincante que si l'on conçoit le tiers exclu comme le principe selon lequel tout énoncé a une valeur de vérité déterminée par avance. Nous montrerons par la suite qu'il existe une autre manière d'interpréter le tiers exclu, capable de restituer la validité de ce principe sans faire aucun choix entre la position sémantique du réaliste et de l'antiréaliste : la relation entre les concepts de vérité et de connaissance peut être examinée logiquement sans prendre position sur ce sujet.[2]

L'effet de l'anti-réalisme sémantique doit pouvoir se vérifier sur les expressions et les principes de la logique modale épistémique. Que devient notamment la clause de vérité, pour un anti-réaliste ? $K\phi \rightarrow \phi$ demeure valide, dans la mesure où la condition de connaissabilité est supposée satisfaite dans l'antécédent : si ϕ est connu, alors ϕ est vrai puisqu'il a été vérifié. Mais si vérité et connaissance ne font qu'un pour l'anti-réaliste, l'expression de la clause de vérité se transforme en une trivialité de type $K\phi \rightarrow (K)\phi$, où la parenthèse indique que la connaissance de ϕ est présupposée dans l'expression de la vérité de ϕ. Un doute persiste toutefois sur la notion anti-réaliste de vérité : signifie-t-elle la connaissance actuelle ou la simple possibilité de connaître, à un instant quelconque du temps ? Puisque l'avis général favorise la seconde interprétation de l'anti-réalisme, plus modérée, la trivialité demeure mais sous une autre forme, de type $K\phi \rightarrow (\Diamond K)\phi$. Nous reviendrons sur la forme multimodale de la connaissabilité, dans le cas du paradoxe de Fitch. Des changements sont indispensables au sein de la logique modale épistémique, si l'on adopte une position anti-réaliste : par exemple, le paradoxe pragmatique exposé dans [Zemach 1969] produira une circularité *logique* si l'on maintient la clause de vérité tout en adoptant un point de vue comme celui de Dummett. De son point de vue assimilant vérité et croyance justifiée, la vérité du conséquent ϕ dans $K\phi \rightarrow \phi$ implique sa justification et ne peut être affirmée sans elle ; or dans ce cas, la converse de la clause de vérité devient un principe logique valide : $\phi \rightarrow K\phi$, dans la mesure où l'affirmation de ϕ présuppose son assertabilité. Et l'on conçoit l'opérateur K comme une expression de la croyance justifiée, ce que fit [Hintikka 1962] dans sa défense de la thèse KK, on obtient ainsi une équivalence triviale entre vérité et connaissance : $(K\phi \rightarrow \phi) \wedge (\phi \rightarrow K\phi) \rightarrow (K\phi \equiv \phi)$.[3]

La formulation de l'axiome T devient donc problématique et une logique modale épistémique anti-réaliste ne peut plus contenir la clause de vérité, d'inspiration réaliste ; on trouve une logique épistémique anti-réaliste de ce genre dans [Williamson 1990,1992]. Un

---

[2] [Wiredu 1975] affirme que le débat sur le tiers exclu provient d'une confusion entre deux acceptions de la notion de *vérité* : une vérité première, qui exprime la valeur sémantique des énoncés (ϕ est vrai, noté $T\phi$) ; une vérité seconde ou 'comparative', qui exprime un jugement sur la valeur assignée à un énoncé (il est vrai que ϕ est vrai, noté T_2'T_1'ϕ). D'après Wiredu, le tiers exclu a suscité un faux débat parce que le réaliste sémantique s'est mépris sur son interprétation ; le tiers exclu ne prétend pas en réalité que chacun de nous est capable d'évaluer un énoncé à tout instant : $T\phi \vee F\phi$, mais que chacun de nous peut porter un jugement sur la valeur des énoncés : T_2'T_1'$\phi \vee F_2$'T_1'ϕ. L'issue de ce chapitre sera conforme à l'approche de Wiredu car son traitement de la vérité comme constante logique sera maintenue : dans l'énoncé 'il est vrai que ϕ', ϕ est une proposition (ou complétive) et représente une fonction saturée (ou complétée) par la déclaration de vérité T. Mais plutôt que de parler de vérité, nous parlerons d'assertion, avec $A\phi$ correspondant à $T\phi$ et $A\neg\phi$ à $F\phi$.

[3] La parenté entre l'anti-réaliste et le paradoxe de [Zemach 1969] est normale, attendu que leur expression de la vérité est centrée sur le point de vue du sujet. Mais elle l'est pour deux raisons distinctes : chez Zemach, c'est la *déclaration de vérité* de ϕ qui présuppose la connaissance de ϕ par le locuteur ; chez l'anti-réaliste, c'est la *vérité* même de ϕ qui présuppose sa connaissance (devenue croyance justifiée).

choix doit être fait entre l'interprétation anti-réaliste des attitudes épistémiques et leur interprétation réaliste, laquelle a été supposée jusqu'ici dans la lignée de [Hintikka 1962]. Un réarrangement des relations conceptuelles entre vérité et connaissance devrait donner la chose suivante. Dans la définition classique du savoir comme croyance vraie justifiée, $K\phi \equiv B\phi \wedge \phi \wedge J\phi$, la vérité de ϕ est exprimée indépendamment de sa justification et constitue une condition nécessaire à la connaissance. Une telle définition est clairement réaliste, d'autant plus qu'elle s'est affirmée à l'origine contre la définition du savoir faible ou simple opinion vraie. Les expressions de savoir fort et savoir faible sous-entendent ainsi l'indépendance de la vérité à l'égard de la justification. Dans une perspective anti-réaliste, la vérité n'est plus indépendante mais devient synonyme d'une connaissabilité préalable, formalisée plus tard par $\Diamond K$. Or puisque cette connaissance potentielle est assimilable chez Dummett à une forme de croyance justifiée,[4] la définition classique du savoir doit être abolie et laisser place à une autre hiérarchie entre les quatre notions épistémologiques ; le *definiendum* K doit être remplacé par sa version affaiblie $B\phi \wedge J\phi$, et l'on obtient la définition anti-réaliste suivante : $\phi =_{df} B\phi \wedge J\phi$. Le résultat est une réduction symbolique de la connaissance à la croyance justifiée, conformément au savoir fort de [Hintikka 1962] mais sans la clause de vérité.

Les deux points de vue décrits dans ce qui suit favorisent tour à tour le choix puis le rejet d'une sémantique de type anti-réaliste. D'une part, le problème de Gettier montre que la définition classique ou réaliste du savoir est insatisfaisante et conduit à une interprétation faillible de nos connaissances. Une telle interprétation incite à rejeter la clause de succès, si du moins l'on interprète ϕ comme l'expression d'une vérité de fait. D'autre part, la redéfinition de la vérité en termes affaiblis de connaissabilité conduit à une contradiction et remet en question le choix de la symbolisation pour les termes épistémologiques. Examinons ces deux problèmes de conceptualisation du savoir et de la vérité.

3.1.1.1 Un mérite : le problème de Gettier

Le but de [Gettier 1963] est de montrer que les trois conditions classiques du savoir sont peut-être nécessaires mais ne sont pas suffisantes, d'où la nécessité d'ajouter une quatrième condition pour saisir le sens de 'savoir'. Deux exemples sont donnés. Dans le premier, Smith postule avec Jones pour un même job et a de fortes raisons de croire que c'est lui qui l'obtiendra ; de plus, Smith a compté dix pièces de monnaie dans la poche de Jones et en conclut qu'il croit à juste titre la conjonction suivante : (a) 'Jones aura le job, et Jones a dix pièces de monnaie dans sa poche'. De cette croyance justifiée, il en infère la conclusion suivante : (b) celui qui décrochera le job a dix pièces de monnaie dans sa poche (par généralisation existentielle à partir de l'individu Jones). Gettier montre par cet exemple que l'on peut avoir une croyance vraie justifiée sans posséder une connaissance à proprement parler : il se trouve que ce n'est pas Jones mais Smith qui décrochera finalement le job, et que Smith a également dix pièces de monnaie dans sa poche. En conséquence, (a) est faux et (b)

[4] L'idée de Dummett "est de compter pour ni vrai ni faux un énoncé de la science de la nature si aucune procédure n'est connue qui donne de solides arguments empiriques en faveur de sa vérité ou de sa fausseté" ([Quine 1993], pp. 133-134). A quoi [Quine 1993] répond par une conception sémantique ou neutre de la vérité, distinguant ainsi vérité et reconnaissance de la vérité : "La vérité est une chose, la croyance justifiée en est une autre. Nous pouvons gagner en clarté et profiter de l'aimable simplicité de la logique bivalente en tenant compte de la distinction" (p. 134).

est vrai, donc (b) exprime une croyance vraie justifiée qui a été obtenue à partir d'une croyance fausse. Pour cette raison, Gettier estime que (b) ne constitue pas une connaissance à part entière :

> Smith ne sait *pas* que [(b)] est vraie ; car [(b)] est vraie en vertu du nombre de pièce de monnaies dans la poche de Smith, alors que Smith ne *sait pas* combien de pièces de monnaie sont dans la poche de Smith, et il base sa croyance de [(b)] sur un décompte des pièces de monnaie dans la poche de Jones, dont il croit *à tort* qu'il est l'homme qui obtiendra le poste.[5]

L'autre exemple est similaire et produit une croyance vraie justifiée sur la base d'une évidence fausse : Smith a de fortes raisons de croire que Jones possède une Ford parce qu'il l'a souvent vu la conduire ; il croit donc de façon justifiée à l'énoncé (a) : 'Jones possède une Ford'. Smith infère de (a) les trois conclusions suivantes (par introduction de la disjonction : $A \Rightarrow A \vee B$) : (b) soit Jones possède une Ford, soit Brown est à Boston ; (c) Soit Jones possède une Ford, soit Brown est à Barcelone ; (d) Soit Jones possède une Ford, soit Brown est à Brest-Litovsk. Gettier pose ensuite les deux faits suivants : Jones ne possède pas en réalité une voiture Ford, mais il est vrai que Brown se trouve actuellement à Barcelone. La conclusion (b) est donc une croyance justifiée et vraie mais ne constitue cependant pas une connaissance.

De ces deux exemples, [Gettier 1963] conclut que les trois conditions classiques du savoir ne sont pas suffisantes et doivent être renforcées par une quatrième : la justification doit être 'bonne', si l'on veut faire de la croyance justifiée une connaissance. La littérature créée autour du problème de Gettier consiste à rechercher les conditions pour qu'une justification soit appropriée. Pour ne citer qu'un exemple, [Lehrer & Paxson 1969] proposent la distinction entre connaissance basique et non basique, qui sert à compléter la justification par une condition d'indéfectibilité [*undefeasibility*] : soit une connaissance se justifie d'elle-même et n'a pas besoin d'évidence supplémentaire à l'appui, à l'instar du 'fondationalisme' de Chisholm ; soit la connaissance doit être justifiée par une autre affirmation ou preuve à l'appui, à l'instar de l'approche 'fiabiliste' de la connaissance. Pour résumer, l'insuffisance de la définition classique du savoir provient du flou qui entoure la notion de 'justification' dans la définition de la croyance vraie : l'application ci-dessus des règles de généralisation existentielle et d'introduction de la disjonction a produit des croyances vraies justifiées sur la base de jugements faux (Jones n'aura pas le job, Jones ne possède pas de Ford). Autrement dit, une croyance vraie justifiée ne constitue une connaissance que si la justification de départ est valable ; mais le problème porte désormais sur la valeur de la justification : à quel moment une justification peut-elle être considérée comme indéfectible [*undefeasible*] au point de produire une connaissance ? Les exemples ci-dessus de Gettier ont exprimé la définition forte du savoir selon Hintikka, où K désigne une opinion vraie justifiée et dont la justification est *connue* du locuteur : si Smith sait que ϕ, alors il doit être en mesure de dire pourquoi ϕ est vrai et ne pas le croire pour d'autres raisons involontaires, comme ce fut le cas dans les exemples précédents. En d'autres termes, une croyance vraie justifiée de ϕ est une connaissance si elle

[5] [Gettier 1963], p. 122.

repose sur une bonne raison ψ, c'est-à-dire la raison en vertu de laquelle l'agent croit que ϕ. Cette condition se traduit en termes de preuve indéfectible.

Le débat sur les conditions d'indéfectibilité d'une justification nous conduit dans des complications théoriques qui semblent déborder le cadre de la logique épistémique : le savoir fort de Hintikka stipule d'emblée que l'opérateur K est une croyance suffisamment justifiée pour couper court à toute contestation, si bien qu'il présuppose l'indéfectibilité de la justification sans rechercher ses conditions pratiques.[6] Son analyse logique ne dit pas comment cette justification ultime doit être obtenue pour satisfaire $K\phi$. On pourrait répondre que la logique épistémique ne doit pas plus concerner l'épistémologie que la logique classique ne concerne l'ontologie : peu importe les conditions pratiques dans lesquelles une justification est obtenue, si la logique épistémique se contente de poser cette connaissance obtenue par hypothèse. Le problème de Gettier peut affecter toutefois le langage de la logique épistémique de deux manières opposées, selon que l'on prétend lui trouver une solution ou non. Dans le premier cas, on peut proposer de remplacer la relation conditionnelle entre une croyance vraie et sa justification par une relation d'*entraînement*, ou d'implication stricte ; le but de cette révision du conditionnel classique serait d'exprimer la condition suffisante à l'attribution de connaissance par un lien logique plus strict (quasi-causal) entre la croyance et sa justification. Dans le cas contraire, l'absence de conditions d'indéfectibilité suffisantes signifie qu'une croyance justifiée n'est pas une connaissance totalement garantie et que sa vérité n'est que provisoire.

Arrêtons-nous d'abord sur la solution de la relation d'implication stricte ou de *conditionnel contrefactuel*, en vue de garantir le lien entre une croyance vraie et sa 'bonne' justification. Si la connaissance de ϕ a besoin d'une preuve ψ, un moyen d'exprimer la relation de justification entre ψ et ϕ consiste à modifier leur relation logique : la relation des deux exemples de Gettier était trop faible pour garantir l'attribution de connaissance de ϕ, et une implication plus forte pourrait être invoquée en remplacement. Par exemple, une quatrième condition nécessaire à la connaissance de ϕ pourrait être le fait que la preuve ψ entraîne ou *cause* nécessairement ϕ et soit telle que l'implication entre ψ et ϕ est fausse si ψ est faux. Une telle implication n'est pas simplement conditionnelle, comme dans les cas de généralisation existentielle et d'introduction de la disjonction de [Gettier 1963] : elle fait place à une sorte de nécessité causale ou stricte, et le rôle de l'épistémologue consiste alors à interpréter le sens de cette connexion. Cette quatrième condition de la connaissance en termes de conditionnel contrefactuel ne sera pas suivie ici, pour une raison liée à la section suivante et qui a trait au principe de charité : la nécessité causale fait appel à des conditions de vérité en termes de

[6] [Lenzen 1978] a constaté que la frontière entre les questions d'épistémologie et de logique épistémique est difficile à fixer, d'autant plus lorsque l'on constate la part croissante de l'épistémologie formelle. L'auteur précise cependant à deux endroits que la logique épistémique n'est pas plus tenue de refléter une épistémologie particulière que la logique classique n'est tenue de refléter une ontologie particulière : "La recherche de l'analyse correcte du savoir, alors qu'elle est certainement d'une extrême importance et d'un intérêt extrême pour l'épistémologie, ne semble pas affecter de manière significative l'objet de la logique épistémique, c'est-à-dire la question de la validité de certains principes logico-épistémiques" (p. 27) ; "la question de savoir si quelque chose est réellement tel qu'il semble être ou n'est pas réellement tel qu'il semble être est une question qui mérite sans aucun doute que l'on philosophe à son sujet. Mais celle-ci appartient à l'épistémologie plutôt qu'à la logique épistémique" (X). Mais cette question affecte le contenu des règles logico-épistémiques ainsi que le choix de leur forme logique, et ce chapitre 3 a pour but précis de remédier à ces difficultés d'ordre *conceptuel*.

mondes possibles, mais ces mondes ne représentent plus des scénarios imaginés par un agent et font appel cette fois-ci à des conventions scientifiques discutables. La nécessité causale n'est que peu ou prou mobilisée par Hintikka dans son analyse de la connaissance, et nous la laisserons également de côté pour maintenir une approche naturaliste des mots logiques. La nécessité causale ne "fera pas partie de notre nourriture", pour reprendre une formule de [Quine 1977].[7]

Passons maintenant à l'attitude opposée, celle qui consiste à abandonner la recherche de la quatrième condition de K et à assimiler en conséquence les notions de connaissance et de croyance justifiée. Cette démarche peut être suivie afin d'éviter les objections du sceptique face à la faillibilité (ou défectibilité) de nos justifications : comme nous l'avions vu avec [Zemach 1969], la vérité est peut-être atteinte par une croyance justifiée, mais rien ne semble permettre à un locuteur de s'attribuer une déclaration de connaissance *in*contestable. La vérité de ϕ peut être le cas sans que celui-ci n'en ait la preuve ou n'en ait conscience, et ni le sentiment de certitude (proposé par Ayer en remplacement du critère de preuve) ni l'existence d'une justification ne permet de justifier la vérité de ϕ. Une connaissance paraît toujours faillible et ne pas en être une à proprement parler, de ce point de vue.

Un moyen de combler le fossé entre croyance justifiée et connaissance consiste à garantir (et non plus justifier) la vérité d'un énoncé par lui-même et sans preuve supplémentaire à l'appui, que ce soit par des facultés perceptives indubitables (si l'énoncé est un énoncé d'observation) ou par des vérités conceptuelles qui s'imposeraient d'elles-mêmes (si l'énoncé est un énoncé théorique). Un autre moyen consiste à éliminer la notion de vérité dans la définition de la connaissance ou, plutôt, à redéfinir la vérité en termes affaiblis de croyance justifiée. Une croyance peut être vraie et ne pas être justifiée, comme dans les exemples de [Gettier 1963] ; à l'inverse, une croyance peut être justifiée et ne pas être vraie, à l'insu du locuteur. La position anti-réaliste est en mesure d'échapper à ce dilemme, puisqu'elle assimile la vérité à la croyance simplement justifi*able* et fait de la connaissance potentielle un synonyme de vérité. Potentielle, uniquement : le degré de suffisance qu'une justification doit comporter pour conduire à la vérité est le problème du sceptique soucieux de réalité, mais ce n'est pas celui de l'anti-réaliste pour qui la vérité est une construction de l'esprit.

Si l'on suit cette voie réductrice, alors il n'y a plus de différence de nature entre connaissance et croyance justifiée : la vérité d'un énoncé devient le résultat d'une assertion ou affirmation justifiée ou assertion. On obtient ainsi la relation formelle : $\phi \rightarrow (B\phi \wedge J\phi)$, dans laquelle l'anti-réaliste compare la vérité à la notion d'assertabilité garantie ; dans cette perspective, la formulation du problème de Gettier devient le vestige d'une distinction réaliste entre vérité et connaissance. Une fois ces deux notions fusionnées en termes de croyance justifiée, la clause d'introspection positive (ou thèse KK) est confirmée mais devrait être reformulée par un opérateur affaibli d'assertion sans clause de vérité. Si l'on concède au sceptique l'impossibilité de certifier la vérité d'une déclaration de 'savoir' (de croyance justifiée), la différence entre connaissance et croyance s'en retrouve estompée (le critère de vérité qui les distinguait a disparu entre-temps) et repose uniquement sur l'existence d'une

[7] Voir section 2.1.2.4, note 70. Parmi les critères heuristiques susceptibles d'expliquer le choix d'une logique de préférence à une autre, nous suivons avec Quine celui de la simplicité plutôt que celui de l'expressivité, en conséquence de quoi nous n'introduirons pas la nécessité causale pour l'analyse de la connaissance.

justification intermédiaire plus ou moins bonne. Pour échapper à l'objection du sceptique, une autre option consisterait à affaiblir les critères de vérité et à fixer un critère de sévérité minimal pour garantir la vérité ; mais une telle procédure modifierait le sens ontologique que nous attribuons ici à la vérité, c'est-à-dire comme l'expression d'un fait avéré. En résumé, le fossé qui sépare la connaissance et la vérité conduit à une redéfinition des notions épistémiques : vérité, connaissance et croyance justifiée sont tout un.

Le problème de la vérité se transforme ainsi en un problème pour *dire* la vérité : qu'est-ce qui, lors d'une déclaration de connaissance, donne le droit de dire que le contenu de cette déclaration est vrai ? Le passage du premier type de questionnement au second incarne le passage d'une logique d'énoncés à une logique d'énonciation, comparable aux traitements précédents de l'introspection positive et de la clause de vérité. Dans les deux cas, l'*énoncé* du schéma **T** ne fait plus sens dans une interprétation anti-réaliste de la vérité, et seule son *énonciation* en a un : 'si je sais que ϕ, alors ϕ est vrai' est une énonciation douée de sens et échappe à l'objection du sceptique, puisque la déclaration de connaissance est seulement hypothétique ici. Ce changement d'interprétation en vaut-il la peine ? La réduction anti-réaliste de la vérité à la croyance justifiée semble conciliable dans une moindre mesure avec l'interprétation que [Hintikka 1962] avait proposée pour la thèse KK, dans la mesure où celle-ci avait été interprétée en termes d'assertion. Dans cet ordre d'idées, le problème de la justification de la vérité est remplacé par une étude des conséquences logiques d'une déclaration de savoir, quelles que soient les conditions préalables à cette déclaration. Bien que l'acte pragmatique d'assertion qui en résulte puisse avoir un sens légèrement différent de l'assertabilité anti-réaliste dans ses conditions de satisfaction, l'idée de réduire la notion de vérité à un acte d'assertion (de déclaration de vérité) se confirme dans ce travail mais va rencontrer toutefois une difficulté, ci-dessous. Cette difficulté est liée à [Fitch 1963] et montre qu'une approche anti-réaliste des relations entre les concepts de vérité et connaissance produit une conclusion paradoxale.

3.1.1.2 Une limite : le paradoxe de la connaissabilité de Fitch

Le symbolisme logique est-il capable de restituer correctement la thèse du vérificationnisme, contenue dans la position anti-réaliste de Dummett ? C'est cette question que le paradoxe suivant de Fitch[8] a mise en valeur : à partir d'un ensemble de règles de déduction supposées correctes, l'application de la définition anti-réaliste de la vérité implique que toutes les vérités sont non seulement *connaissables* mais *connues*. Tel est le cas du moins si l'on se fie au symbolisme modal et aux règles déductives que pose Fitch, ici. Il n'est pas dans notre intention de dresser un inventaire complet des solutions proposées,[9] mais plutôt de trouver un moyen de régler le problème par une certaine interprétation, performative, de son symbolisme.

Dans l'approche vérificationniste de l'anti-réalisme, un énoncé est vrai s'il est possible d'en donner une preuve ou 'évidence' ; la définition de la vérité en termes d'assertabilité

[8] Ce paradoxe est associé au nom de Fitch parce qu'il est apparu dans [Fitch 1963], mais l'auteur reconnut lui-même que l'argument paradoxal fut découvert par un correcteur anonyme, probablement Alonzo Church. Sur l'origine véritable du paradoxe de 'Church-Fitch', voir J. Salerno : "Who discovered Fitch's paradox, and why won't it go away ?", in Pacific Division of the APA, à paraître.

[9] Nous renvoyons pour cela à la recension de [Brogaard & Salerno 2004].

garantie implique que la vérité est une forme de connaissabilité, symbolisée : $\phi \rightarrow \Diamond K\phi$. Or cette formule conduit à la trivialisation de l'opérateur K : tout ce qui est vrai est connaissable *et* connu : $\phi \rightarrow K\phi$, donc en vertu de la clause de vérité $K\phi \rightarrow \phi$ on obtient $\phi \equiv K\phi$. Nous allons revenir sur les étapes de la démonstration, avant de proposer une lecture du formalisme qui rende acceptable la trivialisation ci-dessus.

[Wansing 2002] propose deux versions du paradoxe. Une première version (non quantifiée) repose sur la définition anti-réaliste de la vérité comme connaissabilité (AR) : $\phi \rightarrow \Diamond K\phi$, la distribution de la connaissance sur la conjonction (**K**) : $K(\phi \wedge \psi) \equiv K\phi \wedge K\psi$, la clause de vérité (**T**) : $K\phi \rightarrow \phi$, et la contradiction impossible de la connaissance (NC) : $\neg\Diamond(K\phi \wedge \neg K\phi)$. On obtient la démonstration suivante :

1. $\phi \wedge \neg K\phi$	prémisse (non-omniscience)
2. $\Diamond K(\phi \wedge \neg K\phi)$	1., (AR)
3. $\Diamond(K\phi \wedge K\neg K\phi)$	2., (**K**)
4. $\Diamond(K\phi \wedge \neg K\phi)$	3., (**T**)
5. $\neg\Diamond(K\phi \wedge \neg K\phi)$	(NC)
6. $\neg(\phi \wedge \neg K\phi)$	4.-5.-1., réduction par l'absurde
7. $\phi \rightarrow K\phi$	6., déf. $\neg\wedge/\rightarrow$

Une seconde version, plus intéressante de notre point de vue pragmatique, part de la prémisse selon laquelle quelqu'un connaît la vérité d'une chose qu'il ne connaît pas, et y ajoute une règle de nécessitation inversée ($\neg\Diamond$) : si les tautologies sont nécessaires, $\vdash\phi \Rightarrow \vdash\Box\phi$, à l'inverse les antilogies sont impossibles, $\vdash\neg\phi \Rightarrow \vdash\Box\neg\phi$, c'est-à-dire $\vdash\neg\phi \Rightarrow \vdash\neg\Diamond\phi$. La conclusion obtenue permet d'obtenir de nouveau un paradoxe et par d'autres moyens.

1. $K(\phi \wedge \neg K\phi)$	théorème épistémique
2. $K\phi \wedge K\neg K\phi$	1., (**K**)
3. $K\phi$	2., (C.$\wedge$)
4. $K\neg K\phi$	2., (C.$\wedge$)
5. $\neg K\phi$	4., (**T**)
6. $\neg K(\phi \wedge \neg K\phi)$	3.-5.-1., réduction par l'absurde
7. $\neg\Diamond K(\phi \wedge \neg K\phi)$	6., ($\neg\Diamond$)

Cette conclusion obtenue est démontrable dans une logique modale épistémique monomodale (de type [Hintikka 1962]) : l'étape 1 est un théorème épistémique, en effet, et puisque la nécessitation inversée ($\neg\Diamond$) s'applique aux formules logiquement contradictoires, l'étape 6 est considérée comme impossible parce qu'elle représente la négation d'un théorème. On en déduit ainsi le paradoxe :

8. $(\phi \wedge \neg K\phi)$	prémisse (non omniscience)
9. $\Diamond K(\phi \wedge \neg K\phi)$	1., (AR)
10. $(\phi \wedge \neg K\phi) \rightarrow \Diamond K(\phi \wedge \neg K\phi))$	1.-2., introduction de l'implication

11. $\neg\Diamond K(\phi \wedge \neg K\phi)$	7.
12. $\neg(\phi \wedge \neg K\phi)$	10.-11., réduction par l'absurde
13. $\phi \rightarrow K\phi$	12., déf. $\neg\wedge/\rightarrow$

Il existe encore d'autres façons de déduire la trivialisation de l'opérateur K à partir de (AR), et nous n'en citerons que quelques-unes. Plusieurs réactions sont possibles face à ce paradoxe, qui conduit à la trivialisation ou réduction de $K\phi$ à ϕ : soit l'on admet que la correction des étapes déductives implique le rejet de (AR), concluant que la réduction vérificationniste de la vérité est contradictoire dans les termes ; soit l'on cherche à maintenir (AR) en montrant qu'elle ne constitue pas en réalité une contradiction dans les termes. Pour ce faire, l'une des étapes de la démonstration doit être réfutée afin de bloquer la conclusion indésirable. Les exemples qui suivent utilisent cette seconde voie.

[Brogaard & Salerno 2004] citent par exemple une définition que Williamson propose pour la prémisse de non-omniscience : l'idée selon laquelle toutes les vérités ne sont pas connues a la forme $\neg(\forall\phi)(\phi \rightarrow K\phi)$; cette version quantifiée a pour effet de bloquer la démonstration puisque, d'un point de vue *intuitionniste*, $\neg\forall\phi$ n'est pas équivalent à $\exists\phi\neg$, donc l'inférence $(\exists\phi)(\phi \wedge \neg K\phi)$ ne peut pas être utilisée et empêche d'obtenir la conclusion $\phi \rightarrow K\phi$ en vertu de la réduction par l'absurde. D'autre part, la réduction par l'absurde est elle-même une règle que les intuitionnistes rejettent et leur donne une raison supplémentaire de refuser les deux versions précédentes du paradoxe de Fitch, avec ou sans quantification. La stratégie de [Wansing 2002] propose de modifier les propriétés de la connaissance à travers une logique multimodale, à la fois épistémique et relevante. Certaines inférences opérées sur K sont dès lors interdites et bloquent la preuve de $\phi \rightarrow K\phi$. Une autre stratégie de révision des constantes logiques, la stratégie *paraconsistante*, consiste à admettre certaines formes de contradiction et à rejeter par conséquent la règle d'inférence (NC).

La question est celle-ci : quel est le prix à payer pour une révision des constantes logiques, et la résolution du paradoxe justifie-t-elle un changement des propriétés de K ? Plus importante encore est la question de savoir si cette stratégie de blocage de la conclusion paradoxale consiste à *résoudre* ou à *dissoudre* le paradoxe de Fitch : le but est-il de bloquer le passage à la conclusion indésirable, c'est-à-dire d'obtenir par tous les moyens logiques concevables la non-validité de l'énoncé paradoxal $\phi \rightarrow K\phi$; ou s'agit-il d'admettre le résultat mais de montrer que le paradoxe repose sur une interprétation trompeuse du symbolisme ?[10] Une résolution passe par la révision des règles d'inférences, quitte à ce que la procédure de blocage soit effectuée de manière *ad hoc*. Or étant donné que, dans le cadre de cette thèse, les propriétés de l'opérateur de connaissance seront maintenues sous leur forme logique *normale* (classique, ou monotone) et justifiées comme telles par la suite, la stratégie paraconsistante ou relevante ne sera pas suivie ici ; le paradoxe de Fitch n'est pas résolu mais dissoluble, de notre point de vue. Plutôt que d'opter pour une révision des constantes, stratégie des logiciens

[10] [van Benthem 2004] classe les réponses au paradoxe de Fitch en deux catégories de personnes : "Certains affaiblissent toujours plus la logique de l'argumentation. C'est comme baisser le volume de votre radio de façon à ne pas entendre les mauvaises nouvelles. Vous n'entendrez pas beaucoup de bonnes nouvelles non plus. D'autres remèdes laissent la logique en l'état, mais elles affaiblissent le principe de vérification lui-même. C'est comme censurer les nouvelles : vous entendez les choses à haute voix et clairement, mais elles peuvent ne pas être si intéressantes que cela" (p. 95). [Wansing 2002] appartient à la première catégorie.

déviants et que nous rejetterons par ailleurs (voir section 3.2), nous proposerons plutôt de revenir sur les deux démonstrations ci-dessus et d'examiner si l'une des étapes déductives empruntée est conforme à notre 'signification de base' de K.

Le paradoxe peut être dissout d'un certain point de vue, le point de vue *performatif* de K, si l'on examine le passage crucial menant de l'étape 1 à l'étape 2. Après substitution de ϕ par $(\phi \wedge \neg K\phi)$ dans (AR), on obtient la formule $(\phi \wedge \neg K\phi) \rightarrow \Diamond K(\phi \wedge \neg K\phi)$ dont le conséquent est réfuté en vertu du théorème 1 de [Fitch 1963] :

> Si α est une classe de vérité qui est close par rapport à l'élimination de la conjonction, alors la proposition $[p \wedge \neg(\alpha p)]$, qui déclare que p est vraie mais n'est pas un membre de α (où p est une proposition quelconque), n'est nécessairement pas un membre de α elle-même.

où α désigne un 'concept de valeur' tel que K.[11] Les règles de (**K**) et (**T**) permettent ensuite de prouver la fausseté de $\Diamond K(\phi \wedge \neg K\phi)$ et de conclure par l'absurde la fausseté de l'antécédent $\phi \wedge \neg K\phi$. Or que signifie $\Diamond K(\phi \wedge \neg K\phi)$, dont la fausseté détermine tout la suite du raisonnement ? Cette formule pourrait vouloir dire qu'il est possible de connaître *dans l'avenir* ce qui est *encore* inconnu au temps présent, auquel cas $\Diamond K(\phi \wedge \neg K\phi)$ serait pourvu de sens. Mais la formule n'utilise pas d'opérateurs temporels ou ne les suppose pas, usant uniquement d'un opérateurs aléthique $\Diamond$ et d'un opérateur épistémique K. Un autre moyen d'admettre la vérité de $\Diamond(\phi \wedge \neg K\phi)$ serait de l'interpréter comme l'expression modale du théorème d'incomplétude de Gödel : il existe dans la théorie des nombres axiomatisée des formules A qui ne sont pas prouvables et dont la négation $\neg A$ n'est pas prouvable non plus. Dans cette approche, K est assimilé à un opérateur de prouvabilité. Mais cette version de l'étape 2 ne semble pas fonctionner davantage : si ϕ désigne dans $\Diamond K(\phi \wedge \neg K\phi)$ la formule non prouvable en question, alors ϕ ne peut pas être *dite* 'vraie' dans le même temps. De plus, la formulation du théorème d'incomplétude de Gödel devrait donner plutôt une autre formulation modale de type $K(\phi \wedge \neg\Diamond K\phi)$: je sais qu'il y a certaines formules ϕ que je ne peux pas démontrer. Une confusion de ce genre sur la portée de $\Diamond$ avait déjà été signalée plus tôt (section 2.2.4.5), au sujet de la clause de vérité et de l'argument sceptique de l'impossibilité de connaissance.

Une autre façon moins coûteuse de comprendre l'impossibilité de $K(\phi \wedge \neg K\phi)$ est l'interprétation *performative* de l'opérateur K, considéré comme une déclaration de savoir ou assertion : pour reprendre le paradoxe pragmatique de [Zemach 1969], a-t-on le droit de *dire* qu'un énoncé est vrai et cependant inconnu ? L'étape 2 renvoie de nouveau au problème du paradoxe de Moore, lorsqu'on l'interprète comme une énonciation et non comme un énoncé. Dans l'analyse du paradoxe de Moore liée à la thèse KK, nous avions expliqué que l'assertion de '$\phi \wedge \neg K\phi$' était inadmissible à moins de situer l'ignorance $\neg K\phi$ dans un temps passé, c'est-à-dire 'ϕ est vrai, mais je ne le savais pas' ; or nous venons de signaler que le paradoxe de Fitch ne prend pas en compte la différence temporelle entre ϕ et $\neg K\phi$ dans $(\phi \wedge \neg K\phi)$.

---

[11] [Fitch 1963], p. 138. La preuve du théorème 1 repose sur une réduction par l'absurde : "Supposons au contraire que $[p \wedge \neg(\alpha p)]$ soit un membre de α; *autrement dit*, supposons que $(\alpha(p \wedge \neg(\alpha p))$". Comment donner un sens à la supposition que $(p \wedge \neg(\alpha p))$ soit un membre de α ? L'interprétation performative convient parfaitement, puisqu'elle revient à nier que l'on puisse affirmer 'p, mais je ne sais pas que p'.

En somme, notre approche de K confirme et donne un sens à la dérivation du théorème 1 de Fitch, confirmant par la même son résultat paradoxal de trivialité : $\phi \rightarrow K\phi$. Mais dans le même temps, la lecture performative de K permet de rendre compte de ce résultat soi-disant paradoxal : cette formule n'aboutit pas à la destruction de la logique épistémique mais à une réinterprétation des énoncés du paradoxe de Fitch. Ce n'est pas en termes d'énoncés mais d'énonciations que le raisonnement doit être interprété, selon nous. Dans cette perspective, la formule $\phi \rightarrow K\phi$ ne signifie pas que toute vérité est connue mais que toute *déclaration de* vérité constitue une *déclaration de* savoir : $K(\phi \rightarrow K\phi)$, justifiée ou non. La clause (**T**) utilisée dans la preuve du paradoxe de Fitch est maintenue, mais dans la mesure où elle signifie non pas une clause de *vérité* (ni même de *véridicité*) mais une clause de *véracité* ou disposition à dire la vérité: dire que l'on sait quelque chose, c'est s'engager sur sa vérité sans l'avoir démontré pour autant. Pour symboliser ce passage de l'aspect locutoire de la connaissance à l'aspect illocutoire des déclarations de connaissance, nous proposerons par la suite (section 3.3) une assimilation des notions de vérité et de connaissance au terme unique de l'assertion. Ce changement dans le symbolisme permettra de rendre compte du résultat $\phi \rightarrow K\phi$ et, surtout, montrera que le paradoxe de Fitch n'en est plus un : dans une logique des assertions, les symboles ϕ et $K\phi$ n'ont plus le sens qu'ils avaient dans une logique modale épistémique courante, signifiant désormais des *déclarations de* vérité et de connaissance. Une logique des assertions devra tenir compte de cet aspect déclaratif dans son symbolisme. De cette façon, nous ne venons pas de résoudre le paradoxe de Fitch mais de la dissoudre, et ce sans affaiblir aucune des règles d'inférence proposées dans l'argument d'origine.[12] Ce ne sont pas les opérations entre les symboles qui ont changé mais le sens des symboles eux-mêmes, ce qui nous conduira par la suite à une reconstruction symbolique des relations entre les concepts de vérité et connaissance.

Le prix à payer pour cette démarche n'est pas excessif, puisqu'elle est conforme aux conditions sous lesquelles la thèse KK a été défendue tout à l'heure. Soit l'approche *cognitive* du système **S4** est restaurée et la clause de vérité est alors confirmée, mais le débat entre internalistes et externalistes est relancé au sujet de la signification de K ; soit l'approche *performative* est maintenue, et la clause de vérité est réinterprétée en termes de véracité de type 'si je dis que je sais que ϕ, alors je considère que ϕ est vrai' et la thèse KK en termes de 'si je dis que p est vraie, alors c'est que je le sais'. Notre préférence pour la seconde option désamorce le paradoxe de Fitch et montre simplement que la définition anti-réaliste du savoir, tout comme la définition classique, aboutit à une impasse tant que les concepts de vérité et connaissance sont distingués.

On peut citer deux autres critiques susceptibles d'appuyer notre interprétation pragmatique de K, qui concernent respectivement la clause de vérité et la connaissabilité.

Dans [Lenzen 1978], l'auteur cite un argument de K. D. Irani allant dans notre sens, selon lequel toute connexion logique entre une déclaration de savoir et une vérité factuelle est dépourvue de sens :

---

[12] Pour reprendre la remarque précédente de [van Benthem 2004], notre analyse fait de nous des 'censeurs' : elle revient à dire que le paradoxe de Fitch n'est pas si intéressant que cela, dès lors que l'opérateur K est réinterprété en termes performatifs. Une telle analyse illocutoire n'est pas *ad hoc* et trouvera d'autres avantages par ailleurs.

A la fin de son article, Irani essaie de réfuter [Kp → p] par une troisième ligne d'argumentation : 'La difficulté dans l'examen de [Kp → p] apparaît, parce que ce que nous voulons dire en disant que [X sait p] est vrai n'est pas immédiatement clair. (...) L'énoncé 'X sait p' est une *assertion*, et la logique des assertions n'est pas la logique de l'implication d'énoncés *factuels*'. Ainsi –c'est la conclusion non déclarée d'Irani –'X sait que p' ne peut pas entraîner 'p'.

A quoi Lenzen répond dans la foulée que la clause de vérité peut avoir un sens indépendamment de son assertion :

Je ne connais pas de 'logique des assertions' et je ne pense pas que nous ayons besoin d'une telle logique pour comprendre et évaluer la condition de vérité, qui nous empêche simplement de savoir qu'un énoncé faux est vrai –que l'on affirme une telle chose ou non. La connaissance qu'une personne a de p est un fait indépendant du fait qu'il affirme savoir ; et c'est un fait qui présuppose que p soit vrai. En dépit de toutes les attaques philosophiques, cela reste un truisme du sens commun que bien tout le monde sache que des énoncés faux sont crus parfois comme étant vrais, personne ne doive croire que des énoncés faux puissent être sus comme étant vrais.[13]

Les deux dernières remarques de Lenzen décrivent des connaissances et croyances générales, ou impersonnelles, qui ne portent pas sur notre propre état de connaissance : symbolisées par les formules respectives $K(\neg\phi \wedge B\phi)$ et $\neg B(K\phi \wedge \neg\phi)$, la première remarque de Lenzen évoque la faillibilité de nos croyances (tout le monde peut se tromper), tandis que la seconde décrit l'impossibilité de parler d'une connaissance fausse. C'est un truisme de dire qu'une connaissance ne peut pas être fausse, mais ce n'est pas un truisme en revanche de dire qu'une personne ne peut pas ne pas avoir conscience de ses propres déclarations de connaissance ; la remarque d'Irani n'est effectivement pertinente que si l'on admet initialement une lecture assertive ou performative de 'X sait p'. Lenzen est en droit de maintenir la clause de vérité comme un truisme de sens commun, mais ce truisme ne sauvegardera pas par ailleurs la thèse KK puisqu'il repose sur une interprétation *im*personnelle de la connaissance (*ce* qui est connu est vrai). L'argument d'Irani suppose ainsi une logique des assertions, que Lenzen dit ne pas connaître mais qui sera développée plus loin comme le point d'aboutissement de notre travail (section 3.3.3). Le truisme de sens commun que représente la clause de vérité ne sera pas rejeté en l'espèce, mais il sera présenté dans un autre sens et sous une autre forme logique.

Un autre argument critique porte sur la forme symbolique du paradoxe de Fitch : celle-ci utilise des modalités qui gouvernent des énoncés, vrais ou faux indépendamment de la connaissance des locuteurs. Or manipuler un symbolisme modal de ce genre ne serait pas tout à conforme à l'esprit anti-réaliste, d'après [Hinzen 2000] :

Une remarque critique plus importante est que le fait de traiter des attitudes épistémiques comme des opérateurs propositionnels modaux *explicites* nous engagera dans l'idée que l'aspect épistémique de la logique, d'un intérêt intuitionniste central, ne sera pas reflété dans une notion de contenu propositionnel ou dans une notion de vérité. Elles ne seront pas *intégrées au sein de* la logique. ◇Kp constitue elle-même un morceau d'information propositionnelle doté de conditions de vérité absolues, non épistémiques. Adopter un langage modal avec une sémantique des mondes possibles standard ne nous débarrassera pas d'une conception

[13] [Lenzen 1978], p. 11; nous soulignons.

parfaitement classique de ce qu'est une proposition. Les notions sémantiques épistémiques présupposeront dans leur définition la compréhension de notions non épistémiques.[14]

En d'autres termes, le paradoxe de Fitch ne proviendrait pas de la définition de la vérité comme assertabilité garantie, mais de sa formulation par un symbolisme qui ne rend pas compte du raisonnement intuitionniste en termes de construction. Il s'agirait donc de modifier les règles de *formation* des formules logiques, plutôt que de chercher à modifier les règles de *transformation* des connecteurs logiques. Si l'on suit Hinzen sur ce point, la littérature portée sur le paradoxe de Fitch manquerait donc la véritable cible en s'attaquant aux connecteurs logiques ou aux règles d'inférence : la révision des propriétés de la négation ou des quantificateurs peut certes bloquer la conclusion paradoxale, mais elle ne touche pas à la raison véritable pour laquelle l'anti-réalisme conduit à un paradoxe. Le problème viendrait de la *formulation* même de sa doctrine vérificationniste, d'où un changement nécessaire du langage-objet et des règles de formation des énoncés modaux (des jugements) pour rendre compte de ce qu'est un raisonnement purement anti-réaliste. La distinction entre une révision des règles de formation des formules et une révision des constantes logiques a déjà été évoquée précédemment : les logiques déviantes sont le résultat du second type de révision, tandis que Hinzen plaide pour le premier.

De ce problème de symbolisation logique de l'anti-réalisme, faut-il conclure que vérité et connaissance n'ont pas de relations conceptuelles clairement définies ? Si, comme Hinzen l'a affirmé ci-dessus, le paradoxe de Fitch repose sur un symbolisme trompeur, cela veut dire que le langage-objet de la logique modale épistémique anti-réaliste ne remonte pas "suffisamment en arrière", pour reprendre les termes de Dummett, ou qu'il l'admet l'usage de formules dont la valeur sémantique est prédéfinie. Comment exprimer 'la connaissabilité de ϕ' sans utiliser l'énoncé ϕ, dont les conditions de vérité sont *préétablies* en logique modale ?

Pour sortir de ce débat entre réalisme et anti-réalisme en sémantique, il nous faut repenser les relations établies entre les concepts de connaissance et vérité. A quel moment un énoncé peut-il être considéré comme pourvu de sens et doté d'une valeur sémantique propre ? Le cas du *principe de charité logique* va être employé pour répondre à cette question : les concepts de signification, de vérité et de connaissance constituent trois aspects indissociables dans le processus de *compréhension* des énoncés des langues naturelles. Notre compréhension de ces énoncés est avant tout le résultat d'une *action* produite par la méthode de traduction, laquelle présuppose des conditions de vérité pour chaque énoncé mais travaille à partir de cas d'*énonciation*. Les points de vue opposés mais défendables de Quine et Hintikka peuvent être conciliés, si l'on considère les attitudes épistémiques par un autre biais : une approche *illocutoire* de ces concepts, et c'est ce que nous allons tenter de démontrer dans ce qui suit.

3.2 Logique et rationalité

L'argumentation qui suit est double : illustrer la façon dont les aspects logiques et épistémiques de la vérité se complètent dans le cadre de la philosophie de la logique de Quine, à partir de sa reconstruction behavioriste (en termes de croyance ou d'acceptation) du

[14] [Hinzen 2000], p. 283. L'auteur est un partisan de la théorie des types constructive de Martin-Löf et, pour cette raison, il définit une proposition comme un ensemble de jugements ou de preuves disponibles pour un énoncé.

processus de traduction ; contester l'idée selon laquelle une description plus convaincante de la croyance rationnelle exigerait une révision des constantes logiques. La question porte sur la logique modale épistémique normale, et les raisons de passer d'une logique épistémique classique (enrichie) à une logique épistémique non-classique (déviante) : pour rendre compte de la non-omniscience et de la limitation des ressources déductives des agents épistémiques, notamment, faut-il modifier les propriétés des connecteurs de la logique classique afin d'obtenir une analyse des attitudes épistémiques plus convaincante ?

Par opposition à la tendance 'tolérante' (carnapienne) actuelle en logique philosophique, consistant à modifier librement les opérations logiques pour résoudre des paradoxes, nous poserons une réserve face à cette attitude de déviance, en raison de notre approche prescriptive et non descriptive des règles de logique. Pour quelle raison maintenir les propriétés de la logique modale *normale*, lorsqu'il s'agit de définir des attitudes épistémiques ? Parce que, selon nous, une analyse satisfaisante des attitudes épistémiques exige non seulement une analyse de leur forme logique mais aussi de leur contexte d'usage : il s'agira de montrer dans ce qui suit que les attitudes épistémiques ont un comportement logique normal lorsque leur énonciation a un but assertif (décrire un état de choses), et que le changement d'usage peut être expliqué sans nécessiter une révision des constantes logiques.

Maintenir l'interprétation classique des constantes logiques nécessite avant toute chose la distinction entre les langages formels et leur application aux énoncés des langues naturelles : les conditions de vérité d'une formule sont relatives aux règles du système choisi, dans le premier type de langage (Tarski, Carnap) ; elles sont déterminées par l'usage dans le second type de langage, même si cet usage est parfois trop variable pour être défini clairement (Quine). Or malgré la variation d'usage de constantes logiques telles que la négation ou le conditionnel, nous prétendrons que les conditions de vérité des mots logiques sont marquées par un certain usage prédominant, une sorte de *signification de base* au sein des langues naturelles. Quel est cet usage ? Dans une section précédente, Ellis a proposé de remplacer une sémantique de vérité par une sémantique de croyance, dans laquelle la vérité ou fausseté d'un énoncé était traduite en termes d'acceptation ou de rejet de cet énoncé dans un ensemble de croyances. Pour rejeter la sémantique *véri*conditionnelle, [Ellis 1976] avait défendu une position qu'il assimile à celle de Quine :

> une structure de langage n'est pas encore une langue. En effet, les composants de la structure du langage pourraient être compris de manières diverses. Comment alors définit-on une langue, un système de croyance dans une langue ? L'approche courante serait : définissez d'abord votre langue en donnant une sémantique de vérité pour les phrases de votre structure de langage, puis définissez votre système de croyance en spécifiant les phrases qui sont admises comme vraies, qui sont considérées comme fausses et qui ne sont pas encore décidées de façon décisive. Je rejette cette approche, parce qu'elle présuppose que les significations des connecteurs, opérateurs, termes, prédicats et phrases d'une langue peuvent être spécifiées indépendamment des modèles d'acceptation et de rejet que les locuteurs de la langue affichent. Je pense que Quine a montré de façon convaincante que les questions de signification ne peuvent pas être réglés d'avance aussi nettement.[15]

[15] [Ellis 1975], p. 188.

Dans ce qui va suivre, nous allons montrer que l'approche charitable de Quine, si elle n'attribue pas en effet la vérité ou fausseté d'un énoncé indépendamment du comportement du locuteur, présuppose toutefois le rôle primordial de la vérité et de la fausseté dans l'étude de la signification des énoncés : la question de la signification des connecteurs logiques *est* déjà réglée avant qu'un énoncé soit compris, puisque ce sont ces connecteurs qui permettent d'interpréter les croyances d'autrui. Signification, vérité et croyance sont donc liées dans le processus de compréhension d'un énoncé, et c'est à partir d'une réflexion sur leur agencement que nous allons repenser le sens de la logique modale épistémique.

Le thème majeur de cette section sera le rôle de la logique dans l'exercice de traduction, et la justification corrélative des opérations en logique classique : c'est sur la base de ce qu'un locuteur *considère comme* vrai ou *comme* faux que nous pouvons déchiffrer progressivement le sens de ses énoncés. Ellis et Quine semblent bien plus en accord que le premier ne semble le penser, dès lors que vérité et supposition de vérité ne sont plus considérées comme des distinctions importantes. La vériconditionnalité pourrait être rebaptisée en termes de vérisimiliconditionnalité, cela ne changerait rien au mode de composition des pensées selon Ellis ou Quine : dans les deux cas, la définition des opérations est *compositionnelle*. De plus, la notion de vérité ne semble pas essentielle en logique puisque 'vérité' et 'fausseté' y sont utilisées comme de simples postulations de valeurs, de simples suppositions : le rejet et l'acceptation désignent des comportements ou croyances ostensives servant à interpréter les significations des énoncés d'autrui, mais ce sont des conditions de *vérité* qui expriment ces comportements en termes de 'considérer comme vrai' et de 'considérer comme faux'.[16] Vérité et fausseté, assentiment et dissentiment, acceptation et rejet participent ensemble au processus de compréhension dans un langage, si bien que la relation d'ordre exigée par [Ellis 1976] entre les notions de croyance et de vérité n'a plus vraiment d'importance à ce stade.

La section qui suit concerne la représentation des systèmes de croyance et le rôle de la logique à cet égard. Une autre façon de poser le problème est la suivante : quelle place la logique occupe-t-elle dans une théorie de la rationalité ? Encore faut-il s'entendre sur la notion de *rationalité* elle-même. Globalement, on peut comprendre celle-ci comme un ensemble de règles à suivre en vue d'un comportement justifié ou correct. Mais l'idée de conformité qui sous-tend la rationalité présuppose un modèle auquel se conformer : lequel ? Lorsque l'on parle de rationalité,

> cette notion est le plus couramment comprise en termes de certains 'critères universels', eux-mêmes compris comme des règles qui spécifient ce qui est considéré comme une *bonne* raison de croire quelque chose.[17]

---

[16] Lorsque [Ellis 1975] prétend que "si le système de croyance d'une personne affiche un modèle d'acceptation et de rejet qui viole ces exigences [de cohérence], alors il existe en apparence du moins un argument pour prétendre qu'il ne comprend pas les opérateurs ou connecteurs de la langue" (p. 189), il ne fait que reprendre en d'autres termes (non-vériconditionnels) une position non charitable : si un locuteur considère comme vraie une conjonction dont un membre est faux, c'est qu'il ne comprend pas le sens de la conjonction. La charité consiste au contrarie à présupposer une compréhension ou compétence logique maximale du locuteur.

[17] [French 1990], p. 121 ; nous soulignons. La question qui domine cette section est la suivante : si l'on considère les logiques modales normales comme l'expression de croyances rationnelles, les logiques *non*-normales sont-elles capables de donner une représentation plus adéquate de la rationalité ?

Quelles opérations sont conformes à un raisonnement 'correct' et doivent guider nos *croyances* ? Si la rationalité est l'étude des règles de bonne conduite pour la pensée, les critères retenus par la logique sont notamment la préservation de la *vérité* et la consistance. Mais cela ne signifie pas que les opérations logiques doivent déterminer la totalité des comportements rationnels, comme nous allons nous efforcer de le montrer dans le cadre du principe de charité.

Il semble admis que le comportement rationnel excède le domaine de compétence de la logique : prendre un parapluie lorsqu'il pleut, choisir dans une salle la place la moins exposée aux températures extrêmes, etc. ne trouveront pas leur raison ultime dans une théorie proprement logique, au sens de la logique des relations de conséquence formelle. Induction et abduction constituent deux méthodes de raisonnement qui ne seront pas pris en compte dans ce qui suit. Le problème concernera uniquement la délimitation de la part du logique dans la description d'un comportement rationnel, ainsi que le rôle d'une logique épistémique dans cette description de la rationalité : que dit-elle à ce sujet, et le dit-elle bien ? Pour répondre aux deux questions du début de section, qui concernent l'*épistémologie de la logique* ou l'étude des justifications des raisonnements déductifs, nous présenterons un type d'argument qui propose une *division du travail* entre la logique et les autres disciplines théoriques : dans le cadre de l'hypothèse de traduction radicale, suggérée par Quine en vue de garantir la compréhension de langues parfaitement inconnues, nous maintiendrons que l'occurrence d'un comportement dit 'irrationnel' n'entraîne pas la révision des règles d'inférence logique mais, au contraire, une révision de la traduction de départ proposée par l'interprète. Cette réponse supposera une frontière définie dans le langage entre une catégorie conceptuelle au sens large, d'une part (le lexique des termes *désignatifs*), et une catégorie logique d'autre part (les mots logiques ou *formatifs*, dont la fonction est de combiner des termes désignatifs). Dans quelle mesure cette position rend-elle justice à la logique modale épistémique de [Hintikka 1962], et constitue-t-elle un argument suffisant face au révisionnisme en logique ? Pour traiter des croyances irrationnelles et de la justification des logiques modales normales, nous aborderons notamment le cas des logiques doxastiques paraconsistantes et suivrons la stratégie quinienne de mutilation minimum : sauver ce qui est *obvie*, dont les règles logiques. Nous verrons par la suite si cette défense de l''obvie' résiste à des projets plus récents en théorie de la rationalité limitée : une théorie des limitations de nos capacités cognitives implique-t-elle une révision du cadre logique ?

3.2.1 La charité logique, entre consistance et paraconsistance

Ce qui suit concerne la logique et ses applications aux langues naturelles. Toutes les langues partagent-elles un fond logique commun, ou peut-on admettre la paraconsistance comme la caution de rationalités au pluriel ? Si l'argument des paraconsistants est convaincant, l'argument de Davidson contre les schèmes conceptuels (l'hypothèse théorique de *Weltanschauungen*, ou manières d'organiser conceptuellement le monde qui varieraient selon les foyers linguistiques) s'effondre puisqu'il devient dès lors possible de donner un sens à d'autres rationalités que la nôtre au sein d'un langage formel. Le révisionnisme logique aurait pour conséquence un rejet de la charité logique, précisément. La question de savoir si,

contre ce principe de charité, une révision des constantes logiques s'impose sera utilisée comme un moyen de mettre quelques principes de logique épistémique à l'épreuve, parmi lesquels la distribution modale : $\Box(\phi \wedge \psi) \equiv (\Box\phi \wedge \Box\psi)$. Ce problème sera illustré ensuite (section 3.2.3) avec quelques problèmes connus tels que les paradoxes de la préface, de la loterie ou de la décitation, et nous discuterons la légitimité de ces prétendus paradoxes.

Contre l'idée d'une révision des particules logiques lors d'une traduction, nous défendrons la sémantique classique (bivalente) comme une étape logique privilégiée en vue d'interpréter le comportement humain. Pour le montrer, nous insisterons d'abord sur la condition d'émergence du principe de charité dans l'interprétation des comportements (section 3.2.2), puis nous passerons en revue quelques contre-exemples au principe (3.2.3) avant de tirer des conclusions sur la relation entre langage et théorie dans l'explication de la rationalité (3.2.4).

La question porte ici sur les fondements épistémologiques de la logique, c'est-à-dire la place qu'occupe cette discipline dans l'organisation générale de nos connaissances. La relation entre langage et monde concerne à l'origine la partie sémantique de la logique. Or cette relation sémantique est compliquée par certaines limitations épistémologiques propres à notre rapport au monde. Les zones d'ombre entre notre langage et les théories scientifiques en sont le résultat, mais la logique n'est pas censée être concernée par ce débat si elle est un pur langage. L'est-elle vraiment ?

Il s'agit de discuter le rôle de la logique en théorie de la connaissance ; la relation entre vérité et vérité logique (entre réalité et langage) ; la portée des sémantiques déviantes (philosophie de la logique) et la forme des raisonnements logiques au sein de la pensée conceptuelle. Pour défendre le primat des règles de la logique classique, Quine avait soutenu la doctrine de l'invariance de la signification des constantes logiques (critère d'identité des constantes logiques). Il y aurait beaucoup à dire sur cette approche favorable au conservatisme en logique, puisque l'éclosion récente des logiques substructurelles et le recours au calcul des séquents tend à affiner la définition des procédures logiques et donner tort à la thèse d'invariance de Quine (cf. [Paoli 2003]).

Le but de cette section est de montrer que les logiques non-classiques (non-bivalentes) sont susceptibles de commettre une 'erreur de catégorie', lorsqu'elles servent de logiques philosophiques appliquées à des fragments d'énoncés naturels : elles tentent de se suppléer à d'autres disciplines en expliquant des incohérences apparentes en termes d'inconsistance admissible ; le pluralisme logique qui en découle peut créer une certaine confusion (voire une collusion) entre deux domaines distincts, à savoir la théorie de la *vérité* (relation de conséquence), d'une part, et la théorie de la *connaissance*, d'autre part (justification de la vérité d'énoncés simples). Puisque la logique épistémique se situe à mi-chemin entre vérité et connaissance, ou entre logique et épistémologie, nous nous servirons de son formalisme afin d'exprimer la position charitable de Quine et Davidson, avec en point de mire une critique des arguments paraconsistants favorables à la révision des procédures déductives.

Par opposition à la famille des logiques paraconsistantes, toutes réunies sous une propriété commune qui est la distinction entre inconsistance et trivialité (ou explosion),[18] les partisans

[18] Une contradiction ne fait pas 'exploser' un système paraconsistant et ne rend pas tout démontrable : on n'a pas $(\phi \wedge \neg\phi) \Rightarrow \psi$ pour tout énoncé ϕ et ψ. Les logiques *relevantes* sont paraconsistantes, dans la mesure où elles

de la charité prétendent qu'un certain modèle de déduction s'impose à nous pour comprendre autrui :

> Puisque la charité n'est pas une option, mais une condition pour avoir une théorie qui soit efficiente, il est dépourvu de sens de suggérer que nous puissions tomber dans une erreur massive en la soutenant. (...) La charité s'impose à nous ; – que nous le voulions ou non, si nous voulons comprendre les autres, nous devons les considérer comme ayant raison dans la majeure partie des cas. Si nous pouvons produire une théorie qui réconcilie la charité et les conditions formelles d'une théorie, nous avons fait tout ce qui peut être fait en vue de garantir la communication. Rien de plus n'est possible, et rien de plus n'est exigé.[19]

Dans ce qui suit, nous essaierons de développer les raisons pour lesquelles un seul schème conceptuel s'imposerait à nous *du point de vue logique*, ce qu'il est, ainsi que la conséquence de cette position charitable sur l'aspect logique de la rationalité.

3.2.2 Charité bien ordonnée ...

La charité sert d'argument épistémologique en faveur de la *bivalence*. Bien plus qu'une application abstraite d'un ensemble de formules vers un ensemble binaire de valeurs de vérité, cette bivalence est surtout une règle concrète servant à interpréter les croyances d'autrui au sein d'une méthode de traduction radicale, le 'principe de charité'. Pourquoi interpréter les énoncés dans un cadre bivalent ? Quine, Davidson et les partisans de la logique classique considèrent celui-ci comme la contrepartie sémantique d'une méthode basée sur l'observation des comportements (behaviorisme) d'assentiment (oui/vrai) et de dissentiment (non/faux). Elle provient chez Quine d'une épistémologie d'inspiration behavioriste, motivée entre-temps par plusieurs thèses épistémologiques :

- le rejet des *significations* comme entités objectives dans un langage : exprimée dans [Quine 1961b] par les critiques de la distinction synthétique-analytique et de la notion de synonymie, ce rejet signifie que la signification commune à plusieurs énoncés n'est pas donnée mais construite par un jeu d'interprétation entre les langues. Il n'y a pas de *proposition* garantissant l'identité de signification des énoncés, et il n'y a pas de *fact of the matter* (de faits bruts) pour garantir leur vérité ; autrement dit, l'absence d'entités extra-linguistiques et de faits objectifs implique que, dans un schème conceptuel, les valeurs de vérité des énoncés simples ne sont pas données mais doivent être justifiées sur la base de l'expérience. Le refus des vérités analytiques ne plaide ni plus, ni moins en faveur du relativisme ou de l'universalisme en logique : la doctrine linguistique de la vérité de logique de Carnap, favorable aux vérités analytiques, plaidait dans le sens du relativisme en considérant les règles de conséquence comme des conventions sans aucun import ontologique ; Quine rejette ces vérités conventionnelles, mais il n'en écarte pas moins l'idée d'un import ontologique des vérités logiques. Son rejet de l'atomisme épistémologique donne droit à une révision (en principe)

---

n'autorisent à inférer que les formules déjà contenues dans les prémisses. Nous laissons de côté les nuances entre les diverses branches de la logique paraconsistante (incluant le dialéthisme de Priest). Pour un panorama de la famille paraconsistante, voir notamment J.-Y. Béziau : "What is paraconsistent logic ?", in *Frontiers of paraconsistent logic*, D. Batens et al. (eds.), Research Studies Press, Baldock (2000), pp. 105-117.

[19] [Davidson 1973a], p. 19.

des combinaison possibles entre les 'faits' ; en l'absence de significations objectives, le lien entre logique et ontologie est indistinct et les vérités logiques doivent chercher leur justification du côté de l'épistémologie ;

- l'*inscrutabilité de la référence* : thèse liée à la précédente, l'inscrutabilité signifie que la connexion entre un terme et son référent est indéterminée dans le langage (voir l'exemple de Gavagai et l'hypothèse de la lapinité, in [Quine 1960]), nécessitant une reconstruction behavioriste des relations de désignation ; cet argument sceptique concerne le problème du découpage conceptuel et notre manière d'organiser le champ de l'expérience par des termes et relations du langage. Encore une fois, rejeter l'objectivité de ces relations de significations (pas de désignation rigide entre un terme et l'objet désigné, ici) semble conduire Quine à une relativité des schèmes conceptuels, contrairement à ce que nous constaterons plus loin chez lui et Davidson.
- la *sous-détermination des théories* : deux théories logiquement incompatibles peuvent être empiriquement équivalentes et convenir pour les mêmes expériences ; en d'autres termes, aucune théorie ne prévaut pour décrire correctement notre champ de l'expérience. Chaque énoncé étant considéré comme un mélange indistinct de termes d'observation et de termes théoriques (holisme sémantique), il est possible de modifier ces termes et d'exprimer cependant un même état de choses.[20]
- le rejet du *vérificationnisme* (un des deux 'dogmes de l'empirisme') : liée au holisme sémantique, il signifie la signification d'un énoncé n'est pas vérifiable isolément et réfute la toute distinction nette entre des faits bruts et des concepts théoriques.

Nous avons rappelé ces thèses épistémologiques de Quine afin de donner raison à sa reconstruction behavioriste des énoncés en termes d'*assentiment* et de *dissentiment* ; sans ces précisions doctrinales, la thèse de la charité n'aurait pas lieu d'être ou céderait face au réalisme ou au révisionnisme en logique, deux positions antinomiques que Quine récuse à part égale. Malgré son holisme sémantique et l'abandon de l'hypothèse d'une correspondance entre faits et énoncés, les partisans de la charité maintiennent pourtant que la logique classique est *le* modèle du raisonnement déductif. Considérant que la thèse de la relativité de l'*ontologie* n'entraîne pas une relativité de la *logique* chez Quine, nous poserons deux questions afin de mieux saisir sa position conservatrice en ce domaine : comment la réalité nous est-elle accessible, et comment résister au relativisme en logique sans présupposer que les structures logiques disent quelque chose du monde ?

Quine base sa reconstruction des énoncés vrais sur la base d'énoncés d'observation ou holophrastiques, c'est-à-dire de simples tokens d'énoncés ('eau', 'maman') sans forme prédicative et dont le but est d'indiquer un événement d'observation occasionnel par ostension (ceci est de l'eau, ceci est ma maman). Il s'agit pour lui de procéder à une reconstruction causale des étapes menant de ces énoncés d'observation aux énoncés

[20] Pour exprimer son holisme sémantique et le rejet de la distinction entre faits et théories, Quine déclarait que notre champ de connaissance constitue un "savoir gris pâle, blanc par les faits et noir par les conventions (...) je n'ai pas trouvé de raisons substantielles pour en conclure qu'on y trouve des fils parfaitement noirs, ou des fils parfaitement blancs" ([Quine 1954], p. 406). Pour exprimer sa thèse de la sous-détermination des théories, il évoque également sa notion de 'fonctions déléguantes' (*proxy functions*) : des énoncés théoriques peuvent êtres traduits les uns en termes des autres tout en référant à un même état de choses.

théoriques permanents du langage scientifique. Au cours de ces étapes d'abstraction progressive de la prédication, la relation entre langage et réalité passe par la médiation de ce qui est appelé un 'schème conceptuel'. Suite au rejet des projets réductionnistes tels que la théorie vérificationniste de la signification, l'étude du rapport entre langage et réalité se complique et le but de la charité est de trouver une explication à cette entreprise de structuration du monde par un schème conceptuel :

> Les schèmes conceptuels, nous dit-on, sont des manières d'organiser l'expérience; ce sont des systèmes de catégories qui donnent forme aux données de la sensation ; ce sont des points de vue à partir desquels les individus, les cultures ou les époques contemplent le spectacle en cours. Il se peut qu'il n'y ait pas de traduction d'un schème à un autre, auquel cas les croyances, désirs, espoirs et portions de connaissance qui caractérisent une personne n'ont pas de contrepartie véritable pour celui qui souscrit à un autre schème. La réalité elle-même elle relative à un schème : ce qui est considéré comme réel dans un système peut ne pas l'être dans un autre.[21]

Quel rapport avec notre sujet, lié à la logique et l'épistémologie ? Un rapport de raison à conséquence entre la notion de schème conceptuel et le choix des inférences logiques, c'est-à-dire du comportement logique des agents. Dans un schème conceptuel, nous distinguerons entre plusieurs fonctions attribuées au concept, c'est-à-dire un mot du langage dont la fonction est d'apporter une information. Premièrement, la fonction du concept peut être *désignative* (ou sémantique) lorsqu'il vise à dire quelque chose sur un objet ou une référence observable dans notre champ d'expérience. Parmi ces termes désignatifs, les uns sont référentiels et les autres sont descriptifs : les termes singuliers (noms propres, déictiques, etc.) sont des concepts dont la fonction est *référentielle* ; les termes généraux (adjectifs, descriptions définies[22]) sont *descriptifs*, ce sont également des termes désignatifs dans la mesure où ils disent quelque chose sur le monde en attribuant une qualité à un terme singulier ou référence. Deuxièmement, la fonction du concept peut être *formative* (ou syntaxique), lorsque celui-ci n'a plus pour but de désigner un objet du monde mais de structurer ou combiner des concepts désignatifs entre eux : les mots logiques, ou syncatégorèmes en général, ont cette fonction formative puisqu'ils ne disent rien sur notre champ d'expérience. La thèse d'incommensurabilité de Kuhn portait sur des termes désignatifs, puisqu'il était question de douter des conditions d'identité possibles entre certaines notions de la physique d'une époque à une autre ; ce n'est pas de cette incommensurabilité que nous traiterons ici, mais de celle des concepts formatifs ou connecteurs logiques. L'allusion à des schèmes conceptuels différents implique-t-elle une différence dans l'usage des concepts formatifs, autrement dit : peut-on penser par le biais d'autres logiques ou d'autres manières de combiner des énoncés (syntaxe) pour dire la vérité (sémantique) ? Aux deux premiers modes d'usage désignatif et formatif des concepts, nous en ajouterons un troisième : la fonction *intentionnelle* du concept, selon que le mot est utilisé en vue de décrire un fait du monde

[21] [Davidson 1973a], p. 5.

[22] Les descriptions définies sont une catégorie de termes désignatifs particuliers, partagées selon le contexte d'usage entre une fonction référentielle et une fonction descriptive. Voir à ce sujet la relativité contextuelle de la fonction des descriptions définies, in [Donnellan 1966], que Hintikka a reprise à son compte en termes de connaissance par description ('le maître d'Aristote', pour décrire l'individu Platon) et par accointance ('l'homme avec un verre de martini à la main', pour indiquer une personne située en face du locuteur).

(fonction *locutoire*), d'accomplir un acte de langage (fonction *illocutoire*) ou de produire un effet sur l'interlocuteur (fonction *perlocutoire*). La complexité du processus de traduction réside dans la capacité du traducteur à distinguer dans un acte de discours entre les fonctions désignative, formative ou intentionnelle des mots. Le débat sur la révision des mots logiques, par exemple, consiste à insister sur la variété d'usage d'un seul et même concept formatif, un connecteur logique. Pour les partisans de la charité, les problèmes de traduction ne doivent jamais entraîner une révision de ces concepts formatifs mais, au contraire, trouver leur cause dans la variété d'usage des concepts désignatifs ou intentionnels. Pour notre part, nous insisterons sur la division du travail que nécessitent ces trois aspects de la pensée conceptuelle : il est possible de maintenir un usage stable des termes formatifs, même si une révision de ces termes est toujours permise en principe.

D'après Davidson, la notion de schème conceptuel est infondée aussi bien pour les concepts désignatifs ou formatifs, parce qu'elle repose sur un *troisième dogme* : la distinction entre un schème conceptuel et un contenu empirique, qui suggère que deux schèmes distincts se distinguent par leur manière d'organiser les données empiriques. Or en l'absence de données brutes de ce genre, dont l'absence est justifiée par le holisme sémantique de Quine, l'idée d'un langage déterminant la structure des faits du monde ne trouve aucune garantie et devient hypothèse gratuite. Comment comprendre le rapport langage-monde, et que signifie l'hypothèse de schèmes incommensurables ? La comparaison des catégories de langage à des catégories de pensée donne lieu à une alternative, selon que le langage détermine ou est déterminé par l'objet de discours. Soit un schème linguistique informe (donne forme à) des données brutes par le biais de concepts formatifs qui relient des concepts désignatifs, soit il s'accorde à des données préalables. La relation établie entre les catégories de langage et les objets du monde a pour but de répondre à un problème de traduction pratique, d'attribution de signification au discours d'autrui : comment comparer les observations de deux locuteurs appartenant à des communautés de langage dont les schèmes seraient différents ? A moins de restaurer la distinction entre énoncés analytiques et énoncés synthétiques, l'idée de schème conceptuel relatif pourrait entraîner un *relativisme conceptuel*, voire un relativisme ontologique si la vérité de nos énoncés dépend de notre manière d'organiser le tableau du monde.[23] Les gens pensent-ils ou perçoivent-ils les choses différemment en fonction de leur communauté de langage, conformément à la théorie des linguistes structuralistes Sapir et Whorf ? Si tel est le cas, quels pourraient être les critères d'identité entre différents schèmes conceptuels ? La thèse du relativisme culturel a été mentionnée également pour justifier le relativisme logique : l'appartenance de logiciens à des schèmes différents pourrait expliquer leur manière différente de penser le monde et de l'organiser en termes de relations conceptuelles. Or selon Davidson, la notion même de schème conceptuel repose sur un dogme

[23] Pour des arguments en faveur du 'littéralisme', notion due à [Engel 1989b] et selon laquelle d'autres cultures manifesteraient d'autres formes de pensée, voir notamment V. Bazhanov : "The rationality of Russia and rationality of the West", in: *Rationalitat und Irrationalitat. Beitrage des 23 Internationalen Wittgenstein Symposiums. 13 -19 August 2000. Kirchberg am* Wechsel, vol. 1(2000), pp. 58-63. Pour un argument relativiste de ce genre en faveur de la paraconsistance, voir également Jonardon Ganeri : "Jaina logic and the philosophical basis of pluralism", *History and Philosophy of Logic* **23**(2002), pp. 267-281.

empiriste radical injustifié.[24] Ni l'idée d'organisation des données empiriques par le biais du langage, ni l'idée d'adéquation à des faits bruts n'a de sens. Pourquoi ?

Annonciatrice du principe de charité, l'argumentation de Davidson repose sur une assimilation entre être vrai et être vrai *dans un langage*. Contre l'idée d'un accès direct aux choses et d'une théorie de la vérité correspondance qui en découle, Davidson explique que

> la totalité des évidences sensorielles est ce que nous voulions, à condition que ce soit la totalité des évidences qu'il y ait ; et la totalité des évidences qu'il y a est précisément ce qu'il faut pour rendre *vrai* nos énoncés ou nos théories. Rien, cependant, aucune *chose* [*nothing* (...) *no thing*] ne rend nos énoncés et nos théories vrais : ni l'expérience, ni les irritations de surface, ni le monde, ne peuvent rendre un énoncé vrai. *Que* l'expérience prenne un certain cours, que notre peau soit échauffée ou contusionnée, que l'univers soit fini, ces faits (...) rendent les énoncés et les théories vrais. 'Ma peau est chaude' est vraie si et seulement si ma peau est chaude. Il n'y a aucune référence ici à un fait, un monde, une expérience, ou une portion d'évidences.[25]

En d'autres termes, la vérité sert d'intermédiaire entre la croyance et la signification : ce que je crois du monde est exprimé par des énoncés, mais la vérité est préalable à l'expérience puisque c'est en vertu des propriétés du vrai que nous sommes en mesure de saisir la signification des actes de langage d'autrui et de lui attribuer des croyances. Prenant acte du rôle pré-empirique de la vérité, la convention T de Tarski caractérise cette vérité comme la relation *intra-linguistique* d'adéquation matérielle entre la réalisation d'un fait *s* (exprimée dans le métalangage) et son expression par l'énoncé '*s*' (exprimée dans le langage-objet) :

> d'après la Convention T de Tarski, une théorie de la vérité satisfaisante pour un langage L doit avoir comme conséquence, pour tout énoncé de L, un théorème de forme '*s* est vrai si et seulement si p' [= énoncé T], où 's' est remplacé par une description de *s* et 'p' par *s* lui-même si L est le [français], et par une traduction de *s* dans la langue [française] si L n'est pas le [français]. Celle-ci n'est pas une définition de la vérité, bien évidemment, et ne cache pas le fait qu'il y a une seule définition ou théorie qui s'applique aux langues en général. Néanmoins, la Convention T suggère *sans pouvoir l'établir* une importante caractéristique commune à tous les concepts de vérité spécialisés. Elle y parvient en faisant un usage essentiel de la notion de traduction dans une langue que nous connaissons.[26]

Autrement dit, la vérité ne désigne pas une relation entre un langage et le monde mais une relation entre deux niveaux de langage. Partant de là, la notion de schème conceptuel devient un non-sens dès lors qu'elle prétend décrire la relation entre nos langage et les faits du monde. L'interprétation de la connaissance chez Quine admet pourtant l'idée d'un schème conceptuel, contrairement à Davidson ; mais ce schème n'est pas absolu et admet des révisions dans son contenu. Dans le même registre que Davidson, le rejet des significations objectives,

---

[24] A la critique de l'empirisme radical par Davidson, [Quine 1981] précise toutefois dans son article "On the very idea of a third dogma" (en part. p. 40) qu'il défend lui-même un empirisme non pas pour fonder une théorie de la vérité, ce que reproche Davidson, mais une théorie de la croyance justifiée : l'expérience sert à justifier des affirmations, mais croyance justifiée et vérité restent deux thèmes nettement distingués chez Quine, contrairement à la position antiréaliste de Dummett.

[25] [Davidson 1973a], p. 16.

[26] [Davidson 1973a], p. 17 ; nous soulignons (pour insister sur la distinction entre vérité et justification de la vérité). La vérité est donc une notion intermédiaire entre plusieurs langages distincts, par exemple un langage formel et une langue naturelle telle que le français.

l'inscrutabilité de la référence et le rejet du vérificationnisme amènent Quine à interpréter l'acquisition de la connaissance par induction sur le comportement des sujets locuteurs. Autrement dit, la connaissance s'acquiert faute de mieux à travers des énoncés d'observation (d'abord pris holophrastiquement), puis par le biais d'énoncés progressivement chargés de théorie. L'acquisition et l'usage des énoncés est un processus de nature empirique, mais le contenu de ces énoncés reste foncièrement indéterminé au sens où l'interprétation qui en est faite n'est pas définitive. C'est à ce stade qu'intervient le rôle des inférences logiques dans la construction du schème conceptuel, dont la nature est *intersubjective* et non objective : pour construire des combinaisons d'énoncés et pour saisir l'intention de signification d'un locuteur (sa façon de voir les choses, dirions-nous), la plupart des énoncés théoriques reposent selon Quine sur une méthode de traduction à base de *fonctions de vérité*.[27] Ces fonctions abstraites sont essentielles pour déterminer les conditions de vérité d'un énoncé complexe, et pour la raison suivante. Tout d'abord, le behaviorisme consiste à assigner une signification à un mot ou énoncé en fonction du comportement ostensible, par empathie et en vertu de stimulations sensorielles (les 'significations-stimuli' dont use Quine pour naturaliser la signification des énoncés d'observation) ; or puisque les stimulations sensorielles sont d'autant moins vives lorsque les énoncés portent sur des sujets d'autant plus théoriques, la part d'observation s'amenuise lorsque le locuteur utilise des termes de plus en plus généraux et abstraits.

Pour saisir dans une telle situation la relation entre la croyance d'un locuteur et le sens de son discours, Quine propose de traduire ce que le locuteur considère comme vrai en termes d'attitudes d'assentiment et de dissentiment. Ces deux réactions observables indiquent que le locuteur tient l'énoncé pour vrai ou pour faux, respectivement (accepter ϕ, c'est croire que ϕ est vrai).[28] La bivalence se manifeste dans l'opposition symétrique entre les paires vrai/faux et assentiment/dissentiment. Quant à la théorie des fonctions de vérité, elle intervient en guise de processus d'interprétation des croyances portant sur des énoncés complexes : la signification des énoncés est devinée à partir de cette base comportementale ainsi que de quelques règles de traduction logique, c'est-à-dire les conditions de vérité des connecteurs, lorsque les énoncés sont complexes. La bivalence est le cadre sémantique dans lequel sont déterminées les fonctions de vérité, dont le but est de combiner des énoncés simples. Ces fonctions reconstruisent les langues naturelles sur le modèle d'un langage formel, et une telle reconstruction s'effectue sur la base du comportement et des stimulations ostensibles des locuteurs au sein d'un environnement ; avec l'indice comportemental, la notion métalogique de bivalence trouve donc une application d'ordre épistémologique.

Pour insister sur le lien entre logique et épistémologie, notons également que le rejet de la théorie de la vérité correspondance naïve (postulant un lien direct entre des propositions de la

---

[27] D'après Davidson, la *compositionalité*, dont la théorie des fonctions de vérité est une application particulière, participe à la traduction : "Une théorie satisfaisante en vue d'interpréter les déclarations d'une langue, y compris la nôtre, révélera une structure sémantique significative : l'interprétation des déclarations d'énoncés complexes dépendra systématiquement de l'interprétation de déclarations d'énoncés plus simples, par exemple". ([Davidson 1973b], p. 317).

[28] La traduction des croyances en termes d'assentiment et de dissentiment implique une théorie *déflationniste* ou *redondance* de la vérité : dire que ϕ, c'est accepter la vérité de ϕ, mais ce détail n'apporte rien au contenu de l'énoncé. Le rôle subalterne de la vérité entre les croyances et les significations des énoncés justifie en partie cet aspect trivial de la notion, mais une telle réduction peut poser des problèmes par ailleurs. Pour une discussion de la théorie déflationniste de la vérité, voir notamment Pascal Engel : *Truth* (2002), en part. Chapitre 2.

logique et des faits du monde) entraîne un rapprochement des notions de vérité et de croyance : l'absence de faits objectifs donnés oblige l'interprétant à se reporter sur les informations obtenues à partir du comportement observable de l'interprété, de sorte que le fait d'*être vrai* est réductible à celui de *tenir* un énoncé *pour vrais*. Parmi les énoncés interprétés, les fonctions de vérité tiennent un rôle à part et justifient la place privilégiée des vérités logiques au sein du schème conceptuel :

> (...) en référence à l'assentiment et au dissentiment, nous pouvons établir des critères sémantiques pour les fonctions de vérité ; autrement dit, des critères pour déterminer si un idiome indigène donné peut être interprété comme exprimant la fonction de vérité en question. Le critère sémantique de la *négation* est qu'il transforme tout énoncé court auquel on donnera notre assentiment en un énoncé auquel on ne le donnera pas, et vice versa. Celui de la *conjonction* est qu'il produit des complexes auxquels (tant que les énoncés complexes sont courts) on est enclin à donner son assentiment lorsque et seulement lorsque l'on est enclin à donner son assentiment à chaque membre. Celui de la *disjonction* est similaire, avec l'assentiment changé à deux reprises en dissentiment.[29]

Le problème reste celui de savoir si la traduction comportementale des énoncés d'autrui en termes de fonctions de vérité est suffisamment fiable ; c'est à ce stade que le principe de charité intervient, puisqu'il consiste à garantir la rationalité des interprétés sur la base d'un schème conceptuel, celui du traducteur. En s'instituant comme modèle de rationalité par défaut, celui-ci suggère donc un accord maximal entre les locuteurs en ce qui concerne leur usage des fonctions de vérité. La Convention T (de structure booléenne) est censée garantir un tel accord. Malgré cela, Quine rappelle que tout énoncé est révisable et peut passer de la vérité à la fausseté, si besoin est ; les énoncés logiquement vrais sont concernés au même titre par le rejet de la distinction analytique-synthétique et, pour cette raison, le succès de la traduction est toujours relatif à un choix d'interprétation. Pour éviter un changement constant dans l'interprétation théorique des énoncés d'un langage, le logicien peut se doter toutefois de règles stables dans le fonctionnement de ses connecteurs à partir du principe de *mutilation minimum* : en cas de désaccord avéré entre la signification d'un concept formatif (mot logique) et son application à un énoncé indigène particulier, le problème peut être surmonté en faisant appel à des *hypothèses analytiques*. Pour rendre compte d'un tel désaccord au sujet des croyances du locuteur, l'interprétant ajoute alors des hypothèses afin de restaurer l'accord au sujet des énoncés d'observation impliqués. Comment se manifeste un tel désaccord en ce qui concerne les mots logiques ? Pour conserver le caractère vérifonctionnel des vérités logiques et ne les modifier qu'en dernière instance, Quine invoque comme hypothèse analytique l'argument de l'*erreur de traduction* : toute interprétation qui contesterait la signification d'un mot logique doit être considérée comme erronée plutôt que propice à la

[29] [Quine 1960], pp. 57-58. L'auteur restitue ici les conditions de vérité des opérateurs booléens (classiques) en termes de conditions d'assentiment et de dissentiment : donner son assentiment à un énoncé ϕ, c'est refuser sa négation : $v(\neg\phi) = 0$ ssi $v(\phi) = 1$; donner son assentiment à la conjonction $\phi \wedge \psi$, c'est donner son assentiment à la fois à ϕ et à ψ, soit $v(\phi \wedge \psi) = 1$ ssi $v(\phi)$ et $v(\psi) = 1$; exprimer son dissentiment vis-à-vis de la disjonction $\phi \vee \psi$, c'est exprimer son dissentiment vis-à-vis de ϕ et de ψ, soit $v(\phi \vee \psi) = 0$ ssi $v(\phi)$ et $v(\psi) = 0$. A noter que la définition ci-dessus de la négation serait acceptée par l'intuitionniste, puisqu'elle correspond à la double négation $\phi \rightarrow \neg\neg\phi$ et non à sa converse $\neg\neg\phi \rightarrow \phi$: $v(\neg\phi) = 0$ ssi $v(\phi) = 1$.

révision du mot logique. Il est préférable de changer l'application d'une théorie plutôt que la théorie elle-même, ce que [Quine 1975] exprime par sa formule : 'sauver ce qui est obvie'.[30]

Concernant le rôle de la logique (et sa sémantique formelle) au sein de ce débat sur la rationalité des locuteurs, nous venons d'observer qu'elle sert à encoder les croyances d'autrui en termes de conditions de vérité supposées admises par tout le monde (ou presque). Aucun accord absolu ne peut être garanti au sein d'une analyse comportementale du discours, donc faillible, mais les vérités logiques sont censées être admises par un maximum de locuteurs. La charité logique concerne ainsi la plus générale des théories, portant sur la totalité des faits ; la 'théorie' logique est la plus générale et ne porte sur aucun objet en particulier, elle concerne l'inférence et son expression sémantique en termes de conséquence et de conditions de vérité. Pour cette raison, réviser les règles logiques en cas de désaccord sur la traduction reviendrait à 'mettre la charrue avant les bœufs', dans la mesure où la charité logique est employée comme une condition *préalable* à toute attribution de rationalité :

> La recommandation méthodologique consistant à interpréter de façon à optimiser l'accord ne devrait pas être conçue comme reposant sur une hypothèse charitable qui pourrait s'avérer être fausse, en ce qui concerne l'intelligence humaine. Si nous ne pouvons pas trouver de moyen d'interpréter les énonciations et autres comportements d'une créature comme révélant un ensemble de croyances largement cohérentes et vraies en vertu de nos propres standards, nous n'avons aucune raison de considérer cette créature comme rationnelle, comme ayant des croyances, ou comme disant quoi que ce soit.[31]

Mais la charité connaît quelques difficultés d'application, car certains types de phrases semblent être en conflit avec la méthode d'interprétation en termes de fonctions de vérité : les énoncés modaux, parmi lesquelles les *attitudes propositionnelles*, les conditionnels irréels, etc., d'où une application discutable et discutée de la charité logique au sein des langues naturelles. Par ailleurs, l'interprétation behavioriste néglige cet autre type de comportement intermédiaire qu'est l'*abstention*. Si croire ϕ, c'est croire que ϕ est vrai, peut-on dire également que ne pas croire que ϕ revient à croire que ϕ est faux : $\neg B\phi \rightarrow B\neg\phi$? Dans un cadre bivalent, toute alternative entre le vrai et le faux est exclue ; mais dans le cas d'une interprétation des comportements d'autrui, un élargissement de ce cadre semble de rigueur. A certains endroits, Quine semble se faire lui-même le porte-parole d'un cadre *trivalent*, insérant l'abstention entre l'assentiment et le dissentiment. Pour exprimer le fait que les fonctions de vérité semblent insuffisantes en vue de déchiffrer les croyances de l'interprété, l'Américain invoque le cas des *fonctions verdictives*, comparables à la valeur de vérité tierce ½ de Łukasiewicz et Kleene :

---

[30] Considérant qu'aucune traduction n'est garantie absolument et que l'erreur est toujours possible, est dit 'obvie' [*obvious*] ce qui prête le plus à l'assentiment des locuteurs ; la traduction 'obvie' a été sans doute préférée à celle d''évidence' pour éviter toute connotation psychologique ou internaliste : ce qui est obvie est déterminé par des critères de reconnaissance publics, nullement introspectifs. Or les vérités logiques font partie selon lui des vérités les plus obvies : "Il nous échoit, en interprétant une langue étrangère, de faire en sorte que les énoncés obvies de cette langue se transforment en des énoncés français qui soient vrais et de préférence obvies eux aussi. C'est le point que nous avons noté. Or cette règle 'Sauver ce qui est obvie' suffit pour établir, *en ce qui concerne leurs conditions de vérité au moins*, nos traductions de quelques-uns des énoncés d'à peu près chaque petite branche de la connaissance ou du discours (...) à cet égard la logique est à part : toute vérité logique est obvie, en fait ou potentiellement." ([Quine 1975], pp. 122-3 ; nous soulignons).

[31] [Davidson 1973a], p. 324.

> Une opposition apparaît ainsi entre les fonctions de vérité et quelque chose de plus, les fonctions verdictives. La logique des verdicts est une logique trivalente, les trois verdicts étant l'assentiment, le dissentiment, et l'abstention (...) La négation est à la fois une fonction verdictive et une fonction de vérité ... La conjonction, en revanche, est une fonction de vérité qui n'a pas tout à fait droit au titre de fonction verdictive. Sa table de verdict est incomplète, comme le montre ce qui suit.

p \ q	assentiment	abstention	dissentiment
assentiment	assentiment	abstention	dissentiment
abstention	abstention	?	dissentiment
dissentiment	dissentiment	dissentiment	dissentiment[32]

Faut-il déduire de ce constat qu'une interprétation comportementale des termes formatifs plaide en faveur de la logique intuitionniste ? Après tout, l'abstention est un comportement analogue à l'absence de preuve chez les intuitionnistes (la non-prouvabilité $\neg B$ dans [Gödel 1933]) et, par conséquent, susceptible d'invalider le tiers exclu. Néanmoins, la bivalence peut être pensée et admise *sans* le tiers exclu, dans la mesure où l'abstention n'exprime pas un comportement ostensible de la part de l'interprété. La négation que Quine a définie plus haut consistait à déduire le *dissentiment* vis-à-vis de ϕ sur la base d'un comportement ostensible, à savoir l'*assentiment* vis-à-vis de ϕ, et non à partir d'une *abstention* au sujet de ϕ ou $\neg\phi$. Pour le dire autrement, l'état d'abstention n'est pas définissable au sein d'une interprétation comportementale des croyances. En effet, ce n'est pas l'abstention qui permet d'attribuer le dissentiment, sur la base de l'inférence $\neg B\phi \rightarrow B\neg\phi$; au contraire, c'est le dissentiment qui permet de déduire une certaine abstention sur la base de l'inférence inverse $B\neg\phi \rightarrow \neg B\phi$. Cette relation logique entre dissentiment et abstention, entre croyance négative et non-croyance se manifeste notamment dans le carré logique des modalités : elle correspond à la relation de subalternation entre le côté E de la négative universelle et le côté O de la particulière négative, montrant de ce fait qu'une certaine structure logique s'exprime à travers notre interprétation des croyances. L'abstention n'est donc pas un antécédent sur laquelle le traducteur peut s'appuyer et, dans la mesure où toute interprétation doit s'effectuer sur la base d'un comportement ostensible, la bivalence prévaut dans tout processus d'interprétation : en l'absence de signes d'assentiment ou de dissentiment, de stimulations perceptibles ou de tout autre signe positif d'engagement de l'interprété vis-à-vis de l'affirmation d'un énoncé, il n'y a tout simplement pas d'interprétation : de l'état d'abstention $\neg B$, rien ne peut être inféré. La bivalence comportementale ne prétend pas à l'applicabilité universelle mais sert simplement d'outil d'interprétation minimal, dans le cadre de l'épistémologie naturalisée de Quine. Pour cette raison, nous ne pensons pas que le cas de l'abstention plaide en faveur d'une révision des foncteurs de la négation ou de la disjonction, notamment : l'abstention incite à ne pas interpréter du tout, plutôt qu'à réinterpréter les mots logiques.

[32] [Quine 1974], p. 77. La zone indéterminée (marquée '?' et qui correspond à la valuation ½ des matrices trivalentes) apparaît également pour le cas de la disjonction : celle où le locuteur s'abstient sur les deux membres disjoints. Les cas de la conjonction et de la disjonction donnent les mêmes résultats ici que les tables trivalentes de Łukasiewicz : $v(\phi \vee \neg\phi) = v(\neg(\phi \wedge \neg\phi)) = ½$ lorsque $v(\phi) = ½$. La notion de 'verdictif' apparaît également dans [Austin 1962], où elle désigne la classe des actes d'assertion.

Le problème d'application des fonctions de vérité tient finalement à un certain désaccord sur les conditions de vérité des mots logiques et sur l'ordre des inférences applicables. Seul le sens 'ostensible' de la double négation a été admis ci-dessus : $\phi \rightarrow \neg\neg\phi$, interprétable doxastiquement comme l'axiome du système modal **D** : $B\phi \rightarrow \neg B\neg\phi$. Le rejet de l'autre sens de la double négation comportementale : $\neg\neg\phi \rightarrow \phi$, ou $\neg B\neg\phi \rightarrow B\phi$, montre une parfaite adéquation entre les oppositions en logique modale et le processus de traduction radicale ; la conséquence est effectivement un affaiblissement du tiers exclu, au sujet duquel Quine concède aux intuitionnistes qu'il n'incarne que de façon 'partiellement obvie' le comportement des mots logiques 'ou' et 'ne ... pas' :

> nous apprenons en effet qu'une disjonction est impliquée par ses membres, avec le même apprentissage du mot 'ou'; et tout cela est très bien, car c'est une loi logique que les intuitionnistes ne contestent pas. Je suggère que la loi du tiers exclu, laquelle ils contestent bel et bien, n'est pas liée de façon similaire à l'apprentissage même de 'ou' et 'ne ... pas' ; elle se situe plutôt dans la zone trouble de la disjonction.[33]

Malgré de possibles divergences et nuances dans l'interprétation des connecteurs à partir du comportement observable, un certain modèle prévaut chez Quine pour assigner des croyances aux interprétés et ce modèle est la logique classique, de préférence à toute imputation d'irrationalité ou de rationalité logiquement différente chez l'interprété. Mais comment justifier clairement l'application de cette logique classique ? Le passage ci-dessus a signalé qu'une interprétation des mots logiques suppose un cadre bivalent, et que ce cadre ne tient pas compte du cas de l'abstention. Mais puisque nous avons considéré l'abstention comme n'ayant aucun effet pratique sur la vérifonctionnalité lors d'un processus de traduction comportementale, la difficulté se situe ailleurs pour la défense des procédures de la logique classique : il existe en effet des problèmes d'interprétation pour les croyances d'autrui au sein d'une même langue naturelle, et la logique épistémique l'a illustrée plus tôt avec le cas de l'*opacité référentielle*. L'irrationalité, d'un côté, la croyance fausse de l'autre sont-elles des cas de figure injustement ignorés par le partisan de la charité ? Au contraire, nous allons montrer que l'attitude charitable ne consiste pas à les ignorer mais à admettre au préalable l'imputation de croyances consistantes et vraies, comme une première étape en vue de la construction d'une théorie de la rationalité.

3.2.3 ... commence par soi-même

Le cas des croyances fausses *at home* (lorsque l'interprété et l'interprétant font partie de la même communauté linguistique) montre que de possibles exceptions parmi les lois logiques soi-disant 'obvies' apparaissent fréquemment. Non seulement de récentes logiques non-classiques indiquent qu'un système formel peut être informatif, donc rationnel, sans respecter certaines règles d'inférence *classiques* : si en effet, comme nous l'avons admis plus tôt, la rationalité est synonyme de gestion réglée d'informations en vue d'une prise de décision,

---

[33] [Quine 1974], p. 80. L'auteur suggère ici que la loi d'introduction de la disjonction : '$\phi \Rightarrow \phi \vee \psi$' est *plus* obvie que le théorème du tiers exclu '$\phi \vee \neg\phi$', parce que celui-ci présupposerait un engagement sur la vérité ou fausseté de ses membres en vertu de la signification de '$\vee$'. On rappellera toutefois l'interprétation plus conciliante du tiers exclu par [Wiredu 1975] (voir note 2) : un énoncé ϕ ou sa contradictoire $\neg\phi$ est vrai *si* l'on est capable de déterminer sa valeur de vérité.

alors la *paraconsistance* semble être l'argument le plus redoutable contre la charité puisqu'elle prétend élargir notre champ d'action raisonné.[34] Mais de plus, notre propre communauté de langage laisse supposer que la charité logique n'est pas soutenable dans la mesure où tous les logiciens ne souscrivent pas à la théorie des fonctions de vérité, parmi lesquels les pluralistes non-classiques.[35]

Selon Quine, il semble impossible d'appliquer la méthode de traduction radicale, sur la base du comportement ostensif des locuteurs, par d'autres moyens que les fonctions de vérité :

> De ce que nous considérons comme une logique, la partie vérifonctionnelle est la seule partie dont nous sommes en mesure d'identifier la récognition à des critères comportementaux.[36]

Dans le cas de la relation conditionnelle, une observation des comportements exigerait que le locuteur donne toujours son assentiment à ψ après l'avoir donné pour ϕ, lorsque $\phi \rightarrow \psi$ est la traduction visée. Mais comment distinguer le conditionnel du bi-conditionnel, sur la base de ces comportements ? La réponse de Quine est double, visant à maintenir le sens classique du conditionnel tout en tenant l'implication stricte à l'écart du langage-objet (le langage-cible de la traduction). Soit l'on interprète la conditionnalité en termes d'inclusion du conséquent dans l'antécédent, parallèlement à l'interprétation ensembliste du connecteur : $\phi \rightarrow \psi$ si et seulement si tout ce l'assentiment de ϕ implique toujours l'assentiment de ψ, observant cette inclusion constante par *induction* sur les comportements observés et sans jamais faire intervenir de termes explicatifs tels que l'analyticité ou la synonymie. Soit cette fonction de vérité est considérée comme impossible à distinguer de la conjonction ou du bi-conditionnel et, dans ce cas, on applique la méthode vérifonctionnelle de définition du conditionnel en termes de négation et de conjonction : "si nous le concevons au sens matériel 'Non(p et non q)".[37] Astreindre toute traduction au régime des fonctions de vérité implique que l'on laisse de côté des notions telles que l'implication stricte ou les conditionnels irréels : la relation *causale* entre l'immersion d'un sucre dans l'eau et sa dissolution ne serait déchiffrable par aucun comportement ostensible, sinon par une relation de concomitance fréquemment observée entre deux événements. Quine a une réponse, radicale, face à cette restriction indésirable des moyens de traduction : la notion de causalité est évacuée du langage scientifique sur lequel repose sa notation canonique. Mais on pourrait reprocher à cette réponse de centrer la méthode de traduction sur une conception positiviste du monde qui ne tient pas compte des représentations causales d'autres foyers linguistiques. En ce sens, le rejet des modalités dans [Quine 1993] n'est pas une réponse à la causalité mais un moyen d'éluder le problème de sa traduction. Il constitue dans le même temps une forme de défection à l'égard de la sémantique

---

[34] La logique paraconsistante ne cherche pas à justifier l'hypothèse de mentalités pré-logiques (Lévy-Bruhl), considérée tout d'abord comme une preuve d'*irr*ationalité des indigènes, mais à défendre l'existence de forme de rationalités différentes (d'autres logiques) ; il s'agit de montrer qu'un critère métalogique tel que la *consistance* (si S $\vdash \phi$, alors on n'a pas S $\vdash \neg\phi$) n'est pas une condition nécessaire à la rationalité, contrairement au 'dogme' légué par la logique aristotélicienne et qui constitue la cible du dialéthiste Graham Priest.

[35] Ainsi [Lokhorst 1998] : "Les anthropologues culturels prétendent d'habitude qu'il n'y a pas de preuve de l'existence de cultures adhérant à des logiques différentes de la nôtre. Mais je trouve que c'est là une déclaration étrange. Tout d'abord, même dans mon pays il y a un groupe de gens qui essaient d'adhérer à la logique intuitionniste plutôt qu'à la logique classique". (p. 57).

[36] [Quine 1960], p. 61.

[37] [Quine 1960], p. 65.

des mondes possibles puisque, dans l'exemple du sucre ci-dessus, un traitement relationnel reviendrait à dire que le sucre fond dans tous les mondes possibles où le sucre est immergé dans l'eau. Cette approche apporte-t-elle une condition suffisante à la traduction de la relation causale ? Nécessaire, tout au plus : l'immersion implique la fonte dans tous les cas, mais rien ne nous prouve que la réciproque sera vraie. Rien ne certifie que la réciproque soit toujours valable, car nous ne pouvons pas parcourir la totalité des situations possibles pour vérifier qu'un sucre immergé dans l'eau doit fondre. Est-il impossible qu'un sucre immergé dans l'eau reste solide dans un monde possible (un jour, quelque part) ? Nous sommes dans l'impossibilité de répondre à cette question inductive, et ne pouvons répondre affirmativement qu'à la question inverse : une relation causale est-elle vraie dans toute situation possible ? Mais la relation de causalité est déjà établie, dans ce cas de figure. En somme, concevoir la relation de causalité comme un conditionnel toujours vérifié équivaut au fait de réduire l'intension à une extension multiple, ou la déduction à l'induction. Comment s'assurer que le sucre fondera toujours après immersion dans l'eau ? Répondre à cette question, c'est répondre à l'hypothèse de Hume que le soleil puisse ne pas se lever demain ; nous n'avons pas de réponse autre qu'inductive à cette question et, de ce fait, le défi de Hume affecte la traduction des énoncés causaux.

Faut-il voir dans cette carence de la traduction une contre-preuve manifeste à la politique de charité, que ce soit en faveur d'un élargissement des mots logiques ou d'une révision de ceux déjà présents ? Non : d'après nous, cela indique seulement que la charité n'épuise pas l'étude de la rationalité et doit être complétée par d'autres méthodes d'interprétation des croyances. Quant aux logiques déviantes ou affaiblies (non-bivalentes en général), nous prétendrons que leur application est acceptable mais dans des contextes autres que la traduction en termes d'assentiment et dissentiment, foncièrement bivalente ; la logique quantique ou la logique linéaire, par exemple, justifie partiellement une révision de la conjonction mais dans un contexte tout à fait différent des attributions de croyance ordinaires. En d'autres termes, une révision des mots logiques est faisable mais non souhaitable dans un contexte de traduction, conformément à la politique de mutilation minimum de Quine. Mais cela reste à démontrer, dans la mesure où certaines fonctions de vérité telles que le conditionnel ci-dessus sont sujettes à caution.

Une étude du cas des croyances 'irrationnelles' (illogiques) *locales* –lorsque l'interprété et l'interprétant appartiennent à la même langue source– peut aider à compléter l'interprétation logique par des hypothèses issues d'autres disciplines. Trois exemples de 'croyances irrationnelles' semblent compromettre les interprétations charitables en termes de fonctions de vérité liées à la contradiction : l'*inconsistance*, l'*opacité référentielle*, et la *duperie de soi*. Ces cas de figure nécessitent-ils une révision des connecteurs logiques, ou peuvent-ils être résolus par d'autres moyens ? Nous plaiderons en faveur de cette dernière approche, charitable.

La relation logique entre des croyances oppose logiciens classiques et non-classiques, selon qu'ils admettent des formes de contradictions plus ou moins explicites. On peut dénombrer trois expressions contradictoires distinctes dans une logique de croyance, dont l'acceptation est au cœur de notre débat présent :

(12) $B(\phi \wedge \neg\phi)$,

(13) $B\phi \wedge B\neg\phi$,

et

(14) $B\phi \wedge \neg B\phi$.

Pour le partisan de la charité, les relations logiques (12), (13) et (14) sont toutes inacceptables parce qu'elles expriment des ensembles de croyances irrationnels. D'une part, il est impossible de tenir des croyances portant sur des paires d'énoncés contradictoires, en vertu du principe de non-contradiction. D'autre part, l'opérateur modal B se distribue sur la conjonction et (12) implique donc (13). D'après la règle tacite de *décitation* admise par le traducteur charitable, dire 'ϕ' c'est croire que ϕ est vraie ; or en vertu de la condition pragmatique de sincérité évoquée dans le cas du paradoxe de Moore, dire 'ϕ et $\neg\phi$' c'est croire à la fois la vérité de ϕ et de $\neg\phi$, c'est-à-dire la vérité de ϕ et à la fausseté de ϕ (en vertu de la définition bivalente de la négation). Admettre (13) revient à admettre l'assentiment et le dissentiment simultanément, mais cela va à l'encontre du principe fondamental (bivalent) de la traduction.

Il existe toutefois des objections d'ordre logique à la charité, selon lesquelles les formes (12), (13) et (14) ci-dessus ne seraient pas toutes aussi fausses l'une que l'autre. En résumé, la stratégie de ces logiciens non charitables (non-classiques) consiste à élargir le champ des expressions rationnelles en rejetant l'une ou l'autre des règles logiques qui sous-tendent la charité et en montrant qu'elles ne sont pas équivalentes. Par un jeu de combinatoire, il s'agit d'assouplir les normes du discours rationnel en éliminant l'une ou l'autre des règles d'inférence utilisées par des auteurs tels que Quine.[38] Dans (12), nous avons une modalité dont le contenu est contradictoire ; dans (13), nous avons deux modalités dont les contenus se contredisent ; dans (14), ce sont les modalités elles-mêmes qui se contredisent. En vertu de la dualité des opérateurs modaux, puisque $B\neg\phi \rightarrow \neg B\phi$ nous en concluons que (13) $\rightarrow$ (14) ; or des logiciens tels que [da Costa & French 1989] valident (12) mais rejettent (14) : $\neg(B\phi \wedge \neg B\phi)$ est satisfiable, donc (13) n'implique pas (14). [Williams 1981] rejette par ailleurs la relation d'équivalence (12) $\equiv$ (13) car (13) n'implique pas (12), selon lui. Une forme de contradiction modérée est introduite par Williams, qui distingue *inconsistance* et *contradiction* : il est possible de croire ϕ et de croire $\neg\phi$ séparément, donc (13) est satisfiable, c'est-à-dire vraie dans au moins un modèle ; mais il n'est pas possible de croire ϕ et $\neg\phi$ conjointement, donc (12) n'est pas satisfiable, c'est-à-dire fausse dans tout modèle. Admettre (13) sans (12) revient à rejeter la règle d'introduction de la conjonction, ou *adjonction* : $\phi, \psi \vdash \phi \wedge \psi$. En d'autres termes, Williams rejette la règle que nous pourrions qualifier de règle de *factorisation modale* : $(B\phi \wedge B\psi) \rightarrow B(\phi \wedge \psi)$, ce qui lui permet d'admettre des cas de croyances inconsistantes tout en excluant les cas de croyances contradictoires.[39] (13) et (14)

[38] Ce jeu de combinatoire apparaît clairement dans l'article de J.-Y. Béziau : "Are paraconsistent negations really negations", in *Paraconsistency: The Logical Way to the Inconsistent* (Proceedings of the World Congress held in São Paulo), pp. 465-486, dans lequel les différentes catégories de logiques paraconsistantes sont classées selon la ou les règles déductives qu'elles excluent (substitution, adjonction, résiduation, etc.). La procédure est similaire aux logiques substructurelles, où les systèmes logiques se définissent négativement par rapport à l'opération séquentielle qu'elles rejettent.

[39] "Si un individu croit que s_1, croit que s_2 et croit que $\neg(s_1 \wedge s_2)$, alors il ne tient pas des croyances *contradictoires*, ce qu'il ferait s'il croyait à la fois que $(s_1 \wedge s_2)$ et que $\neg(s_1 \wedge s_2)$. Il ne tient pas non plus une croyance auto-contradictoire [*self-contradictory*], ce qu'il ferait s'il croyait que $[(s_1 \wedge s_2) \wedge \neg(s_1 \wedge s_2)]$. (...) la seule raison que l'on puisse avoir de supposer que l'inconsistance est la même chose que la contradiction, c'est l'idée, consciente ou inconsciente, que *tenir une conjonction de croyances c'est croire une conjonction*" ([Williams 1981], p. 600) ; nous soulignons. La règle d'adjonction est la cible de cet article.

sont des cas de négation préposée ou postposée à l'opérateur modal et ont déjà été évoqués dans les deux versions du paradoxe de Moore. Mais alors que la solution de l'indéfendabilité doxastique de Hintikka consistait à dénoncer l'aspect contradictoire, *donc* irrationnel (indéfendable) de leur énonciation, les analyses non-classiques qui viennent d'être citées supposent au contraire qu'il est possible de tenir pour vrais des énoncés contradictoires selon leur forme d'expression : (12) serait impossible à tenir mais (13) serait admissible, pour Williams, ce qui invalide la règle de distribution des modalités *normales* : $\Box(\phi \wedge \psi) \equiv (\Box\phi \wedge \Box\psi)$. Le différend porte sur la clôture modale sous conjonction du système **K**, donc la question est : des croyances contradictoires peuvent-elles être acceptées ou non selon le degré de complexité d'une formule conjonctive, par analogie avec l'argument précédent des tautologies de surface de Hintikka ? Si tel est le cas, alors l'assentiment donné à plusieurs croyances contradictoires ne signifierait pas l'assentiment donné à une formule contradictoire unique ; il faudrait plutôt y voir l'incapacité du locuteur à percevoir l'ensemble de croyances contradictoires à laquelle il souscrit, mais *à son insu*. Il ne serait pas question d'un comportement irrationnel ou d'une autre forme de rationalité de type paraconsistant, dans ce cas, mais d'un cas symptomatique d'acuité logique limitée. Autrement dit, admettre une inconsistance n'obligerait pas à réviser les règles logiques elles-mêmes mais inciterait plutôt à comprendre pourquoi l'agent ne les respecte pas à la lettre.

En revanche, la position de [da Costa & French 1989] semble plus radicale puisqu'elle admet que (12) puisse être vraie, indépendamment de la règle d'adjonction des croyances. Mais avant d'en arriver à l'admission de croyances contradictoires, les arguments de [French 1990] en faveur de (12) reposent par exemple sur une distinction préalable entre deux acceptions de la vérité : la vérité *correspondance* ou 'factuelle', où l'énoncé est vrai s'il exprime un état de choses monde ; la vérité *pragmatique* ou 'représentationnelle', où l'énoncé est dit 'vrai' s'il appartient à un ensemble d'affirmations d'une théorie. On retrouve cette distinction chez plusieurs logiciens paraconsistants : [French 1990] mais également [Priest 1985], où il est question de théories scientifiques mutuellement incompatibles mais acceptables pour des besoins différents : la révision de la négation s'effectue dans ce cas pour satisfaire aux conditions de vérité *représentationnelles* des énoncés.[40] Le cas de la théorie des atomes de Niels Bohr, l'utilisation complémentaire de la mécanique classique et quantique selon le type d'objets étudiés, et le cas des vérités scientifiques en général servirait à justifier la satisfiabilité de (12) ; mais ce n'est pas au sens de cette vérité représentationnelle que Quine défendait pour sa part le principe de charité ; contrairement à une typologie des croyances telle qu'on la retrouve chez [Engel 1989a], où l'auteur insiste sur l'aspect *intentionnel* des concepts en distinguant entre des croyances littéralistes (factuelles), intellectualistes ou symboliques, les attributions de croyance faites sur la base du principe de charité reposent uniquement sur une interprétation factuelle de la vérité, si bien qu'une distinction entre plusieurs sens de la vérité entraîne un possible malentendu sur les critères de rationalité des interprétants. En bref, la vérité pragmatique ou 'quasi-vérité' renvoie à des ensembles d'énoncés incomplets, déductivement ouverts (par opposition aux ensembles clos

[40] Notons toutefois que, pour le dialéthiste Graham Priest, la révision de la négation se justifie également pour la vérité factuelle (correspondance) : il y a selon lui des états de choses contradictoires. Sur les arguments allant dans ce sens, voir notamment [Priest 1992, 1999] et [Beall 2000].

de la sémantique relationnelle) et dont la supposition de vérité *provisoire* dirige nos actions quotidiennes ; à l'inverse, la vérité factuelle ou classique est complète, déductivement close et prétend à une représentation exhaustive des faits du monde.[41] La vérité pragmatique admet ainsi des formes de croyance inconsistantes, mais dans la mesure où les énoncés crus se rapportent à des systèmes théoriques distincts, donc à des ensembles de croyance *disjoints* : il n'y a pas un ensemble total d'énoncés vrais ou faux dans un seul et unique langage clos, pour le paraconsistant ; il y a plutôt un ensemble d'énoncés consistants divisés entre plusieurs sous-structures théoriques, et le rôle d'une théorie rationnelle serait de décrire les relations entre ces énoncés sans aboutir à la conclusion catastrophique de l'explosion ou trivialité : admettre une contradiction, ce n'est pas admettre la vérité de tout et n'importe quoi, contrairement au raisonnement classique de l'*ex falso sequitur quodlibet*.

Cela étant, proposer une distinction entre des sous-structures d'énoncés inconsistants ne crée pas une contradiction au sens où l'entendent les logiciens classiques, qui présupposent un domaine d'objets unique : admettre la vérité de ϕ dans une sous-structure s_1 et la vérité de $\neg\phi$ dans une autre sous-structure s_2, n'est-ce pas parler de deux mondes possibles distincts pour l'énoncé ϕ, et ne serait-il pas préférable alors de parler de la vérité de ϕ-dans-s_1 et de $\neg\phi$-dans-s_2 ? Si tel est le cas, alors il n'y a plus de contradiction explicite mais une conjonction de deux formes symboliques non-contradictoires $\phi_1 \wedge \neg\phi_2$, où ϕ_1 = p-dans-s_1 et $\neg\phi_2 = \neg$(p-dans-s_2). Les arguments paraconsistants basés sur l'emploi d'ensembles de théories scientifiques inconsistantes ne semblent plus en mesure de favoriser le dialéthisme, selon lequel l'on admet que $v(\phi \wedge \neg\phi) = 1$ dans un seul et même monde (modèle théorique) : si une sous-structure d'un langage équivaut à un modèle particulier, alors la vérité pragmatique ne signifie pas que ϕ et $\neg\phi$ appartiennent au même modèle mais à deux modèles distincts. La condition de consistance sémantique (C.$\neg$) de [Hintikka 1962] est violée malgré tout, puisque le partisan des croyances inconsistantes admet qu'un ensemble de modèles puisse être contradictoire. Si l'on interprète chaque modèle théorique comme un monde possible, alors la vérité pragmatique nécessiterait une révision des conditions sémantiques de la logique modale doxastique. Soit W l'ensemble des mondes compatibles avec une théorie particulière (la mécanique classique, par exemple), et soit W* l'ensemble des mondes compatibles avec une autre théorie scientifique (la mécanique quantique) ; dans ce cas, Bϕ est (pragmatiquement) 'vraie' dans un monde de W si et seulement si ϕ est considéré comme vraie dans tous les mondes w* accessibles à w dans W ; en ajoutant cette clause d'accessibilité entre plusieurs ensembles de mondes théoriques, il devient possible de croire un énoncé dans un système de croyances particulier et de croire sa négation dans un autre système. Contre cette variation de la notion de vérité, qui suppose donné un système de croyances complexe et doté de sous-systèmes de mondes possibles, nous maintiendrons l'approche factuelle dont s'inspire le traducteur charitable : l'acte de compréhension ne peut pas se baser sur une organisation

[41] La quasi-vérité rappelle le cas d'indécidabilité Xϕ de [Ellis 1976], lorsque ϕ n'est considérée ni comme vraie (Tϕ) ni comme fausse (Fϕ). On obtient dans ce cas l'acceptabilité de la forme contradictoire $\phi \wedge \neg\phi$, à condition que celle-ci s'exprime sous la forme X$\phi \wedge$ X$\neg\phi$. Mais être indécis (X) et considérer comme vrai ou comme faux (T ou F) sont deux paires d'attitudes distinctes, car nous n'admettons pas les contradictions de forme T$\phi \wedge$ Fϕ ou T$\phi \wedge$ T$\neg\phi$. A quelle forme contradictoire fait allusion la quasi-vérité : à X$\phi \wedge$ X$\neg\phi$, à T$\phi \wedge$ T$\neg\phi$? Sur la notion de 'quasi-vérité, voir notamment Mikenberg, I, Chuaqui, R., & da Costa, N.C.A. (eds.), "Pragmatic truth and approximation of truth", *Journal of Symbolic Logic* **51**, pp. 201-221.

préalable de la connaissance aussi complexe, et d'autres moyens que l'analyse logique doivent exister pour distinguer un acte primitif d'assentiment de la croyance scientifique.[42]

Comment distinguer les différentes intentions de communication de l'interprété, selon qu'il cherche à décrire un fait, exprimer une idée symbolique non factuelle, faire de l'ironie, etc. ? Chacune de ces intentions nécessite-t-elle la création d'une variété de logiques doxastiques ? Au lieu de prendre en compte les différentes formes de croyances qui en résultent, nous plaiderions pour une division du travail dans l'interprétation des croyances, c'est-à-dire une répartition des tâches : le logicien a pour but de proposer un premier type d'interprétation des croyances en termes de condition de vérité et de fausseté factuelle, et sa seule fonction est de déterminer les règles d'interprétation pour les phrases dont l'intention de communication est de type *déclaratif* (dire ce qui est le cas) ; quant aux autres intentions de communication, elles sont laissées à disposition des sociologues, anthropologues ou spécialistes d'autres types d'usage conceptuel. En bref, nous plaiderions pour une théorie minimale et déflationniste de la vérité : dire ϕ, c'est croire que ϕ est vrai, c'est-à-dire que ϕ décrit un état de choses réel et s'utilise le plus généralement comme un concept *désignatif* à usage *locutoire* ; en cas de désaccord entre un comportement et notre traduction, la révision n'affectera pas les concepts *formatifs* mais portera soit sur les concepts désignatifs (lorsque les termes de couleurs sont plus ou moins nuancés entre des langues différentes, par exemple), soit sur l'intention du locuteur (acte locutoire, illocutoire ou perlocutoire). Dans le cas où une traduction de première main, celle du logicien, ne semble pas conforme au comportement du locuteur, alors le logicien devrait céder sa place et laisser place à d'autres types d'interprétations des formes de croyances plutôt que de réviser ses propres concepts formatifs ; si le principe de charité peut s'avérer insuffisant lors d'une traduction, c'est parce que l'intention de dire la vérité ne constitue pas la seule intention de communication possible et parce qu'une phrase peut faire appel à d'autres conditions de satisfaction que la vérité ou la fausseté.[43] Le logicien propose de traduire ce que nous appelons ici les *énoncés* ; rien de plus.

C'est cette division du travail que nous allons illustrer dans ce qui suit en faveur du principe de charité logique, dans le but de maintenir les logiques modales normales (à superstructure classique) mais sans prétendre à ce que cette charité s'applique *sans restriction* : une division du travail entre les interprétants signifie que l'approche logique de la traduction, limitée par nous à l'aspect locutoire des langues naturelles, ne doit pas constituer le seul angle d'approche de l'interprétant ; au contraire, la diversité des intentions de communication exige des traductions graduelles et complémentaires, puis une sélection des traductions les plus plausibles. La stratégie envisagée pour obtenir une traduction la plus adéquate possible consisterait à réviser dans l'ordre la fonction intentionnelle de la phrase, ses termes désignatifs puis, en dernière analyse, ses concepts formatifs. En d'autres termes, il est

[42] "Nous nous méprendrions sur la situation si nous négligions le rôle social et si nous insistions sur la charité logique. Parce que la langue peut avoir des fonctions autres que la *communication de vérités*, la traduction ne peut pas toujours être charitable" ([Thagard & Nisbett, 1983], p. 254).

[43] Ainsi les intuitionnistes préfèrent-ils la notion d'assertabilité à celle de vérité, estimant que leur critère de satisfaction d'un énoncé justifie la révision de certaines conditions sémantiques ; un acte illocutoire n'est pas non plus vrai ou faux, mais fructueux ou infructueux, etc. Nous suivons ici la démarche de Quine, qui consiste à réserver au domaine logique la recherche de conditions de satisfaction particulières des énoncés : les conditions de *vérité*. La réduction de la vérité à l'assertabilité reflète un choix philosophique dans la représentation de notre pensée et, en cela, elle affecte la question de la traduction radicale.

toujours plus plausible d'expliquer un problème de traduction par une autre raison que le dysfonctionnement de nos connecteurs logiques. 'Sauver ce qui est obvie' n'est pas une réponse *ad hoc* de Quine aux logiques déviantes et au pluralisme logique, selon nous, puisqu'elle tient compte de la variété de ce que l'on peut vouloir dire par le biais d'une seule et même expression.

Rappelons que le choix entre une rationalité consistante ou paraconsistante repose sur un choix essentiellement *heuristique*, relatif à l'idée de logique qui anime les partisans ou opposants de la charité : pour que la logique demeure la science des vérités les plus générales qui soient et indépendamment de l'objet étudié, le maintien d'une politique de mutilation minimum répond à ce souci de généralité de la logique. L'étude des structures les plus générales *et minimales* passe par une restriction dans le choix des opérations logiques, par opposition à l'attitude méthodologique des pluralistes globaux.[44]

Pour illustrer notre défense 'charitable' de la charité, revenons notamment sur le rejet de la règle d'adjonction, prôné par certains logiciens afin de distinguer les formules précédentes (12) et (13). Tel que nous comprenons la procédure de traduction, et sur la base du principe de décitation, nous supposons encore une fois que (13) implique (14) en vertu de la règle de non-contradiction : dire '$\neg\phi$' implique $B\neg\phi$; or $\neg(\phi \wedge \neg\phi)$, donc non '$\phi$', donc $\neg B\phi$. Pour mieux comprendre les cas de 'conjonctivite'[45] ci-dessus (violation de la règle de distribution modale sous conjonction), on peut distinguer deux sortes de croyance : une croyance contradictoire *involontaire*, d'une part, où le contenu contradictoire de la croyance apparaît à l'insu du locuteur et manifeste une forme d'incompétence linguistique de sa part ;[46] une croyance contradictoire *volontaire*, d'autre part : c'est cette sorte de croyance qui est susceptible d'entraîner un véritable cas d'irrationalité, puisqu'elle est prononcée au vu et au su du locuteur et ne peut donc pas être faite à son insu. Pour faire la distinction entre les deux types de comportement plus ou moins fidèles à la norme rationnelle, des auteurs tels que [Williams 1981] associent par exemple le nom d'inconsistance à la contradiction du premier type, involontaire ; dans ce cas de rationalité limitée, la croyance consiste à croire *à tort* la consistance d'un énoncé inconsistant, comme dans le cas des croyances relationnelles consistant à accepter et refuser à la fois d'attribuer une propriété à un seul et même individu :

> Notez que le traitement de ce cas de figure n'engage pas dans la faute de raisonnement *intentionnelle* consistant à supposer que [l'on] tient la croyance contradictoire selon laquelle ce

[44] En ce sens, l'argument de [Paoli 2003] et des partisans du pluralisme n'est pas faux en tant que tel mais déconseillé, si l'on veut 'sauver ce qui est obvie' et éviter une multiplication des opérations logiques suivant le contexte de discours. A cet égard, les arguments de la logique linéaire, quantique ou relevante ne sont pas erronés mais indésirables pour le partisan de la charité.

[45] Jeu de mots en référence à un article de Henry Kyburg, "Conjunctivitis" in Marshall Swain, ed., *Induction, Acceptance, and Rational Belief* (Boston : Reidel, 1970), pp. 55-82. Notre conjonction ne souffre d'aucun dysfonctionnement, et nous cherchons à montrer ici qu'il y a simplement un malentendu sur les symptômes de la maladie (de la mauvaise traduction).

[46] Faisant allusion dans son article "Sens et dénotation" au problème des croyances relationnelles, [Frege 1971] précisait qu'un langage idéal n'admettrait pas une telle incompétence linguistique dans la mesure où à chaque terme serait assigné une et seule extension : "Le sens d'un nom propre est donné à qui connaît suffisamment la langue ou l'ensemble des désignations dont il fait partie ; mais la dénotation du signe, à supposer qu'elle existe, n'est jamais donnée en pleine lumière". (p. 104) La traduction doit tenir compte de ce problème, en revanche, puisqu'elle s'effectue entre des langues naturelles 'impures'.

> qui est impossible est possible. Néanmoins, [on] constate par la suite que des énoncés contradictoires résultent de ce que [l'on] croit, auquel cas ce que [l'on] croit est contradictoire.[47]

Autrement dit, des énoncés peuvent être contradictoires parce que le locuteur ne réalise pas qu'il porte deux jugements opposés sur un seul et même sujet ; la logique épistémique quantifiée traite ce genre de situation, mais nous avons vu plus tôt qu'elle rapporte cet échec à l'usage *désignatif* des concepts et n'impliquait en rien une révision des concepts formatifs, c'est-à-dire les constantes logiques. Les exemples de Bernard J. Ortcutt ou de Tom, illustrant chez Quine l'opacité référentielle dans des contextes épistémiques, avait pour but de montrer que deux termes désignatifs d'individus co-référentiels ne sont pas substituables *salva veritate* dans n'importe quel contexte d'énonciation. Dans ces cas de figure, c'est l'application du principe de non-contradiction qui permettait de montrer le caractère non-extensionnel de la croyance et d'invalider la loi d'indiscernabilité des identiques : étant donné qu'un locuteur ne peut pas croire deux choses contradictoires, il doit réviser l'usage d'un des deux termes désignatifs incriminés. Si en revanche on admettait ici une révision du terme formatif de la *négation*, auquel la non-contradiction doit son statut de principe, le traducteur aurait le droit de supposer que l'interprété tient des croyances contradictoires. En d'autres termes, la charité logique consiste à déplacer le problème de traduction sur l'opacité référentielle : la difficulté réside dans la fonction désignative des termes d'individus, non dans la fonction formative de la négation ou de la conjonction.

Si notre hypothèse analytique de l'incompétence linguistique peut sauver les meubles, ici, elle semble plus discutable en revanche face au problème de l'*omniscience logique* ; comment gérer ce cas de figure dans le processus de traduction charitable ? Si l'on applique la charité quelle que soit la complexité des expressions à traduire, l'interprété sera considéré comme un agent logiquement parfait, capable de suivre toutes les conséquences de ses affirmations et de ne jamais se tromper sur les relations logiques entre ses croyances. Or nous savons que cette condition de rationalité est excessive :

> La seule raison que l'on pourrait avoir de nier [la présence de croyances contradictoires], c'est l'idée selon laquelle on croit toujours les conséquences logiques de ce que l'on croit. Elle est fausse.[48]

Mais la procédure de traduction, avait prévenu [Quine 1974], ne porte avant tout que sur des énoncés 'courts' et dont la complexité est telle que la longueur des raisonnements serait comparable à celle des *tautologies de surface*. De plus, la traduction entraîne une affirmation *explicite* des relations entre les énoncés de la part de l'interprété, ce qui signifie que la relation déductible entre les énoncés doit être connue du locuteur ; nous pouvons donc en conclure que, au cours d'une traduction, la règle de nécessitation n'est pas employée et seule la règle de distribution sur l'implication est supposée, en vertu de laquelle la relation de conséquence est assertée et donc connue. Formellement, cela signifie que la règle logique $\vdash(\phi \rightarrow \psi) \Rightarrow B(\phi \rightarrow \psi)$ n'est pas valable pour toute formule ϕ et ψ et que, en revanche, la règle logique $B(\phi \rightarrow \psi) \Rightarrow (B\phi \rightarrow B\psi)$ est valable dans la mesure où toute croyance est énoncée explicitement,

[47] [Williams 1982b], p. 282 ; nous soulignons.
[48] [Williams 1982b], p. 284.

donc *consciemment*. Le cas contre-intuitif de l'omniscience logique (l'agent connaîtrait toute forme de conséquence logique, y compris celles qu'il n'asserte pas) n'affecte donc pas le processus plus réaliste de traduction radicale. Le traducteur charitable doit plutôt tenir compte dans son application du problème de l'opacité référentielle, c'est-à-dire de la possible incapacité du locuteur à reconnaître une même extension sous deux termes d'individus distincts.

Une charité intégralement appliquée exigerait donc que les objets d'attitude propositionnelle soient toujours correctement décrits dans les croyances explicites, c'est-à-dire les *assertions* du locuteur, au risque de prétendre qu'il n'y a pas d'opacité dans le choix de l'extension ; or il est possible de croire des choses contradictoires d'une même personne, tout au moins à son insu. Dans ce cas, la croyance est rationnelle mais limitée et involontaire. Lorsque la croyance inconsistante est due à une relation ambiguë entre un objet et le symbole qui l'exprime, la croyance est dite *de re* ou *relationnelle*. Mais l'inconsistance d'une croyance n'est pas forcément de nature linguistique et peut être due également à un défaut de perception ou d'attention. Lorsque de telles confusions nous font prendre une personne pour une autre, sommes-nous en présence d'un cas de croyance irrationnelle qui réfute l'interprétation charitable ? Pour conserver la non-contradiction des croyances et la règle d'adjonction, donc pour rejeter toute distinction entre contradiction et inconsistance,[49] rappelons que la logique épistémique quantifiée propose la distinction entre deux types d'attribution de croyance : une croyance *de dicto* ou notionnelle, dans laquelle l'objet de croyance est celui que vise l'interprété au sein de son discours ; une croyance *de re* ou relationnelle, dans laquelle l'objet de croyance n'est pas forcément celui que vise l'interprété :

> L'acceptabilité des attributions est clairement due à ce que Quine appelle leur forme *relationnelle*. Et ce qui explique cette acceptabilité c'est simplement que, en assignant de telles attitudes à une personne, il n'y a aucune présomption que les concepts (prédicats) que le rapporteur invoque pour identifier l'entité sur laquelle la croyance rapportée est dite porter soient eux-mêmes –ces concepts –crédités par le sujet croyant. Le rapporteur emploie *ses* concepts pour identifier un objet auquel il considère que le sujet croyant est cognitivement lié. Le fait que ces concepts ne soient pas supposés être possédés par le sujet croyant est indiqué en conservant le vocabulaire que l'on utilise pour exprimer les concepts situés en dehors de la portée de l'opérateur de croyance.[50]

Lorsqu'un désaccord apparaît au sujet de la traduction d'une croyance, expliquer ce désaccord par un cas d'opacité référentielle revient à montrer une incompatibilité entre les interprétations *de dicto* et *de re* : $(\exists x)BFx \wedge \neg B(\exists x)Fx$, ou $B(\exists x)Fx \wedge \neg(\exists x)BFx$. Mais cette incompatibilité ne produit aucune contradiction formelle : la distinction entre les portées de la

[49] [Prior 2002] s'accorde avec [Williams 1981] sur cette distinction : "Il est même rationnel d'adopter un ensemble de croyances qui ne peuvent pas être toutes vraies, encore qu'on puisse difficilement jusqu'à une paire de croyances mutuellement contradictoires (...)". (p. 94). Les deux auteurs admettent donc l'éventualité d'un ensemble de croyances $\{p, \neg p\}$ mais rejettent l'énoncé $p \wedge \neg p$, ce qui revient à supprimer entre-temps la règle d'adjonction. [Prior 2002] fait allusion également aux logiques modales normales de [Hintikka 1962], ajoutant qu'"il n'est pas vraisemblable de réduire avec Hintikka la rationalité à la cohérence" (p. 97).
[50] [McGinn 1977], p. 525.

négation permet de l'éviter, et c'est cette distinction qui a été invoquée par Quine pour rendre compte de l'opacité référentielle.[51]

Le choix de la révision des mots logiques est donc évité avec succès, ici : dès que le sujet croyant s'aperçoit qu'il tient deux croyances contradictoires au sujet d'un même individu, il sera enclin à réviser ses croyances et à avouer qu'il s'était trompé sur l'une d'elles. Dans le cas où le sujet admet ϕ et $\neg\phi$, on supposera qu'il a changé d'avis entre-temps et que l'ensemble de ses croyances (de ses mondes doxastiquement possibles) doit être révisé et contracté ; à cet égard, il semble qu'une admission du modèle inconsistant $\{\phi,\neg\phi)$ ne tienne pas compte de l'*unité de temps* des ensembles de croyances.[52] Le même genre d'argument conservateur peut être utilisé pour le traitement de paradoxes tels que la *préface*, la *loterie* ou la *décitation* de Kripke.

Sans entrer dans les détails de leur argumentation, chacun de ces cas de figure consiste à faire état d'une inconsistance au sein d'un ensemble de croyances différents et à discuter ses effets sur les règles de logique. La question est : une paire de croyances contradictoires peut-elle être admise par expansion au sein d'un même système de croyances, ou toute contradiction implique-t-elle une contraction de ce système à un seul des termes contradictoires ? Dans l'argument de la préface, le préfacier d'un ouvrage déclare qu'il croit à l'ensemble des affirmations qu'il y a faites mais que, l'erreur étant humaine, il est possible qu'une ou plusieurs erreurs se soient glissées lors de la rédaction ; l'argument de la loterie est similaire, dans lequel chacun des participants croit qu'un des tickets existants est le ticket gagnant mais personne ne croit que ce ticket est le sien. Dans les deux cas de figure, l'inconsistance réside dans l'opposition logique des deux croyances respectives : la croyance du préfacier que tout ce qu'il a écrit est sincère et la croyance des joueurs de loterie qu'aucun de leur ticket n'est gagnant s'expriment toutes les deux par une assertion de forme logique $\forall xB(Fx)$, où B symbolise l'opérateur de croyance, F la propriété d'être vrai ou d'être perdant, et x une affirmation quelconque du livre ou un ticket de loterie quelconque. Par opposition, la croyance du préfacier qu'au moins une erreur s'est glissée dans la publication et la croyance des joueurs de loterie qu'au moins un des tickets est le ticket gagnant donnent la forme logique $\exists xB(\neg Fx)$, où $\neg F$ désigne la propriété d'être vrai (donc d'être faux) ou de ne pas être perdant (donc d'être gagnant). Si l'on traduit respectivement les quantificateurs universels et existentiels ci-dessus en termes de formes réduites conjonctives et disjonctives, on obtient les ensembles de croyances contradictoires : $B(Fx_1 \wedge Fx_2 \wedge \ldots \wedge Fx_n) \wedge B(\neg Fx_1 \vee \neg Fx_2 \vee \ldots \vee \neg Fx_n)$, de forme schématique $\phi \wedge \neg\phi$ et où *n* symbolise le nombre des affirmations du livre du préfacier ou le nombre des tickets de loterie. L'ensemble doxastique ci-dessus correspond à une forme logique non pas contradictoire mais inconsistante, de type $BFx \wedge B\neg Fx$, mais la

[51] Pour distinguer les croyances *de dicto* et *de re*, Quine parle d'une "interface opaque entre deux mondes : celui de l'homme mentionné dans l'attitude, si profonde que soit son ignorance, et celui du sujet qui attribue l'attitude. Parfois une brèche apparaît dans l'interface" ([Quine 1993], p. 104). C'est cette brèche qui produit l'opacité référentielle et dont l'interprétant doit tenir compte pour comprendre le comportement de l'interprété, en apparence irrationnel.

[52] A quelle condition révise-t-on ses croyances, ou change-t-on d'avis ? Nous renvoyons ici (sans plus de détail) à la théorie des systèmes de croyance de [Gardenförs 1988] et l'approche *dynamique* de la logique épistémique, qui tient compte des *expansions* et *contractions* des ensembles de croyances. Dans le cas présent, nous posons qu'une conjonction de croyances contradictoires entraîne la contraction de cet ensemble, contrairement aux paraconsistants qui admettraient une extension de l'ensemble à des termes contradictoires.

relation d'implication $B\neg\phi \rightarrow \neg B\phi$ permet d'obtenir un ensemble de croyances contradictoire de type $BFx_n \wedge B\neg Fx_n$ (pour au moins cas de croyance *n*).

Faut-il admettre des contradictions vraies, étant donné l'aspect convaincant des deux scénarios exposé plus haut ? Contre la stratégie de révision des opérations logiques qui en découle, nous poserons au contraire que ces scénarios n'ont pas été correctement symbolisés parce que toute instance de x doit être explicitement assertée pour être véritablement 'crue' : un agent croit que ϕ s'il est prêt à affirmer la vérité de ϕ dans une assertion publique. Or ce critère charitable n'est pas observé dans les paradoxes de la préface et de la loterie : le préfacier prétend qu'au moins une de ses affirmations est fausse pour un certain x, mais sans préciser quelle est la valeur de ce x ; les joueurs de loterie admettent qu'au moins un ticket x est gagnant, mais sans dire quelle est la valeur précise de ce ticket. En l'absence d'affirmations explicites sur la valeur du x qui est F, peut-on parler de croyances réellement inconsistantes à son sujet, et ne doit-on pas évoquer plutôt un ensemble de croyances vagues, ou mal définies ? Nous optons pour cette dernière interprétation des raisonnements inconsistants, dont l'aspect paradoxal repose sur une formulation équivoque des ensembles de croyances : les sujets soi-disant inconsistants s'engagent séparément sur la vérité de ϕ et $\neg\phi$ mais sans jamais asserter $\phi \wedge \neg\phi$, ce qui confirmerait à première vue le rejet de la règle d'adjonction $\phi, \psi \vdash \phi \wedge \psi$ en logique doxastique ; cependant, les énoncés disjoints ϕ et $\neg\phi$ ne sont *assertables* ni par le préfacier ni par les joueurs de loterie, dans la mesure où leurs affirmations contradictoires portent sur des objets de croyances floue et sans imprécis. Pas de sens sans identité, pas d'entité sans identité : en l'absence d'objets de croyance bien identifiés, rien n'interdit l'interprète charitable de dissoudre ces paradoxes en refusant d'attribuer un sens à ces attribution d'inconsistance. De notre point de vue, un énoncé sans assertion explicite à l'appui n'exprime pas une croyance bien définie, et notre argument d'une symbolisation incorrecte des croyances peut être invoqué de préférence à la création d'une logique de l'inconsistance et contradiction en logique.

L'origine des paradoxes reposerait ici sur la transcription des quantificateurs en termes de formes réduites conjonctives ou disjonctives : certes, l'universelle est une conjonction de particuliers, l'existentielle est une disjonction de particuliers ; or lorsque le préfacier déclare qu'il croit à la totalité des affirmations de son livre, il n'est cependant pas en mesure d'asserter, c'est-à-dire d'affirmer la vérité de chacune des affirmations particulières de l'ouvrage ; pour cette raison, il les croit indistinctement mais sans que cette croyance vague ne contredise sa seconde affirmation ('au moins une erreur a pu se glisser lors de la publication'). C'est cette identification explicite entre croyance et assertion que nous développerons dans la suite de l'essai, ce qui nous conduira à une réinterprétation illocutoire des expressions épistémiques du langage ;[53] une telle identification a pour conséquence de maintenir la règle d'adjonction : $B\phi, B\neg\phi \vdash (B\phi \wedge B\neg\phi)$, ainsi que la relation d'équivalence

[53] [Barcan Marcus 1983] renvoie à une théorie de la croyance linguistique similaire chez Ramsey, tirée de "Facts and propositions", pour en tirer la conséquence que, chez Ramsey, "'donner son assentiment à p', 'asserter que p' et 'croire que p' deviennent pour lui des usages interchangeables" (pp. 326-7). Cela ne revient pas à refuser toute attribution de croyance *de re* aux animaux dépourvus de langage symbolique, mais à affiner les attributions spécifiques pour les êtres humains : les cas d'opacité référentielle ne concernent que nous, humains, et le programme de traduction radicale peut très bien concilier un critère d'assertion publique et un critère strictement comportemental.

entre inconsistance et contradiction : $(B\phi \wedge B\neg\phi) \Leftrightarrow B(\phi \wedge \neg\phi)$. La logique des croyances demeure donc classique, tant qu'une croyance doit porter sur un objet bien identifié ; l'acte d'assertion est considéré comme le critère nécessaire au moyen duquel une croyance peut être attribuée à autrui ; l'acte de discours est considéré comme nécessaire pour fonder une logique des attitudes épistémiques.

Quant au paradoxe de la décitation de Kripke, on peut l'assimiler plus simplement à un cas d'opacité référentielle et diagnostiquer la séparation entre deux objets de croyance distincts. Dans [Kripke 1979], l'auteur met à l'épreuve sa théorie de la désignation rigide des noms propres et cite le cas d'un Français parti vivre à Londres, Pierre, dont l'ensemble des croyances serait problématique : d'un côté, il a hérité de la France la croyance coutumière selon laquelle Londres est une jolie ville et l'asserte publiquement : 'Londres est jolie' ; d'un autre côté, il vit désormais dans un quartier insalubre de Londres et a appris à s'exprimer dans la langue locale : 'London is not pretty', exprimant son aversion pour le quartier dans lequel il vit. En vertu des principes de traduction et de décitation, toute assertion d'une langue naturelle trouve sa traduction dans une autre langue et exprime une même croyance de la part de son locuteur (énoncer ϕ-en-français équivaut à énoncer ϕ-en-anglais et à croire que ϕ est vrai). Les deux énoncés ci-dessus peuvent être interprétés ainsi : 'London is not pretty' signifie en français l'énoncé 'Londres n'est pas jolie', donc Pierre croit à la fois que Londres est jolie et que Londres n'est pas jolie : $B_{Pierre}(\phi \wedge \neg\phi)$. Conformément à notre explication précédente en termes d'opacité référentielle, une réponse possible à ce paradoxe consiste à dire que Pierre ne fait pas allusion à la même ville lorsqu'il parle de 'Londres' et 'London' : ces deux noms propres ne renvoient pas au même objet dans le monde doxastique de Pierre, et l'utilisation des noms propres dans un contexte épistémique produit un cas hyperintensionnel contraire à la théorie de la désignation rigide de Kripke. A supposer que la co-référence de 'Londres' et 'London' soit le cas dans tous les mondes *métaphysiquement* possibles, elle ne tient pas dans un contexte de croyance où l'on distingue la nécessité métaphysique de la contingence épistémique. Ce paradoxe de la décitation met en évidence le problème de la forme logique des attitudes propositionnelles : consistent-elles dans une relation entre un sujet et une proposition objective, un état psychologique particulier, une représentation privée ? Nous excluons ce dernier cas de figure, dans la mesure où deux croyances ne pourraient jamais porter sur un même objet si l'on définissait celles-ci en termes de représentations frégéennes ; nous excluons également le premier cas, qui ne tient pas compte des conditions ambiguës dans lesquelles un sujet croyant peut utiliser des termes désignatifs. La signification de 'Londres' et 'London' n'est pas totalement dans la tête de Pierre et ne dépend pas de lui seul, conformément au critère de l'usage public des termes d'un langage ; cependant, un même nom propre peut renvoyer à différents objets dans des mondes épistémiquement possibles, et c'est sur la base de cette opacité que le paradoxe de la décitation peut trouver une réponse : la croyance met en relation un sujet avec un *énoncé*, et les conditions d'identité de ces énoncés dépendent de la référence qu'elles prennent *pour l'auteur de l'énonciation*. Les conditions d'identité des énoncés sont donc plus précises que celles des propositions, définies pour n'importe quel sujet locuteur, et cette distinction est la solution invoquée par [Lewis 1981] : lorsque Pierre asserte ses deux énoncés de langue française et anglaise,

> ses voisins pourraient dire de lui 'Il croit que Londres est jolie' ; mais ce n'est pas à *nous* de le dire, compte tenu de ce que 'Londres' veut dire dans nos bouches et compte tenu de son isolement à la fois de notre Londres et de notre 'London'.[54]

Autrement dit, les attributions de croyance doivent être renforcées par un critère d'assertion publique : les propriétés F et ¬F doivent être prédiquées d'un seul et même terme *asserté* par un seul et même locuteur, que ce terme unique soit 'London' ou 'Londres' ; en l'absence de la prémisse auxiliaire évoquée dans [Hintikka 1962], B_{Pierre}(Londres = London), il est possible que Pierre associe des objets distincts aux deux noms propres dans son ensemble de croyance (Londres ≠ London) ; dans le même ordre d'idées, [Barcan Marcus 1983] suit cette solution de préférence à une conjonction de croyances contradictoires, hypothèse qu'elle juge impossible :

> Mon intuition concernant l'usage de ma propre langue est que, dans des cas de figure comme celui de Pierre, dès qu'une impossibilité a été *révélée*, je dirais que j'ai seulement déclaré que Londres est différente de Londres. En effet, cela équivaudrait au fait de croire qu'une chose n'est pas identique à elle-même, et je ne pourrais sans doute jamais croire cela.[55]

Un autre contre-exemple à l'interprétation charitable est le phénomène psychologique de la duperie de soi. En gros,

> La duperie de soi [*self-deception*] existe, dirai-je, lorsqu'une personne se ment à elle-même, c'est-à-dire se persuade de croire ce qu'elle sait ne pas être le cas. En bref, la duperie de soi a pour conséquence que B croit à la fois p et non-p dans le même temps.[56]

Ce problème est similaire au paradoxe de Moore, dans lequel des énoncés de forme (4) et (5) sont dits absurdes sans être explicitement contradictoires. Le cas des croyances inconsistantes ci-dessus a montré que toute attitude peut être considérée comme rationnelle lorsque son contenu contradictoire échappe à l'attention du locuteur et tant que celui-ci croit suivre la norme. Par opposition, les auteurs paraconsistants cités précédemment ne recourent pas à des hypothèses analytiques pour préserver la consistance des ensembles de croyance, préférant imposer leurs propres schémas d'inférence inconsistants. Comment reconnaître si un ensemble de croyances est volontairement inconsistant ou s'il manifeste simplement une erreur ou incompétence logique de la part de l'interprété ?

---

[54] [Lewis 1981], p. 289. La décitation des croyances de Pierre doit respecter ainsi des conditions d'énonciation explicite de sa part : il faut s'assurer que "Londres" et "London" renvoient *pour lui* à la même ville. Le paradoxe de la décitation ne se distingue en rien des exemples de [Quine 1953], dans ce cas. Un autre problème est celui des conditions d'individuation des objets de croyance. On trouve dans [Bäuerle & Cresswell 2003] une référence à la théorie des représentations mentales de Fodor : deux croyances portent sur un même objet si elles correspondent à la même image mentale chez le sujet. Cette hypothèse suppose un critère psychologique de type introspectif que nous n'avons pas admis jusqu'ici, dans notre approche behavioriste des attributions d'attitude.

[55] [Barcan Marcus 1983], p. 330. D'après Barcan Marcus, l'issue du paradoxe de la décitation consiste donc à trouver un moyen de démontrer que Pierre associe des choses différentes à 'London' et 'Londres', puis à réviser en conséquence l'ensemble de croyances inconsistant.

[56] [da Costa & French 1989], p. 442.

Le problème du choix de la traduction réside dans notre critère d'*explication* de la rationalité : la logique épistémique a repéré parmi les ensembles de croyances des difficultés qui garantissent la rationalité en cas d'interprétation contradictoire ; lorsqu'un sujet dit des choses contradictoires, par exemple, il se trompe la plupart du temps et ne se contredit pas volontairement.[57] Une interprétation de ce genre n'explique plus dans ce cas l'inconsistance par une erreur de traduction, mais par une incompétence linguistique de la part du sujet croyant ; dans les deux cas, il s'agit de sauver ce qui est le plus obvie en cas de contradiction. La position défendue ici va en faveur d'une application *modérée* de la charité : il y a des exceptions à la règle, mais ces exceptions n'exigent pas une révision des règles, en l'occurrence logiques. A partir d'une division du travail entre plusieurs niveaux d'interprétation (selon que les termes sont désignatifs ou formatifs, et selon l'intention de communication), nous prêchons donc en faveur d'une charité *complémentaire*. Comme le signale le passage qui suit, l'application de la charité est modulable et peut s'appliquer selon plusieurs degrés de sévérité :

> On peut distinguer au moins cinq niveaux de sévérité, engendrant les principes suivants qui vont du plus faible au plus sévère : (a) Ne pas supposer *a priori* que les gens sont irrationnels ; (b) Ne pas accorder de priorité spéciale à l'interprétation selon laquelle les gens sont irrationnels ; (c) Ne pas juger les gens comme irrationnels tant que vous n'avez pas une explication empiriquement justifiée de ce qu'ils font, lorsqu'ils violent les standards normatifs ; (d) N'interpréter les gens comme irrationnels qu'en vertu d'évidences manifestes ; (e) Ne jamais interpréter les gens comme irrationnels.[58]

Le cas d'opacité référentielle est traité par un principe de charité de degré (c), lorsque l'irrationalité (contradiction) apparente du locuteur est corrigée en termes d'échec de la désignation. De préférence à une révision parcellaire et locale des procédures logiques, la charité logique s'inscrit donc dans une démarche d'interprétation théorique qui contient plusieurs strates, parmi lesquelles le théoricien tente de trouver un consensus cohérent pour les différentes données observables :

---

[57] La charité n'exclut pas d'office tout comportement irrationnel, mais il s'agit pour le traducteur de choisir entre plusieurs explications plausibles des énoncés d'autrui : "Le principe de charité, là où il s'applique, peut être subsumé adéquatement sous un principe plus fondamental et qui s'applique partout dans la construction des manuels de traduction : traduire de façon à maximiser l'explicabilité des croyances et actions attribuées à des sujets (...) ce 'principe d'explicabilité' nous autorise à attribuer l'irrationalité à nos sujets dans certaines situations. Je ne conçois pas tant cet article comme répudiant le principe de charité que comme affinant les formulations actuelles de ce principe". ([Henderson 1987], p. 227)

[58] [Thagard & Nisbett 1983], p. 252. Une application du critère de charité maximal (e) impliquerait la soumission normative des sciences cognitives à la logique classique, ce qui n'est guère concevable : "Les psychologues cognitifs contemporains étudient la pensée humaine comme un type de procédure d'information. Ils demandent : quels mécanismes de procédure expliquent notre performance cognitive ? On pourrait présumer qu'un mécanisme essentiel serait un système inférentiel qui sert à dériver de nouvelles croyances à partir d'anciennes, conformément aux principes de la logique. Si les gens avaient un tel mécanisme, alors les principes de charité inférentiels forts, au niveau (d) ou (e), pourraient être de rigueur ; nous ne devrions pas donner d'explications de la procédure d'information humaine qui supposent que les gens violent les règles logiques". ([Thagard & Nisbett 1983], pp. 255-6. Mais la violation des lois logiques ne signifie pas pour autant une *révision*, c'est-à-dire une modification fondamentale de ces lois pour expliquer nos raisonnements quotidiens ; comprendre cette violation conformément à notre approche modérée de la charité (qui se situerait au niveau (d) des degrés de charité ci-dessus), c'est expliquer par une autre discipline que la logique ce qui empêche le sujet de suivre les règles de la pensée 'pure'.

alors que le principe de charité est contraignant dans les premières étapes de la construction de manuels de traduction, il n'est pas contraignant dans les étapes ultérieures, où nous nous engageons dans une traduction [affinée]. En effet, il ne s'applique pas à cet endroit. Partant d'un manuel de traduction de première approximation qui conduit à des attributions de croyance irrationnelle apparente et apporte une information pour les expliquer, l'investigateur n'exige plus le principe de charité. Dans les étapes ultérieures de la construction du manuel de traduction –dans la 'construction d'un manuel de traduction [affiné]'– interpréter les usagers du langage source comme ayant raison ou étant consistants est *hors de propos*. Avec l'accès à une information généralement correcte concernant les croyances des locuteurs du langage source, le temps est venu de passer au travail des sciences sociales : l'explication.[59]

Au final, l'application du principe de charité et la solution au problème de l'irrationalité des croyances dépend en partie de la façon dont les traducteurs conçoivent la fonction de la logique en général : selon qu'ils la considèrent comme une discipline normative fixe ou comme une discipline descriptive variable, l'accent sera plus ou moins mis sur la division du travail entre différents types d'explication de l'inconsistance ou sur la révision des constantes logiques, respectivement. Pour comprendre la relation entre croyance, vérité et signification, le principe de charité a été invoqué comme un moyen d'illustrer le thème de la sémantique formelle et des conditions de vérité des énoncés ; le choix du cadre sémantique dépend ainsi des frontières imposées à la discipline logique.

3.2.4 Des limites à l'analyse logique

Suite à l'argument des 'mondes possibles impossibles', utilisé pour bloquer la conséquence de l'omniscience logique, [Rantala 2003] avait reconnu lui-même que sa sémantique des modèles d'urnes a un prix en philosophie de la logique, à savoir le choix d'une approche *descriptive* de la discipline et opposée au 'sens obvie' que Quine voulait maintenir en logique. Dans le cas des logiques d'attitudes, notamment, "il y a manifestement une tension entre deux buts différents dans une telle logique. Un but est celui de constituer une théorie explicative et empirique (réaliste) (...) l'autre est celui de rendre une logique aussi respectable que possible –en dépit du fait que le premier but conduit à de nombreuses caractéristiques non-extensionnelles.[60] Le choix d'une révision des sémantiques pose un problème en philosophie de la logique, à savoir : quelle fonction accorder à la logique en général ?

Jusqu'ici, nous avons plaidé pour une approche normative et 'obvie' des règles sémantiques selon Quine, favorable à une représentation charitable et non-contradictoire des ensembles de croyances.[61] Plusieurs logiques philosophiques étendent la logique à des opérations non-vérifonctionnelles, mais une distinction s'impose entre quelques-uns de ces

[59] [Henderson 1987], pp. 232-233 ; nous soulignons pour constater ici la différence avec la charité extrême de [Davidson 1973a], dont un passage précédent disait que les interprétés doivent être considérés comme "ayant raison dans la majeure partie des cas" (p. 19).

[60] [Rantala 2003], p. 196.

[61] On retrouve la même distinction méthodologique au sujet de la logique zandée, liée au thème du relativisme logique : "Nous pouvons classer [les réponses possibles aux croyances inconsistantes] en deux catégories distinctes. D'un côté, nous trouvons des propositions basées sur l'idée que, pour résoudre ce problème, une *logique différente* devrait être attribuée aux Zandés (...) A l'évidence, une telle logique devrait nous permettre de comprendre le statut des inconsistances. D'un autre côté, dans un style plus conservateur, il y a des solutions alternatives caractérisées par le fait de maintenir la *même logique –classique–* et d'expliquer le phénomène par un autre moyen" ([da Costa & Bueno & French 1998], p. 43).

échantillons : entre les logiques épistémiques, non-monotones, paraconsistantes, quantiques, etc., certaines sont enrichies et d'autres sont déviantes. Les cas de logique déviante sont les seules susceptibles de menacer la charité, parce qu'elles rejettent le manuel de traduction élémentaire basé sur les fonctions de vérité de la logique classique. Les problèmes sont d'ordre sémantique : qu'est-ce qu'une *valeur de vérité*, et comment l'interpréter (par un indice de comportement verbal et observable, ici) ? Que signifie une 'contradiction vraie' : vraie en vertu d'autres conditions de vérité des énoncés complexes, au sein d'une logique déviante (argument du pluralisme) ; ou vraie au sens d'être établie et expliquée par d'autres disciplines que la logique (argument de la charité modérée) ? En guise d'hypothèse analytique pour conserver la théorie logique vérifonctionnelle, on peut supposer que la logique constitue une étape préparatoire dans l'interprétation des comportements humains (par le biais de règles de traduction récursives, régulières, monotones, etc.). Contre l'argument de l'erreur de traduction, la paraconsistance constitue-t-elle une solution favorable à l'inconsistance ? Pour revenir sur l'argument de la quasi-vérité ou vérité pragmatique de [French 1990], celui-ci est considéré parfois comme une raison supplémentaire de maintenir la charité ; ce qui suit distingue notamment l'expression logique des croyances et les raisons *psychologiques* de leur non-omniscience ou inconsistance :

> Mais rien n'oblige un sujet (...) à croire toutes les conséquences logiques de [ses] croyances. Les lois de la logique nous disent ce que nous *pouvons* inférer, non ce que nous *devons* inférer. De plus, il y a des raisons psychologiques de supposer que nos croyances sont organisées en sous-systèmes, de manière telle qu'une inconsistance pourrait être maintenue dans un sous-système sans affecter d'autres sous-systèmes (...) une inconsistance pourrait exister entre des sous-systèmes sans introduire nécessairement une inconsistance à l'intérieur des sous-systèmes, ou à un niveau de généralité et sans introduire d'inconsistance aux autres niveaux de généralité. On peut donc inclure de façon intelligible une contradiction dans un aspect de la pensée, tout en opérant en conformité avec la logique dans d'autres aspects.[62]

Une charité logique mal interprétée ou surestimée, comprise à un degré de sévérité maximal (e) de [Thagard & Nisbett 1983], peut être corrigée non par une révision des critères logiques de la rationalité mais par une division du travail scientifique qui limiterait la portée de l'analyse logique. Lorsqu'elles admettent un pluralisme local, les logiques philosophiques rejettent même le niveau de charité minimale (a) puisqu'elles admettent la vérité d'énoncés inconsistants, de préférence aux hypothèses analytiques exigées par les niveaux (c) et (d). Ces logiques n'ont plus de rien de normatif ni même de prescriptif, puisqu'elles n'incitent pas à maintenir l'application des inférences classiques ; au contraire, elles semblent plutôt descriptives et admettre une variation des procédures logiques selon le contexte d'usage. Est-on ici est présence d'une autre idée de la logique, selon l'interprétation du comportement ? Encore une fois, le problème est que nous ne savons pas au juste quel crédit accorder à ces révisions logiques dans notre problème épistémologique de la représentation des systèmes de croyances. Contre un relativisme débridé et sans limites, ce que la révision constante des règles sémantiques pourrait entraîner, on peut reprocher à certaines logiques déviantes de ne pas s'appuyer sur une caractéristique propre aux langues naturelles et historiquement

[62] [Thagard & Nisbett 1983], p. 254.

construites : l'*intuition logique* dont parlait [von Wright 1952], c'est-à-dire la signification dans l'usage ou tendance à appliquer certaines formes d'inférence de préférence à d'autres. Les cas de l'introduction de la disjonction ou de l'élimination de la conjonction, par exemple, semblent peu discutables et inciteraient davantage à la recherche d'une hypothèse analytique pour les défendre qu'à leur propre révision, en cas de désaccord durant la traduction. L'usage plaide en faveur d'une charité modérée et contre le choix d'une révision paraconsistante des ensembles de croyances :

> (...) dans le cadre d'une logique paraconsistante, on manque précisément d'une telle analyse simple pour les connecteurs (techniquement, cela signifie que les systèmes sont construits syntaxiquement et qu'on n'a pas de sémantique *intuitive* pour les différents langages obtenus). En conséquence, de tels systèmes prêtent particulièrement le flan à la critique quinienne : le coût de leur adoption est rédhibitoire. La fantaisie du logicien est donc plus limitée qu'il ne pourrait sembler au premier abord : il ne s'agit pas seulement de formaliser un ensemble de croyances logiques, il faut en même temps montrer comment ses croyances s'enracinent dans la maîtrise d'une langue.[63]

Contre l'exploitation du Principe de Tolérance de Carnap au service des logiques déviantes,[64] nous maintiendrons ainsi une distinction en sémantique formelle entre *discours* et *discours de vérité* : les langues n'ont pas pour unique fonction de produire des énoncés et de dire le vrai ou le faux, et l'extension du domaine de la logique dans le but d'interpréter les croyances risque de provoquer une confusion trompeuse entre ces différentes fonctions. Le problème est que le sens de la notion de vérité ne peut pas être établi par une référence à des faits objectifs, ce qui la réduit dans les sémantiques formelles au rôle de simple valeur algébrique. Qu'est-ce que la vérité, et n'y a-t-il pas une confusion des ordres lorsqu'une interprétation divergente des croyances conduit à une révision des lois logiques ? Si l'attribution d'une vérité exprime le succès d'une relation désignative entre un énoncé d'un langage et un état de choses, l'idée de concevoir des quasi-vérités comme des vérités 'partiellement' satisfaites au sein de théories incompatibles est ambiguë. [Quine 1975] avait déjà protesté contre ce genre de complications commis par la logique intuitionniste, en ce qui concerne l'assimilation de la vérité à la croyance justifiée.[65] Pour maîtriser le courant relativiste en logique et comprendre la multiplication des modèles, une réflexion s'impose sur la notion même de vérité et ses diverses interprétations (en termes d'attitude de tenir pour vrai, ici). Nous avons trouvé dans la bivalence une base épistémologique stable pour le raisonnement logique (maintenue dans la plupart des systèmes classiques 'enrichis', ou non-déviants), et le reproche adressé ici aux sémantiques déviantes concerne la cible qu'ils choisissent en cas de divergence entre un énoncé et sa méthode de traduction. En somme, admettre une charité modérée revient à faire une distinction entre la logique en théorie et dans

---

[63] [Laugier & Bonnay 2003], p. 64 ; nous soulignons.

[64] Voir notamment les articles de Beall et Restall, logiciens de l'école paraconsistante australienne favorables au pluralisme logique ; dans [Restall 1997] et [Restall 2002], notamment, l'auteur ne dit mot sur le critère d'usage des langues naturelles, réduisant tout comme Carnap la question des opérations logiques à une question strictement *interne*, c'est-à-dire relative à un langage formel défini par des conventions. Restall base son argumentation pluraliste sur une distinction entre plusieurs types de relations de conséquence logique ; or c'est le droit à une libre variation des règles de préservation de la vérité que nous avons contestée ici.

[65] Voir section 2.1.2.4, note 66.

la pratique : en théorie, la révision d'une constante logique est concevable et s'effectue au sein des recherches en logique abstraite, ou non-interprétée ; dans la pratique, cette révision doit satisfaire des conditions plus sévères et passer notamment le critère de l'usage, même si celui-ci aboutit toujours à un choix relatif dans l'interprétation des pensées d'autrui.

Un faux argument en faveur de la logique classique consisterait à confondre les aspects théorique et pratique évoqués ci-dessus. C'est cette erreur que commet [Quine 1975], lorsqu'il prétend par exemple que celui qui croit à une contradiction ne peut pas comprendre ce qu'il dit. Pour démontrer le lien entre inconsistance et irrationalité, Quine traite la non-contradiction comme une condition préalable à la pratique de l'inférence :

> (...) qu'arriverait-il si quelqu'un rejetait la loi de non-contradiction et acceptait comme vrais à la fois un énoncé pris au hasard et sa négation ? On répond en général que cela nullifierait tout le savoir. Une conjonction quelconque de la forme '$p \wedge \neg p$' implique logiquement tout énoncé que ce soit ; par conséquent l'admission d'un énoncé et de sa négation comme vrais nous forcerait à admettre comme vrai tout énoncé, et c'en serait fait de la distinction entre le vrai et le faux.[66]

Une telle objection est facilement surmontée par le logicien paraconsistant, puisqu'il est possible de construire en théorie des systèmes logiques dans lesquels la contradiction n'implique pas n'importe quoi et, donc, ne rend pas vraie toutes ses conséquences.

Les logiciens paraconsistants ont montré contre Quine que le fait d'admettre une contradiction n'est pas impossible : toute prémisse contradictoire n'entraîne pas la propriété d'explosion selon laquelle croire quelque chose de contradictoire entraîne la croyance triviale à n'importe quel énoncé. Dans cette réponse théorique des langages formels, un mauvais emploi de la logique par les paraconsistants consiste à changer les règles sémantiques du langage formel pour 'bloquer' un résultat indésirable, dont celui de l'explosion ou trivialité consécutive du système formel. En excluant par exemple la règle du syllogisme disjonctif (ϕ ou ψ, or $\neg\phi$, donc ψ),[67] par exemple, ou en modifiant les règles de valuation des énoncés : dans la logique du paradoxe de Priest, la valuation n'est plus traitée comme une fonction binaire stricte (à une seule valeur) de type $f(x) = y$ mais comme une relation de type xRy, en vertu de laquelle certains énoncés peuvent être considérés comme vrais et faux à la fois.[68] Dans ce cas de figure, certaines croyances inconsistantes pourraient être admises comme des

---

[66] [Quine 1978], p. 121.

[67] La preuve de trivialité est la suivante :

1.	$\phi \wedge \neg\phi$	prémisse
2.	ϕ	1., élimination de la conjonction
3.	$\phi \vee \psi$	2., introduction de la disjonction
4.	$\neg\phi$	2., élimination de la conjonction
5.	ψ	3.+4., syllogisme disjonctif
6.	$(\phi \wedge \neg\phi) \rightarrow \psi$	2.+5, modus ponens

Un moyen de bloquer la conclusion triviale de l'étape 6 consiste à rejeter la règle du syllogisme disjonctif, utilisée dans l'étape 5. La citation ci-dessus de Quine reflète son attachement aux règles sémantiques de la logique classique, parmi lesquelles ce type de syllogisme. La discussion entre classiques et non-classiques porte sur la valeur *heuristique* du syllogisme disjonctif, et une révision des conditions de vérité doit être motivée.

[68] On distingue dans les sémantiques non-classiques entre des logiques excédentaires [*glutty logics*], où un énoncé peut être vrai, faux ou vrai et faux, et des logiques lacunaires [*gappy logics*], où un énoncé peut être vrai, faux ou ni vrai ni faux. On obtient à la sortie des systèmes paraconsistants et paracomplets, respectivement. Sur l'utilisation de sémantiques excédentaires ou lacunaires pour traiter des cas d'antinomie, voir [Priest & Armour-Garb 2004].

croyances 'vraies' dans la mesure où les propriétés de consistance (aucun énoncé ne peut être vrai et faux à la fois) ou de complétude (tout énoncé est soit vrai soit faux) deviennent dispensables au sein du langage-objet. Or cette solution, si elle est techniquement correcte, ne semble pas 'intuitivement correcte' puisqu'elle enfreint la distinction comportementale du vrai et du faux. Rien n'empêche le logicien de réviser les conditions de vérité de son système, en théorie, mais la pratique du langage quotidien incite sans doute à plus de prudence ou de charité dans l'emploi des notions de vérité et de fausseté. Le cas des énoncés lacunaires (ni vrais ni faux) n'était pas un argument en faveur d'une révision, avons-nous expliqué plus haut au sujet de la loi de double négation ; le cas des énoncés excédentaires (vrais et faux à la fois) est encore moins plausible puisqu'il ne semble s'accorder avec aucun de nos usages sémantiques familiers et sert plutôt à régler des problèmes techniques d'inconsistance au niveau du métalangage. Règle-t-on le problème des antinomies en modifiant les règles de sémantique, comme semblent l'affirmer toute une série de logiciens non-classiques allant de Bochvar à Priest ?[69] Face à l'argument du menteur d'Epiménide le Crétois, ou en cas d'incapacité d'un locuteur à percevoir la fausseté d'un énoncé contradictoire, les logiciens déviants stipulent que ces énoncés sont valables sur la base de nouvelles stipulations de conditions de vérité, plutôt que d'expliquer cet échec de vérification par des disciplines annexes comme la psychologie ou en distinguant plusieurs aspects de la communication (locutoires et autres).

Notre conclusion sur les fondements épistémologiques de la logique est conservatrice, optant pour une place restreinte de la logique dans la représentation des systèmes de croyance. L'échec d'une inférence classique peut s'expliquer par d'autres disciplines que la logique elle-même ; d'où l'appellation de *charité modérée*, limitée au rôle d'interprétation préparatoire des énoncés d'autrui en termes d'assertions factuelles ou locutoires. Le différend entre classiques et déviants porte sur la ligne de démarcation entre irrationalité et rationalité différente. On parle de logiques chinoises, orientales, zandées ou jainas, afin d'illustrer des formes de rationalité distinctes de la logique aristotélicienne, c'est-à-dire occidentale.[70] Mais bien que Bazhanov ait pris au premier degré la logique imaginaire de Vasil'év en distinguant une rationalité occidentale d'une rationalité russe, aucune expérience cruciale ne permet d'établir l'existence de ces modes d'inférence distincts selon l'appartenance à un schème conceptuel. La place de la logique dans la constitution de la rationalité est au cœur de notre débat sur le pluralisme logique, mais l'impossibilité d'observer une différence de schèmes conduit ainsi à des résultats hypothétiques, tant et si bien que le choix entre le maintien ou la révision des inférences classiques repose avant tout sur un critère heuristique : quel est le prix à payer pour chacune des stratégies conservatrice et déviante ? En ce qui concerne

[69] Sur le traitement sémantique des antinomies, voir notamment la sémantique lacunaire de Bochvar : "On a 3-valued calculus and its application to the analysis of paradoxes of classical extend functional calculus", *Matématičeskij Sbornik* **4**(1938), pp. 287-308 (en russe) ; ainsi que la sémantique excédentaire de Priest : "The logic of paradox", *Journal of Philosophical Logic* **8**(1979), pp. 219-41.

[70] Voir notamment le cas de la logiques zandée, population du Soudan et du Zaïre à laquelle certains logiciens paraconsistants attribuent des croyances inconsistantes sur la base d'un argument de sorcellerie héréditaire : soit on en conclut l'existence d'une logique zandée, par opposition à la logique occidentale ou 'aristotélicienne' ([da Costa & French 1995], [da Costa & Bueno & French 1998] ; soit le raisonnement est accusé d'ambiguïté et reconstruit fidèlement au principe de consistance [Triplett 1988].

l'épistémologie de la logique défendue dans cette section, le choix a porté sur le recours à des hypothèses analytiques et une division du travail dans l'acte d'interprétation des croyances, mais rien ne semble plaider définitivement dans ce sens ou dans celui des logiques déviantes ; seul le critère de l'usage ou enracinement épistémique [*epistemic entrenchment*] des inférences classiques a été exploité ici, avec pour conséquence un maintien complémentaire du principe de charité et une distinction entre les aspects théorique et pratique de la logique : il est possible de construire des structures logiques différentes, en témoignent les systèmes modaux non-normaux, mais le problème des logiques philosophiques porte sur la pertinence de ces structures déviantes lorsqu'elles visent à l'interprétation des phrases des langues naturelles.[71] D'après nous, réviser les mots logiques ne permet pas de mieux *comprendre* les cas de croyance irrationnelle ou inconsistants : cela permet de rendre ces ensembles vrais par calcul, mais c'est précisément cette vérité des ensembles inconsistants que nous avons rejetée au nom de la charité. Le recours à d'autres systèmes de logique ne résout rien, permettant simplement d'exprimer une préférence pour des ensembles de croyances non-classiques mais sans justifier cette préférence pour autant. La comparaison entre la valeur heuristique des opérations classiques et non-classiques est loin d'être simple, voire impossible à déterminer en présence d'une épistémologie naturalisée ; mais notre choix va clairement ici en faveur du conservatisme de Quine et, par voie de conséquence, des logiques modales normales ou extrinsèques.

Pour exprimer notre position plus clairement en ce qui concerne la relation entre vérité, signification et croyance, nous allons revenir et insister dans ce qui suit sur l'aspect *illocutoire* des énoncés de croyance : le lien entre croyance et assertion explicite, mentionné ci-dessus pour garantir l'attribution d'une croyance à autrui, va nous conduire naturellement de la logique épistémique des énoncés à une logique illocutoire des assertions. Peut-on éliminer toute référence ambiguë à des croyances dispositionnelles ou implicites, c'est-à-dire réduire la logique des croyances à une logique des assertions ; ou au contraire, une étude exhaustive de la croyance doit-elle tenir compte de croyances irréductiblement implicites ? L'identité entre croyance justifiée et assertion, évoquée dans le chapitre précédent lors du traitement du paradoxe de Moore et l'interprétation pragmatique de l'introspection positive, dépend de la réponse à cette question.

3.3. Pour une logique modale illocutoire

Une assimilation est proposée entre deux systèmes de logique en apparence distincts : la logique *épistémique*, composée des opérateurs de connaissance K et de croyance B, et la logique *illocutoire*, composée d'opérateurs de force pragmatique. Le résultat est un système de logique modale illocutoire $\mathcal{L}_{f.AS}$, basé sur un opérateur primitif d'assertion. Pourquoi une

[71] [Putnam 1969] opte clairement pour une révision des constantes logiques, comme le montrent les passages suivants : "(...) je pense qu'il est plus probable de dire que la logique classique ne va pas que de dire qu'il y a soit des variables cachées, soit des 'coupures entre l'observateur et le système', etc.'" (p. 232) ; "Ce que le point de vue classique néglige, c'est que l'apriorité de la logique et de la géométrie s'évanouit dès que des logiques alternatives et des géométries alternatives commencent à avoir une application physique sérieuse" (*loc. cit.)*. Or ses remarques portaient sur la logique quantique, qui est encore un autre problème que l'application des lois logiques classiques lors d'une traduction radicale. Il n'est pas question pour nous de prétendre à l'apriorité de ces opérations, mais à leur enracinement. Nous renvoyons en cela à la notion de 'force de la règle' de Bouveresse, utilisée pour décrire l'idée wittgensteinienne de lois logiques nécessaires *et* contingentes.

telle réinterprétation des attitudes épistémiques, et dans quelle mesure peut-on caractériser les modalités épistémiques comme des actes de discours ?

3.3.1 Croyance justifiée et assertion

Après avoir évoqué le problème de [Gettier 1963] et le paradoxe pragmatique de la connaissance de [Zemach 1969], nous avons constaté des relations distinctes entre les notions de vérité et de connaissance, d'un côté, et celles de déclaration de vérité et de connaissance de l'autre : toute connaissance est vraie, $K\phi \rightarrow \phi$, mais toute déclaration de vérité équivaut à une déclaration de savoir, $\ulcorner\phi\urcorner \equiv \ulcorner K\phi\urcorner$. De plus, nous avons concédé qu'une déclaration de savoir est réfutable mais que toute déclaration de ce genre ne peut pas être réfutée *au moment de son énonciation* sans être auto-annulée [*self-defeated*] par la même occasion. Pour cette raison, la clause dite de vérité (de 'véracité') ne vaut à proprement parler que pour les *déclarations* de savoir, et celles-ci correspondent à des actes d'assertion. En conséquence, les axiomes de la logique modale épistémique ne sont réellement valables que lorsqu'ils sont interprétés dans un sens illocutoire. La différence entre connaissance et croyance ne réside pas dans la vérité factuelle du contenu de connaissance, mais dans la présupposition de vérité de la déclaration de type 'je sais'.

Or deux actes illocutoires étaient déjà disponibles avant l'apparition de la logique épistémique, portant également sur des attitudes ou engagements à l'égard de la vérité : l'*assertion*, analogue à la connaissance de la vérité, et la *supposition*, qualifiée de considération chez Frege et assimilée par nous à la simple croyance.[72] Dans la mesure où la différence entre connaissance et croyance n'est que verbale et exprime des degrés d'engagement distincts de la part du locuteur, nous réduirons pour cette raison la notion de connaissance à celle de croyance justifiée, dans le même esprit que les anti-réalistes mais sans aboutir à la conséquence du paradoxe (locutoire) de Fitch.

Nous proposerons de réduire les deux opérateurs de connaissance et de croyance à une seule opération primitive : un opérateur de croyance justifiée ou *assertion*, dans le cadre d'une approche illocutoire des attitudes épistémiques et dans le but de modifier la forme logique des énoncés. Le lien entre connaissance et vérité est donc affaibli en termes d'affirmation illocutoire, ou reconnaissance de vérité : le locuteur s'engage par une assertion sur la vérité de son discours, mais aucune preuve absolue de la vérité du contenu propositionnel n'est à sa disposition dans une théorie de l'information. Une fois le lien logique rompu entre le langage et le monde (suite à notre rejet de la clause de vérité), que devient la logique épistémique ?

Sur la base du problème de Gettier, et suite à nos analyses pragmatiques de l'introspection positive et du paradoxe de Moore, nous allons opter ici pour une interprétation *illocutoire* des attitudes épistémiques en termes de jugement puis d'assertion.[73] A rebours des logiques

---

[72] On pourrait objecter que la notion de *considération* (équivalente à l'acte de saisie de la pensée chez Frege) est totalement neutre et n'indique aucun engagement de la part du locuteur, seulement la considération d'une pensée. Nous expliquerons plus loin le sens que nous donnons à cette considération, assimilée à un engagement minimal et orientée vers la notion de *supposition* de Russell qui, elle, connote une forme de croyance minimale : supposer ϕ, c'est considérer la possibilité de vérité ou de fausseté d'une pensée ; par opposition, asserter ϕ c'est s'engager sur la vérité ou la fausseté de cette pensée.

[73] La remarque suivante de [Frege 1919b] sur le caractère indexical des jugements concorde avec l'analyse du paradoxe de Moore et l'aspect 'quasi-performatif' des modalités épistémiques : "Si le jugement est un acte, il se produit à un moment déterminé, il appartient ensuite au passé. Un acte comporte un acteur, et l'acte n'est pas

modales actuelles, qui tendent vers une combinaison de différents opérateurs (les logiques multimodales),[74] la logique modale illocutoire qui suit ne propose pas d'associer les propriétés de différents opérateurs mais de les réduire tous à un seul. Pour ce faire, nous allons étudier tout d'abord les origines de l'assertion en logique moderne, puis nous établirons les conditions dans lesquelles les attitudes épistémiques peuvent être rattachées à l'assertion frégéenne. Les notions de question et réponse (logique érotétique), de déclaration de connaissance, d'assertion (logique épistémique) et de contenu propositionnel (logique assertorique) seront toutes inscrites au sein d'un formalisme simplifié, dans la lignée de la théorie du jugement constituée dans [Frege 1971]. Nous baserons notre étude historique préliminaire de la notion d'assertion sur plusieurs textes logiques de Russell et Frege recensés dans [Vernant 2003, 2005, 2006].

L'assertion est au jugement ce que l'énoncé est à la proposition : l'assertion est l'expression *épistémique* d'une reconnaissance de vérité, l'énoncé est l'expression *ontique* d'une valeur de vérité. Mais contrairement à une certaine tradition réaliste héritée de Frege, nous affirmerons que, dans un énoncé, ces deux aspects de l'expression sont indissociables car toute proposition n'est que l'expression abstraite d'un jugement particulier. Revenons d'abord sur les termes de la théorie du jugement, avant de défendre cette position nominaliste.

Lorsqu'une sémantique porte sur le contenu des énoncés, elle est de type *locutoire* et décrit un point de vue ontologique ou, plus précisément, un point de vue ontique, c'est-à-dire l'ensemble des énoncés que le locuteur considère comme vrai ou faux en fonction de ses engagements ontologiques. Par exemple, la logique classique est de type locutoire parce qu'elle prétend que, quel que soit le locuteur et l'état du monde, la formule $\neg\phi$ doit être fausse si ϕ est vraie. Lorsqu'une sémantique ne porte plus sur le contenu des énoncés mais sur les *actes* d'énonciation, que l'on appellera affirmations ou assertions, elle est de type *illocutoire* et décrit un point de vue épistémique, c'est-à-dire l'ensemble des actes de discours que le locuteur est susceptible d'effectuer en fonction de ses engagements ontologiques de départ.[75] Les deux types de sémantique portent sur le même objet de discours, l'*énoncé*, mais elles le considèrent d'un point de vue distinct. Par exemple, la logique des assertions spécifiée dans [Searle & Vanderveken 1985] est de type illocutoire puisqu'elle explique que le locuteur doit rejeter $\neg\phi$ s'il asserte ϕ. Mais ces deux actes ne sont pas déterminés sans tenir compte des valeurs de vérité des énoncés, et l'on insistera ainsi sur le lien persistant entre les deux aspects ontique et épistémique des logiques locutoire et illocutoire : l'assertion de ϕ est un acte de nature illocutoire qui décrit un fait locutoire, puisque asserter consiste à considérer quelque chose comme *vrai*. Le lien entre ontologie et épistémologie n'est donc pas tout à fait

---

entièrement connu si l'acteur n'est pas connu" (p. 205 note 1) ; le caractère indexical des jugements concorde avec celui des énoncés.

[74] Un exemple de logique multimodale est donné par A. Costa-Leite : "Fusions of modal logics and Fitch's paradox", *Croatian Journal of Philosophy* (2006), dans lequel l'auteur propose de fusionner les opérateurs de possibilité aléthique et de nécessité épistémique pour créer un nouvel opérateur bi-modal : l'opérateur de connaissabilité. Cet opérateur a pour but de bloquer l'inférence paradoxale de Fitch : $(\phi \rightarrow \Diamond K\phi) \vdash (\phi \rightarrow K\phi)$.

[75] Nous empruntons la distinction ontique-épistémique à [Kearns 2004]. 'Ontique' désigne ici le domaine de référence d'un individu, par opposition au domaine du monde réel ; l'expression doit être soigneusement distinguée du sens que lui donne [Gardiès 1979] : chez ce dernier, les *modalités ontiques* désignent les modalités aléthiques de la nécessité et de la possibilité *de re* (un homme est nécessairement bipède, possiblement chauve).

rompu ici, malgré nos réserves précédentes à l'égard de la clause de vérité $K\phi \rightarrow \phi$; le remplacement de l'expression 'ontologique' par 'ontique' permet de souligner que la vérité du conséquent ϕ exprime simplement le *point de vue du sujet connaissant* sur l'état du monde ('être vrai' et 'tenir pour vrai' sont interchangeables).

Chez Frege comme chez Russell, une telle logique du jugement n'avait pas lieu d'être puisque seule comptait pour eux la dimension ontologique du vrai dans l'édification d'un système logique. Pourquoi ce cloisonnement initial des notions de jugement et d'assertion dans le domaine de la psychologie, et comment interpréter le passage progressif au domaine de la logique, à travers la formalisation progressive des actes illocutoires en général et de l'acte pragmatique d'assertion en particulier ?

Dans [Frege 1999], on trouve référence à deux sortes d'actes de pensée en plus des contenus de pensée : la simple saisie d'une pensée, qu'il appelle encore *contenu jugeable* en 1879 et rebaptisera 'pensée' en 1892, est distinguée du *jugement*, et ces deux expressions décrivent deux attitudes distinctes du sujet à l'égard des pensées qu'il exprime. On retrouve sous d'autres termes cette saisie de la pensée, chez Russell (supposition) et chez Meinong (*Annahme*, assomption), mais avec des positions divergentes sur l'aspect psychologique de l'assertion.[76] Nous avons déjà évoqué précédemment (section 2.1.2.1) la distinction fregéenne entre saisie d'une pensée et jugement ; la différence entre les deux servait à montrer que les propriétés des pensées (leurs valeurs de vérité) sont indépendantes des actes de pensée du sujet, d'où la distinction entre un domaine objectif de la logique, centré sur la définition du vrai, et un domaine subjectif de la psychologie qui ne ferait qu'illustrer des actes de pensée. Par conséquent, le jugement est subalterne à la logique, insiste Frege, n'ajoute rien à son contenu et n'en est que l'illustration *particulière* :

> On suivra parfaitement l'usage si l'on entend par jugement l'acte de juger comme un saut est l'acte de sauter.[77]

Pour maintenir la distance entre logique et psychologie, il faut donc éviter toute confusion entre l'instance générale qu'est la pensée et l'acte particulier de penser : les propositions sont des pensées objectives, des entités abstraites et non linguistiques partagées par plusieurs personnes lorsqu'elles s'expriment par le biais des entités linguistiques que sont les énoncés ; quant aux pensées particulières que chacune de ces personnes manifeste, elles s'expriment à travers des actes d'énonciation particuliers, mais de telles instanciations ne garantissent pas d'elles-mêmes la vérité des pensées (propositions) :

> Pas plus qu'un promeneur gravissant une montagne ne crée cette montagne par son ascension, l'homme qui juge ne crée une pensée tandis qu'il reconnaît cette vérité.[78]

---

[76] [Vernant 2005] explique que, pour Russell, l'assertion n'était pas considérée tout d'abord comme une attitude propositionnelle mais comme une propriété intrinsèque des propositions vraies, puis comme ce qui reste de la proposition lorsque l'on a ôté le terme sujet. La critique de Russell portait sur l'accusation de *psychologisme* contre Frege et Meinong : décrire l'assertion en termes d'attitude, c'est introduire des données psychologiques dans l'explication du terme. Sur l'évolution de la théorie de l'assertion de Russell, voir [Vernant 2005].

[77] [Frege 1919b], p. 205 note 1. Le même risque de confusion entre une faculté (*type*) et son instanciation (*token*) apparaît dans la langue française avec les notions de pensée ou de volonté, notamment. La langue allemande exprime cette distinction en donnant à la faculté une forme nominale (ex : *der Gedanke*, *der Wille*) et à l'acte une forme verbale nominalisée (*das Denken*, *das Wollen*).

Pour distinguer entre la simple expression d'une pensée et la reconnaissance de sa vérité, Frege avait proposé un moyen symbolique : le trait horizontal – du *contenu jugeable* ϕ précède ϕ pour exprimer la simple considération par le sujet : '–César a conquis la Gaule' est un fait propositionnel simplement considéré, lorsque le locuteur ne prend pas position sur la vérité ou fausseté de cette pensée ; le trait vertical ou *barre de jugement* | s'ajoute au trait horizontal pour indiquer que le locuteur donne également son opinion sur la valeur du contenu de pensée : $\vdash\phi$ indique qu'il considère la pensée que ϕ *et* la tient pour vraie, l'énoncé 'César a conquis la Gaule' exprimant cette pensée par un acte d'assertion. Pour séparer simple considération (ou supposition) et assertion, Russell et Frege proposaient d'exprimer l'énoncé simplement considéré sous une forme participiale qu'ils appelaient respectivement *concept propositionnel* et *contenu conceptuel* : 'La conquête de la Gaule par César' est le contenu commun aux deux énoncés 'César a conquis la Gaule' et 'la Gaule a été conquise par César'. Dans le vocabulaire de Frege, une telle forme participiale sert à montrer que deux énoncés grammaticaux peuvent exprimer la même pensée et, donc, la même croyance explicite.[79]

L'idéographie de ce dernier avait pour même but d'éviter les ambiguïtés des langues naturelles, notamment en dissociant symboliquement les aspects sémantiques et pragmatiques du langage : le fait d'*être* vrai concerne exclusivement le premier aspect, tandis que le fait de *tenir pour* vrai ou *considérer comme* vrai se rapporte à la relation entre un énoncé du langage et le locuteur. Pour donner un exemple d'ambiguïté manifeste entre les expressions locutoires et illocutoires, la forme nominale des expressions a tendance à créer une confusion entre pensées et actes de pensée, et l'on exprime plus clairement ces derniers par une forme verbale dans la langue française ('affirmer' et 'nier' indiquent plus nettement l'acte que le contenu de pensée). Mais les expressions françaises confondent couramment les formes nominales et verbales, contrairement à certaines langues voisines telles que l'allemand ;[80] d'où l'intérêt d'un langage formel où exprimer clairement ces différences d'usage par un symbolisme iconique. Dans une langue naturelle, la différence entre considération et assertion se manifeste par plusieurs facteurs pragmatiques tels que le ton de voix, ou l'insistance emphatique sur des parties du discours en vue de produire un effet : c'est l'effet *illocutoire* attendu d'une

---

[78] *Loc. cit.*, pp. 205-6. Ce passage exprime clairement une forme de réalisme sémantique, dans la mesure où Frege voulait garantir ici l'objectivité des pensées. L'anti-réaliste concèdera à Frege que la vérité d'une pensée n'est pas révocable selon qu'elle a un non un porteur (quelqu'un pour la juger) ; il rejettera plutôt l'idée que *toute* pensée quelle qu'elle soit puisse être vraie ou fausse sans porteur.

[79] Comment déterminer si deux énoncés grammaticaux expriment une seule et même pensée ? La variation des termes désignatifs peut changer une pensée en une autre pensée, en raison des problèmes d'opacité référentielle ; en revanche, un changement grammatical tel que le passage de la voix active à la voix passive peut être considéré comme sans effet sur l'identité de la pensée. La pensée ne serait donc pas sensible à la structure grammaticale de son énoncé mais à la *compétence linguistique* du penseur : ses pensées sont identiques à condition qu'il soit capable d'identifier les termes désignatifs ou référentiels utilisés. Hormis ce critère individuel, aucun critère général (désignation rigide, isomorphie intensionnelle, etc.) ne nous semble approprié pour identifier une pensée.

[80] En allemand, l'acte d'affirmation se dit 'Bejahung', littéralement 'action de dire oui' (*afirmato*, en latin), et le contenu affirmatif d'un énoncé se dit 'Behauptung' (*afirmo*, en latin) ; l'acte de nier se dit 'Verneinung', littéralement 'action de dire non', et le contenu négatif d'un énoncé se dit 'Negation'. A noter que l'article de [Frege 1919b] porte à l'origine l'intitulé 'Die Verneinung', ce qui peut prêter à confusion étant donné l'anti-psychologisme de l'auteur. Mais Frege entend par le choix de cette expression que toute négation *suppose* une pensée (il n'y a pas de 'faits négatifs', seules les propositions peuvent être négatives). [Vernant 2003] fait référence également (p. 7 note 32) à un article éponyme de Freud, celui-ci insistant au contraire sur le rôle logiquement fondateur de l'acte psychologique de nier.

énonciation qui permet de 'déduire' si le locuteur a l'intention de dire la vérité, ou s'il exprime simplement un fait sans donner son opinion.

La confusion entre les deux aspects locutoire et illocutoire du langage se manifeste notamment dans les notions d'*affirmation* et de *négation*. Il faut distinguer entre une affirmation et une négation locutoires et illocutoires, selon qu'elles expriment respectivement la qualité d'une proposition (forme affirmative, forme négative) ou l'attitude du locuteur à son égard (affirmer, nier).

L'*affirmation* locutoire renvoie à l'opérateur vérifonctionnel redondant ⊢ (le tiret horizontal – de Frege), auquel cas elle exprime la qualité du contenu jugeable (la forme affirmative : Socrate est Grec).[81] L'affirmation illocutoire est synonyme de l'assertion : affirmer, c'est s'engager sur la vérité d'un énoncé : 'Socrate est Grec', sous-entendu que le locuteur le pense, le croit ou le sait. Elle exprime l'acte de jugement et ne dit rien sur la qualité du contenu jugé. Lorsque l'affirmation est illocutoire, on peut affirmer un énoncé de forme négative tel que 'Socrate n'est pas Grec' lorsque l'on s'engage à son sujet : 'J'affirme que Socrate n'est pas Grec'. L'affirmation locutoire d'un énoncé ϕ sera symbolisée ainsi : $\vdash\phi$; quant à la distinction entre l'affirmation illocutoire d'un énoncé de forme affirmative et négative, elle donnera respectivement les deux formes suivantes : $\vdash\phi$ et $\vdash\neg\phi$.

La *négation* locutoire renvoie à l'opérateur vérifonctionnel $\neg\phi$ (le ─┬─ de Frege), auquel cas elle exprime la qualité du contenu jugeable (la forme négative : Socrate n'est pas Perse) ; la négation illocutoire est synonyme de l'acte de nier ou rejeter la vérité d'un contenu jugeable, auquel cas elle exprime l'opposition du locuteur à l'égard du contenu jugeable et ne dit rien sur la qualité du contenu jugé. Lorsque la négation est illocutoire, il est possible de nier un énoncé de forme affirmative par l'expression 'Socrate n'est pas Perse', sous-entendu que le locuteur ne sait pas ou ne croit pas que Socrate est Perse ; mais la différence entre la négation illocutoire d'un énoncé affirmatif et l'affirmation illocutoire d'un énoncé négatif ne ressort pas clairement dans la langue naturelle : affirmer que Socrate n'est pas Perse et nier que Socrate soit Perse sont deux actes illocutoires exprimés par une seule et même déclaration, 'Socrate n'est pas Perse', ce qui peut porter à confusion lorsque le locuteur qui nie refuse simplement de s'engager sur la valeur de vérité du contenu jugeable. La notion de négation illocutoire recouvre donc à la fois la non-assertion et l'assertion de la négation, et le fait d'ajouter l'expression 'Ce n'est pas le cas que Socrate est Perse' ne changera rien à l'affaire puisque celle-ci peut exprimer aussi bien un désengagement vis-à-vis de l'énoncé affirmatif ϕ qu'un engagement vis-à-vis de sa forme négative $\neg\phi$. Une distinction de ces deux acceptions de la négation illocutoire sera importante plus loin, en ce qui concerne les conséquences formelles des opérateurs illocutoires d'affirmation/assertion et de (dé)négation : si le fait d'asserter ϕ implique le fait de nier $\neg\phi$, en revanche le fait de nier $\neg\phi$ n'implique pas toujours le fait d'asserter ϕ car le locuteur peut s'abstenir de juger la valeur de vérité de ϕ. En d'autres termes : si le fait d'affirmer ϕ implique le fait de croire que ϕ est vrai (en vertu du principe de décitation, admis plus haut), en revanche le fait de nier ϕ n'implique pas toujours le fait de croire que ϕ est faux, d'où une distinction au niveau illocutoire entre 'ne pas être

[81] "La valeur de cette fonction sera le vrai si le vrai est pris comme argument, et le faux dans tous les autres cas" ([Frege 1971] : "Fonction et concept", p. 94). Pour ajouter à la confusion ambiante sur cette notion d'affirmation, [Vernant 2005] signale (p. 12 note 62) que l'opérateur classique d'affirmation était malencontreusement qualifié 'assertium' par Lesniewski.

vrai' et être 'faux'.[82] Nous reviendrons plus bas (section 3.3.2) sur l'ambivalence de la négation illocutoire entre le sens de *nier* et *dénier*, car l'équivalence entre l'assertion de $\phi/\neg\phi$ et la (dé)négation de $\neg\phi/\phi$ n'est valable que dans un sens précis et restrictif de la négation. Lorsque l'affirmation est illocutoire, on peut affirmer un énoncé de forme négative tel que 'Socrate n'est pas Grec' lorsque l'on s'engage à son sujet : 'Je sais que Socrate n'est pas Grec' est l'expression affirmative d'un jugement négatif. La négation locutoire d'un énoncé ϕ sera symbolisée ainsi : $\neg\phi$; quant à la distinction entre la négation illocutoire d'un énoncé de forme affirmative et négative, elle sera symbolisée (pour l'instant) par les deux formes respectives suivantes : $\dashv\phi$ et $\dashv\neg\phi$.[83]

Plusieurs similarités se dégagent entre la notion d'assertion et l'analyse logique multimodale : non seulement les modalités épistémiques, mais d'autres formes de modalités (neustiques) liées telles que la question (en logique érotétique, ou logique des interrogatifs) sont anticipées dans la théorie du jugement de Frege, la seule différence portant sur le traitement symbolique de ces notions. Lorsque Frege compare la simple saisie de ϕ à une question et le jugement consécutif sur ϕ à une réponse, non seulement il trouve un moyen de séparer pensée et valeurs de vérité,[84] mais il devance également les commentaires de Hintikka sur la relation complémentaire entre logique épistémique et logique érotétique : une sémantique des questions (closes) et réponses (oui-non) peut être considérée comme l'expression formelle d'un jeu réglé de questions-réponses entre un locuteur et un interlocuteur, où la réponse à la question s'exprime en termes de 'savoir si' (Sais-tu si ϕ est le cas ? Oui, je sais que ϕ ou que non-ϕ / Non, je ne sais pas si ϕ est le cas).[85] En d'autres termes, la formule épistémique de type $K(\phi \vee \neg\phi)$ ne fait que reformuler en termes modaux la représentation idéographique des deux assertions possibles sur la pensée que ϕ : $\vdash\phi$ ou $\vdash\neg\phi$

---

[82] La langue anglaise rend compte de cette subtilité illocutoire : 'untrue' (non-vrai) n'est pas identique à 'false' (faux), tandis qu'aucune notion intermédiaire de ce genre ne semble poindre dans la langue française entre 'vrai' et 'faux'. Le faux implique le non-vrai, mais la converse n'est pas toujours vraie : c'est plutôt une relation d'*implicature* (implication pragmatique, ou conversationnelle) qui permet de passer de 'ne pas être vrai' à 'être faux' : celui qui *dit* qu'une chose n'est pas vraie est enclin à asserter qu'elle est fausse, bien qu'il puisse vouloir dire également qu'elle ne lui paraît pas plus fausse que vraie. Dans ce dernier cas, le locuteur n'utilisera pas la négation impersonnelle 'ce n'est pas vrai' et modérera son expression : 'je ne dis pas que c'est vrai'. La distinction entre 'vrai', 'faux' et 'non-vrai' s'exprime à travers les trois expressions 'dire le vrai', 'dire le faux' et 'ne pas dire le vrai'. Dans une autre acception illocutoire, 'untrue' suggère également le discours délibérément faux, c'est-à-dire le mensonge : dire quelque chose de 'untrue', c'est dire quelque chose de faux avec l'*intention* de tromper son interlocuteur. Le préfixe privatif 'un' porte alors sur la sincérité du locuteur, qui représente en temps normal la condition préparatoire d'un acte assertif.

[83] On trouve également le symbole inversé du rejet $\dashv$ dans des études syllogistiques de Łukasiewicz et Słupecki. Nous reviendrons dans la section 3.3.3.2 ci-dessous sur les raisons de ce choix symbolique.

[84] Pour indiquer qu'une pensée peut avoir du sens sans être vraie, Frege explique que "l'être vrai ne peut pas appartenir au sens d'une proposition interrogative. Cela serait contraire à l'essence de la question (...) le sens d'une proposition interrogative n'est pas quelque chose dont l'être consiste dans son être vrai. La nature même de la question exige qu'on sépare la saisie du sens et le jugement". ([Frege 1919b], p. 197). L'auteur exprimera cette séparation en disant que l'énoncé a une pensée pour *sens* et une valeur de vérité pour *référence* ; quant au jugement, il exprime la pensée que le locuteur associe à l'énoncé.

[85] Sur le lien entre sémantique épistémique et sémantique des questions-réponses, voir notamment [Hintikka 1976] et [Hilpinen 1988]. Les commentaires de [Frege 1919b] devancent également l'actuelle dialogique inspirée de Peirce, ainsi que la sémantique des jeux (*Game Theoretical Semantics*) de Hintikka. [Vernant 2003] insiste sur l'aspect dialogique des notions pragmatiques d'assertion et de dénégation, rappelant qu'une proposition (grammaticalement parlant) correspond à une pro-position [*proposal*] exprimée par un joueur (le Proposant) au sein d'un dialogue.

(symbolisée $\vdash\!\!\top\!\!\neg\phi$ dans [Frege 1999]), dans lesquelles le contenu de l'assertion est soit affirmatif (ϕ) soit négatif ($\neg\phi$). Si l'interlocuteur n'est pas en mesure de répondre à la question, la négation exprime cette absence de jugement par l'ignorance de l'interlocuteur : $\neg K(\phi \vee \neg\phi)$. Le problème à venir sera la forme symbolique de l'ignorance ou absence de jugement : comment distinguer entre l'absence de saisie de la pensée et l'impossibilité de juger par ignorance, sinon en admettant une forme d'assertion négative : $\neg\vdash\phi$? Nous reviendrons plus bas sur les difficultés symboliques liées à la négation illocutoire, et sur sa correspondance avec les modalités épistémiques (notamment le doute, voir section 1.3.1).

Avant d'en venir à la traduction modale des actes illocutoires, relevons quelques passages montrant que Frege et Russell maintenaient les notions de jugement et d'assertion à l'écart de toute formalisation logique. Le lien apparaît clairement entre juger, tenir pour vrai et croire, lorsque Frege déclare que sa barre de jugement symbolise une croyance explicite du locuteur ou assertion [*Behautpung*] :

> en reconnaissant intérieurement quelque chose comme vrai, nous jugeons et, en exprimant ce jugement, nous faisons une assertion.[86]

Mais dans le même temps, l'Allemand rappelait que, dans la perspective du psychologisme, "la vérité est ramenée à ce que les individus tiennent pour vrais" et qu'il faut entendre par "les lois de la logique, non pas les lois psychologiques de ce qui est tenu pour vrai, mais les lois de ce qui est vrai".[87] Le problème de la reconnaissance de la vérité et son lien avec l'objet d'étude de la logique, la vérité, se distingue pour Frege comme le sens d'un objet se distingue de sa référence. Et puisque la logique s'intéresse uniquement à la référence des pensées, le vrai et le faux, le fait que des sujets tiennent des pensées pour vraies ou pour fausses ne doit pas être le souci du logicien. Or si le jugement n'intéresse pas le logicien, pourquoi l'avoir évoqué dans des articles de logique ?

Nous en revenons à la controverse du psychologisme logique, et aux raisons pour lesquelles nous confondons souvent les aspects logiques et psychologiques du raisonnement ; cette ambiguïté est rappelée notamment par Russell à travers la distinction entre *implication* et *inférence*. C'est à partir de ce point historique que nous allons mener notre retranscription modale du jugement : dans le célèbre article de Lewis Carroll sur le paradoxe de l'inférence,[88] le dialogue entre Achille et la tortue porte sur les conditions dans lesquelles un locuteur est autorisé à inférer une conclusion Z des prémisses A et B : si A et B impliquent Z, suis-je certain de pouvoir inférer Z chaque fois que j'admets A et B ? L'article débouche sur une aporie, une régression à l'infini due au refus de l'un des protagonistes (la tortue) de détacher la conséquence des prémisses sans introduire une étape sans cesse intermédiaire. Pour résoudre cette difficulté, [Vernant 2005] rappelle la distinction faite par Russell entre implication et inférence : "la proposition 'p implique q' asserte une implication, quoiqu'elle n'asserte ni p ni q" ; quant à l'inférence, elle porte sur des propositions assertées où

[86] Frege, *Ecrits Posthumes*, trad. P. de Rouilhan et C. Tiercelin (1999), p. 10.
[87] Voir section 2.1.1, note 6.
[88] L.Carroll : "What the Tortoise said to Achilles", *Mind* **4**(1895), pp. 278-280.

"l'hypothèse vraie d'une implication peut être abandonnée et le conséquent asserté".[89] En d'autres termes, l'aporie de la tortue résiderait dans une confusion commise de sa part entre implication et inférence : il est possible de reconnaître une relation d'implication entre deux termes sans reconnaître la vérité de chacun de ses membres en particulier ; pour surmonter l'objection de la tortue, il s'agit donc de préciser que c'est l'inférence (et non l'implication, à laquelle songe la tortue dans son objection) qui permet de détacher la conclusion Z des prémisses A et B. Le parallèle avec la logique doxastique apparaît clairement, si l'on reprend la formulation iconique des deux notions par [Vernant 2003] (p. 8) : à la différence du *principe d'implication*, dans laquelle une relation conditionnelle est assertée entre deux énoncés quelconques ϕ et ψ

$$\vdash(\phi \rightarrow \psi),$$

et signifie 'p implique q', le *principe d'inférence* décrit une relation de conséquence entre plusieurs énoncés assertés séparément :

$$\vdash(\phi \rightarrow \psi) \text{ et } \vdash\phi, \text{ donc } \vdash\psi,$$

signifiant 'ϕ implique ψ ; or ϕ est vraie ; donc ψ est vraie'. Un moyen d'exprimer logiquement les actes de pensée consiste non pas seulement à symboliser ces actes, mais à les introduire dans le langage objet sous la forme d'opérateurs. Or si nous reprenons la comparaison consacrée par Frege entre juger, croire (explicitement) et asserter, le remplacement du symbole iconique $\vdash$ par un opérateur modal est envisageable. De préférence aux notions de connaissance et de croyance, on pourrait substituer à B et K un opérateur d'assertion A exprimant la croyance justifiée, c'est-à-dire l'engagement du locuteur à l'égard du contenu propositionnel qu'il exprime, ce qui donne pour l'implication et l'inférence les deux énoncés suivants, exprimés dans le système modal **K** de Lewis :

(15) $$A(\phi \rightarrow \psi)$$

et

(16) $$[A(\phi \rightarrow \psi) \wedge A\phi] \rightarrow A\psi.$$

Dans (16), le principe métalinguistique d'inférence est introduit dans le langage-objet de la logique d'assertion et consiste à reformuler la relation d'inférence 'donc' par la relation conditionnelle 'si ... alors'. Compte tenu de notre formulation de l'inférence en termes de modalités, nous reprenons l'analogie relevée par [Hocutt 1972] (voir section 2.2.2) et considérons la logique modale épistémique propositionnelle comme une sorte d'internalisation du processus d'inférence : la conclusion $\vdash\psi$ qui est déduite de l'implication $\vdash(\phi \rightarrow \psi)$ et de $\vdash\phi$ est introduite dans le langage-objet en termes de clôture modale sous conditionnel '$[A(\phi \rightarrow \psi) \wedge A\phi] \rightarrow A\psi$'. La question qui occupera la suite sera celle-ci : peut-on identifier à juste titre les notions illocutoires d'assertion et de considération à des

[89] [Russell & Whitehead 1973], resp. §§ 38 (p. 35) et 18 (p. 16). Après remplacement de l'opérateur K par A, on constate que le paradoxe de la Tortue était déjà présent dans [Zemach 1969] : même si l'on admet en principe l'implication $(K\phi \rightarrow \phi)$, ϕ ne peut être détachée de $K\phi$ que si $K\phi$ a été assertée au préalable. La version de la clause de vérité que Zemach et la Tortue admettraient est donc la clause de véracité, de forme itérée $K(K\phi \rightarrow \phi)$, soit $A(A\phi \rightarrow \phi)$. Quant au principe de détachement, il correspond à ce que [Hintikka 1962] appelait une *implication épistémique* : $[A(A\phi \rightarrow \phi) \wedge A(A\phi)] \rightarrow A(\phi)$, c'est-à-dire 'si j'asserte que ϕ, alors ϕ est vrai ; or j'asserte que ϕ ; donc ϕ est vrai'.

modalités, et sinon, quelles différences irréductibles reste-t-il entre logique des assertions et logique modale ? La traduction que nous proposons pour les notions épistémiques au sein d'un système de logique modale illocutoire suppose que, au contraire, aucune différence irréductible n'interdit l'analogie entre A et $\Box$. La distinction marquée chez Frege et Russell entre implication et inférence est en tout cas préservée par les règles de la logique modale normale. En effet, de même que l'inférence d'un conditionnel n'est pas équivalente à l'inférence *conjointe* ou séparée de chacun de ses membres : $[A(\phi \rightarrow \psi) \equiv (A\phi \wedge A\psi)]$ n'est pas valide, de même la logique modale admet seulement la relation d'équivalence *conditionnelle* $[\Box(\phi \rightarrow \psi) \equiv (\Box\phi \rightarrow \Box\psi)]$ est valide (*si* $\Box\phi$ alors $\Box\psi$).

Précisons toutefois qu'il existe deux acceptions distinctes pour le symbole métalinguistique de la barre verticale, qui désigne tantôt l'acte d'assertion tantôt la relation de déduction. Peut-on assimiler l'un à l'autre ? Lorsque la barre désigne la relation d'inférence, $\vdash\phi$ ne signifie pas simplement que ϕ est considérée comme vraie mais que la formule ϕ est un théorème et que, en tant que telle, elle est *logiquement* vraie. C'est dans ce sens de la barre que fonctionne la règle de nécessitation : les formules non-modales marquées $\vdash\phi$ sont nécessaires lorsque ce sont par ailleurs des thèses du calcul standard des propositions ; les énoncés contingents (vrais ou faux dans certains modèles) sont exclus de ce cas de figure, bien qu'ils puissent être parfois considérés comme vrais et marqués ainsi de la barre de jugement $\vdash\phi$. La barre de jugement $\vdash$ porte donc sur n'importe quel énoncé non-modal et signifie que la formule ϕ qui la suit est simplement considérée comme vraie, sans être forcément valide pour autant. Puisque la signification des deux barres est différente, dans quelle mesure peut-on interpréter la déductibilité et la barre de jugement comme exprimant la même chose ? En vertu du métathéorème de déduction de Herbrand, la relation d'inférence $\phi \vdash \psi$ équivaut à la validité du conditionnel $\vdash(\phi \rightarrow \psi)$ (symbolisée aussi par $\phi \Rightarrow \psi$), et c'est cette relation d'implication qui détermine l'ensemble des déclarations *cohérentes* du locuteur : si un locuteur asserte que ϕ et si ϕ entraîne (a pour conséquence logique) ψ, alors ce locuteur est tenu d'asserter ψ en vertu de cette relation de conséquence logique entre ϕ et ψ. Si par exemple ϕ implique ψ, alors ψ ne peut pas être vrai sans que ψ le soit également. Au niveau des actes de discours de l'assertion, cela implique que l'assertion de $\neg\psi$ ne serait pas *correcte* si le locuteur a d'abord asserté ϕ : le locuteur ne peut pas asserter ϕ, c'est-à-dire considérer que l'énoncé ϕ est vrai, sans asserter ψ par la même occasion, c'est-à-dire sans considérer que l'énoncé ψ est vrai également. En d'autres termes, la barre de jugement exprime l'acte de détachement au sein d'une relation de déduction : si $\vdash(\phi \rightarrow \psi)$, alors de $\vdash\phi$ on infère $\vdash\psi$.

Si l'on introduit l'opérateur d'assertion A dans le langage-objet de la logique modale illocutoire, en remplacement de la barre métalinguistique de jugement, alors nous dirons que si ψ implique ϕ alors si $A\psi$ alors $A\phi$: la relation métalinguistique de déduction équivaut à une relation conditionnelle entre deux assertions. Cette traduction présente un avantage pour la logique épistémique : à la différence de la logique modale épistémique de type locutoire, dans laquelle la règle de nécessitation était considérée comme contre-intuitive, cette règle ne pose plus aucun problème dès lors que la déduction s'exprime en termes d'assertion : $\vdash\phi \Rightarrow \vdash\Box\phi$ s'interprète désormais en termes illocutoires d'assertion *explicite* comme la simple loi d'identité : $A\phi \Rightarrow A\phi$. Dans la mesure où $\Rightarrow$ est encore un signe d'inférence du métalangage, nous pouvons l'internaliser également et réduire la règle de nécessitation à l'autre formule **K-**

valide : $A(A\phi \rightarrow A\phi)$. L'un des principaux inconvénients de la logique modale épistémique, la règle de nécessitation, est donc évité dans notre logique modale illocutoire puisque toute formule bien formée élémentaire est soit assertée, soit supposée, donc déclarée en toute conscience par le locuteur. Une conséquence est qu'aucune thèse ne sera assertée qui ne soit déjà explicitement énoncée. Une autre conséquence, symbolique, est que la nécessitation des lois logiques $\phi = \mathsf{T}$ devient une version particulière de la thèse KK de la logique modale épistémique, que l'on exprime désormais sous la forme $A\mathsf{T} \rightarrow AA\mathsf{T}$. Hormis la règle de nécessitation, le système modal **K** conserve son sens dans notre perspective illocutoire : la règle de distribution est maintenue sous la forme $A\phi, A(\phi \rightarrow \psi) \vdash A\psi$, que l'on peut symboliser également par l'inférence du langage-objet $(A\phi \wedge A(\phi \rightarrow \psi)) \rightarrow A\psi$. Et puisque l'*assertion* d'une implication suppose la *connaissance* de ses termes par le locuteur, une telle règle est parfaitement conforme aux usages de l'assertion ordinaire et permet de maintenir un aspect minimal convenable pour le système **K**. Le problème de l'omniscience logique n'a pas été résolu, il a été dissout par le biais d'une reformulation de la règle de nécessitation.[90]

Peut-on considérer également l'assertion comme une modalité susceptible d'*itération* ? Avant de considérer l'interprétation de A relativement aux systèmes modaux **S4** et **S5**, proposons une première ébauche de ce que pourrait être la syntaxe d'une logique modale illocutoire.

Soit $\mathcal{L}_{f.AS}$ un langage logique de l'assertion et de la supposition, composé d'un ensemble de formules élémentaires : ces formules sont des *actes de discours*, c'est-à-dire des déclarations d'énoncés locutoire ou déclaratifs. Une différence avec la logique modale courante porte sur la nature des atomes propositionnels du langage : l'atome à partir duquel une formule acquiert un sens n'est donc pas une formule de type ϕ mais de type $\ulcorner\phi\urcorner$, où les crochets indiquent un acte de discours qui est soit l'assertion A, soit la supposition S ; la formule simple ϕ est donc syntaxiquement mal formée car elle est incomplète et n'exprime pas un énoncé des langues naturelles, toujours marqué d'une force illocutoire et symbolisé par un acte d'énonciation. L'énoncé n'acquiert son sens que par le biais de son énonciation, donc. Comme dans la logique illocutoire de [Searle & Vanderveken 1985], toute énonciation possède des conditions du contenu propositionnel (consistance des énoncés), des conditions préparatoires (but de l'énonciation), des conditions de sincérité (reconnaissance de vérité), et elle applique des degrés de force illocutoire (exprimés ici par les opérateurs fort et faible) sur un contenu propositionnel de base. La condition de consistance (C.¬) de [Hintikka 1962] est donc maintenue en l'état, à laquelle est ajoutée la condition de sincérité et deux degrés de force illocutoire. Les deux formes élémentaires de $\mathcal{L}_{f.AS}$ sont analogues aux formes modales $\Box\phi$ et $\Diamond\phi$: la modalité forte correspond à l'assertion $A\phi$. En guise de contrepartie de la modalité faible $\Diamond$, la supposition tient une *relation duale* avec l'assertion : celui qui asserte ϕ ou $\neg\phi$ ne suppose pas $\neg\phi$ ou ϕ. Si l'on reprend la remarque de Frege selon laquelle la saisie d'une pensée correspond à une question, la relation duale se confirme : $-\phi$ revient à s'interroger sur la valeur de ϕ (ϕ?), contrairement à $\vdash\phi$ qui équivaut à l'assertion de ϕ (ϕ!) ; une question portant sur ϕ équivaut à l'absence d'affirmation catégorique (d'assertion) au

[90] Une tentative de maintenir et *ré*soudre le problème de l'omniscience logique existe, toutefois : on la trouve dans le projet de 'logique faisable' de [Dubucs & Marion 2003], dont le but est de caractériser les ressources déductives limitées des agents humains dans le cadre d'un anti-réalisme radical et par le biais de logiques non-classiques (intuitionniste, linéaire).

sujet de ϕ, et l'assertion de ϕ équivaut à l'absence de question (exprimant un doute) au sujet de ϕ. On proposera ainsi la relation duale entre une modalité illocutoire forte, $A\phi$, et une modalité illocutoire faible, la supposition $-\phi$ ou $S\phi$. Hormis les atomes de $\mathcal{L}_{f.AS}$, les règles de formation des formules seront identiques à celles d'une logique modale : toutes ces formules se combinent par le biais de connecteurs tels que $\wedge$, $\vee$, $\neg$, $\rightarrow$ et $\equiv$.

Un problème se pose donc sur le statut des formules *non modales* de type $\llcorner\phi$ et $\neg\phi$: peut-on les inclure malgré tout dans $\mathcal{L}_{f.AS}$ en combinant des formules illocutoires et locutoires ? Si l'on considère les *actes* de discours ou énonciations comme les unités de signification minimales, le schéma **T** ($A\phi \rightarrow \phi$) devient une formule dépourvue de sens puisque le conséquent ϕ exprime un contenu jugeable mais non jugé (une formule mal formée). L'ambiguïté est d'ordre sémantique, la même ici que celle évoquée dans la section sur la clause de vérité : dans $K\phi \rightarrow \phi$, le conséquent ϕ sous-entendait que ϕ est vraie dans le monde réel, indépendamment des jugements du locuteur. Dans $\mathcal{L}_{f.AS}$, en revanche, la notion d'*être* vrai est totalement remplacée par celle de *tenir pour* vrai (dans un monde illocutoire, celui du discours) : la notion de vérité n'est pas ontologique mais ontique, c'est-à-dire relative à un ensemble d'engagements ontologiques tenus par un locuteur.[91] Or par analogie avec les syllogismes mixtes qui combinent des formes modales et non-modales, l'utilisation d'une sémantique relationnelle pour $\mathcal{L}_{f.AS}$ donnerait un sens à des formules telles que $A\phi \rightarrow \phi$ ou $\phi \rightarrow S\phi$ si l'on considère toute formule (locutoire ou illocutoire) comme relative à un modèle particulier. La solution est la suivante : une formule non-modale telle que ϕ peut figurer dans une formule pourvue de sens si elle figure dans la portée d'un opérateur illocutoire. Autrement dit, les formules $A\phi \rightarrow \phi$ et $\phi \rightarrow S\phi$ ont un sens parce qu'elles évoquent la vérité de ϕ relativement à un monde possible, c'est-à-dire à un état de choses que le locuteur est en droit de supposer. Dans cette perspective, $A\phi \rightarrow \phi$ peut avoir deux sens distincts, selon que le conséquent ϕ exprime une supposition ou que la supposition porte sur la formule d'ensemble : on obtiendra respectivement soit la formule valide $A\phi \rightarrow S\phi$, conforme à la définition modale $\Box\phi \rightarrow \Diamond\phi$ (*A necesse esse ad posse valet consequentia*), soit la formule $S(A\phi \rightarrow \phi)$ qui signifie que toute assertion est supposée dire la vérité. On peut donc combiner des formules locutoires et illocutoire dans la syntaxe de $\mathcal{L}_{f.AS}$, mais à condition que toutes figurent dans la portée d'au moins un opérateur illocutoire.

Dans le système de [Kearns 1997], présenté en Annexe 2, le principe de formation est similaire puisque toute formule est de type illocutoire et aucune forme locutoire ne peut figurer en dehors de la portée d'un opérateur de force illocutoire. Un point de désaccord existe entre $\mathcal{L}_{f.AS}$ et la logique d'assertion et de supposition de Kearns. Outre son refus des formules modales itérées, sur lequel nous reviendrons plus bas, ce point de divergence concerne les règles de formation des formules : pour Kearns, tout connecteur logique doit figurer dans la portée d'un opérateur de force illocutoire, ce qui veut dire qu'une formule telle que $A\phi \rightarrow S\phi$ est mal formée et doit être remplacée par $A(\phi \rightarrow S\phi)$; dans notre système $\mathcal{L}_{f.AS}$, par contre, le connecteur peut figurer hors de cette portée afin d'exprimer le point de vue d'un

[91] La différence entre les aspects ontologiques et ontiques de la vérité était présentée dans [Hanson 1952] en termes respectifs de déclaration de fait [*statement of fact*] et de déclaration factuelle [*factual statement*]. $A\phi$ désigne une déclaration factuelle et non une déclaration de fait : l'assertion telle que nous l'entendons ici exprime simplement le point de vue du sujet sur un état de choses.

interlocuteur : $A\phi \rightarrow S\phi$ signifie que quiconque asserte ϕ suppose ϕ également. Dans cette perspective métalinguistique, la formule ci-dessus exprime une relation générale entre deux actes de discours mais ne porte pas sur un acte de discours en particulier. C'est là une différence principale avec les règles de formation de la logique illocutoire selon Kearns. Pour le reste, les deux systèmes portent un regard similaire sur les actes de discours et accordent un rôle fondamental à l'aspect illocutoire dans la compréhension des énoncés. Tandis que, dans les travaux de Searle & Vanderveken et Vernant, la logique illocutoire est considérée comme une extension de la logique classique, dans l'approche de Kearns l'aspect illocutoire est premier dans l'ordre des significations du langage et toute formule bien formée exprime un engagement du locuteur au sein d'un acte de discours.[92]

La différence majeure concerne la distinction admise par [Vernant 2003] entre des expressions d'engagement (assertion, dénégation) et de non-engagement (considération). Ni le système de Kearns, ni $\mathcal{L}_{f.AS}$ n'admettent cette distinction : dans ces deux logiques d'assertion et de supposition, seules les formules gouvernées par un opérateur de force illocutoire sont considérées comme des formules complètes ou bien formées, donc même les opérateurs 'locutoires' d'affirmation $\llcorner$ et de négation $\neg$ sont traités comme des indicateurs de force illocutoire : l'*affirmation locutoire* de ϕ est interprétée comme le fait de *supposer* ϕ *vraie* et représente une force illocutoire de degré inférieur à l'affirmation de la vérité (l'assentiment, ou assertion positive) de ϕ, symbolisée $\vdash\phi$; la *négation locutoire* $\neg\phi$ est interprétée comme le fait de *supposer* ϕ *fausse* et représente une force illocutoire de degré inférieur à l'affirmation de la fausseté (le dissentiment, ou assertion négative) de ϕ, symbolisée $\vdash\neg\phi$. La distinction entre locutoire et illocutoire n'a plus de sens dans cette logique modale, car tout énoncé exprime un moindre engagement de la part du locuteur. Il n'y a pas d'énonciation neutre (ni positive ni négative), contrairement à la notion frégéenne de 'simple' considération ou saisie de la pensée, car chaque opérateur désigne la conviction manifestée par l'expression d'un énoncé. La traduction russellienne de la notion de simple considération ou saisie de la pensée en termes de 'supposition' conviendra mieux, ici, pour faire ressortir l'aspect illocutoire contenu dans la supposition vraie $\llcorner$ et la supposition fausse $\neg$. Notre position consiste à traiter le *signe* ϕ comme une fonction insaturée, par opposition au contenu jugeable $-\phi$ et au jugement $\vdash\phi$ qui lui donnent un *sens* ; ces deux derniers cas ne sont pas neutres mais expriment deux forces illocutoires distinctes lors de l'énonciation de ϕ. Une telle lecture a un effet sur l'interprétation des opérateurs 'locutoires'. Dans l'énoncé négatif 'Socrate n'est pas Perse', par exemple, la forme logique est $\neg\phi$ mais ce n'est pas le signe lui-même qui est négatif. La combinaison $\neg\phi$ de l'opérateur $\neg$ et du signe ϕ symbolise un énoncé *faiblement nié* : la négation n'exprime pas une caractéristique propre à la fonction propositionnelle ϕ mais fait de cette fonction un énoncé, et cet énoncé exprime en l'occurrence un désaccord du locuteur vis-à-vis du contenu de ϕ. Etre vrai et être faux ne sont donc pas les propriétés du signe énonciatif mais de son énoncé ; or un énoncé est l'expression du point de vue d'un locuteur ; donc être vrai et être faux sont des expressions manifestant les points de vue d'un locuteur. Une autre façon de dire ceci est que la vérité et la fausseté ne sont pas inhérentes ou

[92] Parlant de Searle et Vanderveken, Kearns explique qu'"ils conçoivent la logique illocutoire comme un supplément ou un appendice de la logique standard, et ils se concentrent sur des principes/lois très généraux qui caractérisent des actes illocutoires en tout genre. Par opposition, j'entends la logique illocutoire comme un thème très exhaustif qui comprend la logique standard comme partie à part" ([Kearns 2004], p. 54).

'contenues dans' une phrase déclarative mais expriment le point de vue externe du locuteur sur l'état de choses dont parle la phrase. Puisque la fonction ou intention communicative d'un énoncé (ou phrase déclarative) est de décrire un état de choses, donc de dire une vérité, la négation consiste ainsi à exprimer pour le locuteur son opposition à l'égard de la vérité de ϕ. Une meilleure traduction de l'énoncé 'Socrate n'est pas Perse' serait rendue par l'expression 'Je ne pense pas que Socrate soit Perse', qui éviterait les ambiguïtés et implications ontologiques provoquées par des formulations *im*personnelles telles que 'Socrate n'est pas Perse' ou 'Ce n'est pas le cas que Socrate est Perse'. Quant à l'expression voisine 'Je suis certain que Socrate n'est pas Perse', sa forme logique serait cette fois-ci l'assertion négative de type $\vdash\neg\phi$.

Est-il préférable de rejeter notre lecture totalement illocutoire des actes de discours, qui inclut aussi bien les expressions modales (exprimant des attitudes) que les expressions non-modales ou classiques (n'exprimant pas d'attitudes) ? Pourquoi tout énoncé devrait-il être l'expression d'une modalité illocutoire? On pourrait nous objecter que, après tout, l'expression $\neg\phi$ n'exprime aucune attitude explicite et que, dans la mesure où sa forme logique ne contient ni la barre horizontale du jugement ni la barre verticale de la simple considération, il serait abusif d'assimiler n'importe quel énoncé à l'expression d'une attitude, d'un point de vue. Nous répondrions que cette attitude n'est certes pas manifeste mais présente dans l'énoncé de façon latente : contrairement à l'argument selon lequel un énoncé de type $\neg\phi$ serait neutre et n'exprimerait aucune attitude, il semble que la forme négative $\neg\phi$ soit trompeuse et résulte d'une déformation du symbolisme initial de Frege.

Dans son 'langage formulaire', on distingue en effet deux types de contenus jugeables. Si ϕ symbolise un contenu conceptuel, alors les deux contenus jugeables possibles sont l'affirmation $-\phi$ (symbolisation moderne : $\llcorner\!\!-\phi$) et la négation $-\!\!\top\!\!-\phi$ (symbolisation moderne : $\neg\phi$), où la négation est indiquée par le petit tiret vertical $_{|}$. Or ce tiret est attaché à la partie inférieure de la barre de contenu, la barre horizontale de $-\phi$, donc il porte sur un contenu jugeable. La symbolisation moderne de la négation locutoire par $\neg\phi$ est trompeuse dans la mesure où elle ne montre pas aussi clairement le lien entre négation et contenu jugeable ; dans ce qui suit, nous garderons la symbolisation moderne $\neg$ de la négation, mais sans oublier qu'elle porte sur une supposition et qu'une supposition constitue déjà pour nous un acte de discours effectué par un locuteur.

Par conséquent et si l'on admet notre distinction entre deux types de force illocutoire, l'une exprimant la supposition symbolisée par la barre de contenu $-$, l'autre exprimant l'assertion symbolisée par la barre de jugement $|$, alors la séparation catégorique entre le *sens* et la *référence* d'un énoncé n'est plus valable : dire que la référence d'un énoncé, sa valeur de vérité, n'est pas de même 'nature' que son sens est une chose irrecevable dans notre approche illocutoire, si une telle différence de nature consiste à distinguer dans un énoncé la valeur de vérité qu'il dénote et la *proposition* qu'il exprime. En effet, nous avons renoncé à l'usage d'une entité extra-linguistique telle que la proposition et nous en sommes expliqués plus tôt (voir section 3.2), de sorte que les seuls contenus conceptuels dont nous disposons désormais dans $\mathcal{L}_{f.AS}$ sont des *énoncés*. Nous avons également prétendu plus tôt (section 2.1.2.1) que la proposition correspond à l'expression de ce que Frege appelle le 'sens' de l'énoncé : 'l'Etoile du Matin est une planète' et 'l'Etoile du Soir est une planète' représentent deux propositions

distinctes. Mais si notre approche illocutoire des énoncés est recevable, la distinction entre le sens d'un énoncé (la proposition qu'il exprime) et sa référence (la valeur de vérité qu'il dénote) se traduira dans notre vocabulaire nominaliste par la distinction entre la référence d'un énoncé dans *un* contexte de discours et sa référence dans *tous* les contextes de discours. La proposition est assimilable à un énoncé analytique, c'est-à-dire un énoncé dont tous les locuteurs s'accorderaient sur le sens, par opposition aux énoncés contingents dont les locuteurs peuvent ne pas s'accorder sur le sens (pour des raisons liées aux circonstances d'énonciation, à l'opacité référentielle de ses termes désignatifs, etc.). Nous reprenons ainsi à notre compte l'idée de Hintikka selon laquelle *le sens* (des énoncés) *est la référence multiple* (de leurs énoncés). La supposition $\neg\phi$ (ou $\neg\top\neg\phi$) signifie que l'énoncé ϕ est supposé vrai (ou faux) par son locuteur ou auteur, c'est-à-dire qu'il considère ϕ 'simplement' comme vrai (ou faux) dans au moins un cas ou une situation possible ; l'assertion $\vdash\phi$ (ou $\vdash\neg\phi$) signifie que l'énoncé ϕ est affirmé vrai (ou faux) par l'auteur de l'énoncé, c'est-à-dire qu'il le considère comme vrai (ou faux) dans tous les cas ou situations possibles. La relation logique entre la barre d'assertion et la barre de déductibilité est plus claire, ici : un théorème est considéré comme toujours vrai *de juro* mais peut être considéré comme faux *de facto* par certains locuteurs négligents ; la barre de déductibilité exprime une sorte d'assertabilité impersonnelle, tandis que la barre de jugement incarne une assertion personnelle produite effectivement par un ou plusieurs locuteurs. En somme, la proposition représente l'entité idéale que chaque locuteur saisirait en toute transparence et dont il pourrait déterminer la valeur de vérité sans peine ; les énoncés arithmétiques de base, les définitions analytiques ou les règles logiques élémentaires se rapprocheraient plus ou moins de ces propositions idéales, puisque chaque locuteur semble s'accorder sur la vérité de '2 + 2 = 4', de 'Un célibataire est un homme qui n'est pas marié' ou de 'Si ϕ, alors ϕ ou ψ'.[93] Mais dans la mesure où rien ne prouve qu'une situation dans laquelle ces énoncés serait supposés faux soit impossible, la proposition est rejetée du côté des idées platoniciennes qui n'ont pas cours dans notre approche empiriste, illocutoire du langage.

Trois conséquences formelles résultent de ce traitement purement illocutoire des actes de discours.

Premièrement, si l'on interprète le jugement de Frege comme un jugement fort ou une croyance forte, alors les contenus juge*ables* $\llcorner\phi$ et $\neg\phi$ deviennent dans $\mathcal{L}_{f.AS}$ des contenus 'faiblement jug*és*' : elles expriment des suppositions, dont le degré d'engagement est moins fort que celui des assertions de type $\vdash\phi$ et $\vdash\neg\phi$. En d'autres termes, l'affirmation locutoire $\llcorner\phi$ exprime la supposition de vérité $S\phi$, et la négation locutoire $\neg\phi$ exprime la supposition de fausseté $S\neg\phi$. Supposer ϕ revient à poser formellement ϕ en guise de prémisse, sans prendre

[93] Un énoncé toujours vrai pour n'importe quel locuteur serait en quelque sorte nécessaire, c'est-à-dire vrai dans tous les mondes *métaphysiquement* possibles. Or nous ne savons pas comment interpréter ce type de scénario et nous en tenons à des mondes épistémiquement possibles, relatifs aux attitudes épistémiques des locuteurs dans le cadre de leurs actes de discours. La nécessité illocutoire A concernant l'assertion d'un et un seul locuteur, les conditions dans lesquelles une telle nécessité pourrait s'imposer à tous les locuteurs sont supposées insaisissables. Dire le contraire reviendrait à admettre les vérités analytiques, la synonymie et l'hypothèse de l'identité de signification entre différents énoncés. Le cas des annonces publiques et des connaissances communes donne un traitement empirique qui conviendrait en remplacement des notions analytiques, puisqu'elles proposent de *construire* des assertions intersubjectives par le biais de discours au sein d'une communautés de locuteurs.

parti sur la question de sa vérité ou sa fausseté matérielle ; le seul engagement est d'ordre conditionnel, lorsque le locuteur s'engage à supposer d'autres choses déductibles sur la base de sa supposition préalable de ϕ. Puisque la formule négative locutoire $\neg\phi$ se traduit dans $\mathcal{L}_{f.AS}$ par la supposition de fausseté de ϕ, la symbolisation de $S\neg\phi$ qui en résulte est trompeuse dans la mesure où le signe de négation $\neg$ apparaît également dans l'assertion $A\neg\phi$ et donne ainsi l'apparence de fonctionner comme un opérateur locutoire, indépendant de l'acte de discours correspondant. Mais si l'on admet que toute expression a une teneur illocutoire, négation et supposition de fausseté se confondent symboliquement sous la même forme $\Diamond\neg\phi = S\neg\phi$. La négation $\neg\phi$ sert donc à indiquer l'assignation de fausseté à l'énoncé ϕ par un locuteur ; et bien que le vrai et le faux désignent une relation sémantique objectivement définie entre un énoncé et ce qu'il exprime, cet opérateur de négation ne doit pas être dissocié de l'opération illocutoire sans laquelle il ne ferait pas sens.

Deuxièmement, notre interprétation illocutoire du contenu jugeable suscite une réflexion en retour sur le symbolisme hérité de l'idéographie : si | et – symbolisent deux actes illocutoires distincts, alors l'assertion positive $\vdash\phi$ semble signifier que le locuteur *asserte qu'il suppose* ϕ, symbolisé dans $\mathcal{L}_{f.AS}$ par la formule modale itérée $AS\phi$. Ainsi, l'assertion d'un énoncé négatif signifierait l'assertion d'une supposition de fausseté d'une phrase déclarative. Dans ce sens, la symbolisation idéographique de la négation est préférable à sa contrepartie moderne $\neg$ parce qu'elle permet d'y mettre en évidence à la fois la supposition – et la négation pure $_{|}$. Tout énoncé négatif doit-il être supposé en même temps qu'il est asserté, ou ne pourrait-on pas admettre une assertion sans supposition ? La représentation idéographique de cette éventualité donnerait le symbolisme $|\ _{|}\ \phi$, sans barre de contenu. Pareille symbolisation est dépourvue de sens pour Frege, parce qu'elle reviendrait à admettre la possibilité d'une négation sans pensée. Mais une autre manière d'interpréter dans $\vdash\neg\phi$ la juxtaposition des barres verticale et horizontale consiste à dire que l'assertion *présuppose* la supposition : asserter la fausseté de ϕ, c'est supposer dans le même temps la fausseté de ϕ puisque l'on ne peut asserter que ce que l'on considère au préalable. En termes modaux, la forme modale du jugement $\vdash\phi$ ne donnerait donc pas la formule $AS\phi$ ci-dessus mais la relation conditionnelle $(A\phi \rightarrow S\phi)$. $A\phi$ implique *a fortiori* $S\phi$, conformément à notre hypothèse précédente voulant que $\neg\phi$ et $S\neg\phi$ soient synonymes.[94] On ne peut pas asserter sans supposer *ipso facto*, et la juxtaposition des deux barres est traduisible dans $\mathcal{L}_{f.AS}$ sous la forme d'une relation conditionnelle plutôt que d'une itération modale. La distinction entre possibilité de vérité et détermination de vérité se retrouve exactement dans la distinction entre supposition et assertion, parallèle à celle entre vérité formelle et vérité matérielle mais qui ne doit pas faire oublier la production d'un acte de discours dans chaque cas. Oublier cette

[94] Une meilleure symbolisation de la négation est donnée par [Wiredu 1975] : lorsque la formule ϕ représente un énoncé, $\neg\phi$ symbolise alors une assignation de fausseté à la fonction propositionnelle ϕ et correspond au jugement de second ordre $F\phi$ ('l'énoncé ϕ est faux'), par opposition à l'assignation de vérité $T\phi$ ('l'énoncé ϕ est vrai') ; les formules $\neg\phi$, $S\neg\phi$ et $F\phi$ signifient donc la même chose, de même que pour ϕ, $S\phi$ et $T\phi$. La symbolisation de l'acte *unique* de supposition de fausseté par *deux* symboles S et $\neg$ est trompeuse, d'où l'avantage des opérateurs de [Wiredu 1975] qui permettent de signifier tout à la fois la supposition de S et la fausseté de $\neg$. Malgré tout, nous maintiendrons le symbolisme de $\mathcal{L}_{f.AS}$ non seulement pour rester fidèle à la représentation idéographique de Frege en termes de contenus jugeables – et de jugements $\vdash$, mais aussi parce que Wiredu ne distingue pas l'assignation fable (simple supposition) de l'assignation forte, ou justifiée (assertion).

dimension permettra d'expliquer plus bas des notions logiques telles que la dénégation et dont Frege ne pouvait ni ne voulait rendre compte, pour la raison qu'elles lui semblaient soit superficielles soit nuisibles à la discipline logique en tant que telle.

Troisièmement, le symbole choisi pour l'opérateur vérifonctionnel de l'*affirmation* n'est peut-être pas le plus judicieux qui soit : que ce soit notre symbole ⊢ utilisé jusqu'ici (rotation de la négation ¬ à 180°) ou le symbole ⊣ (rotation de la négation ¬ à 90°), choisi par ailleurs par Vernant,[95] aucun des deux ne semble rendre justice à l'origine historique de l'opérateur de l'affirmation : si l'opérateur vérifonctionnel de la négation ¬ est, comme nous le pensons, un produit dérivé de l'idéographie frégéenne, alors l'affirmation devrait être symbolisée de préférence par ⌐ (rotation de la négation ¬ à 360°), et ce pour la raison suivante. Frege explique ainsi son choix du symbolisme de la négation par ⊤–ϕ :

> La partie du trait horizontal qui se trouve à la droite du trait de la négation est le trait de contenu de [ϕ] et la partie qui se trouve à la gauche du trait de négation est, par contre, le trait de contenu de la négation de [ϕ].[96]

En détachant la partie droite de la partie gauche du trait de la négation, on obtient ainsi les signes symétriquement opposés ¬ et ⌐ ; et puisque le trait de contenu de ϕ dans –ϕ signifie pour nous l'acte de supposer vrai l'énoncé symbolisé par le signe énonciatif ϕ, la partie droite ⌐ de ⊤– devrait être considérée comme l'expression de l'opérateur d'affirmation. Par souci de clarté dans l'exposition, nous maintiendrons toutefois le symbole ⊢ mais garderons à l'esprit la parenté entre négation idéographique, négation locutoire et supposition illocutoire.

La raison pour laquelle Kearns et nous-mêmes traitons tous les *énoncés* comme des actes d'énonciation et tous les opérateurs logiques comme des forces illocutoires tient à une raison d'ordre sémantique, en vertu de laquelle le sens d'un énoncé est indissociable de l'*acte* de discours qui l'accompagne :

> Si nous abordons un système de logique dans une perspective des actes de discours, un langage logique artificiel n'est pas vraiment un langage. En effet, nous ne parlons pas, n'écrivons pas ou ne pensons pas avec des expressions du langage logique. A la place, les expressions du langage logique servent de *représentations* d'actes de langage. Et l'explication sémantique du langage

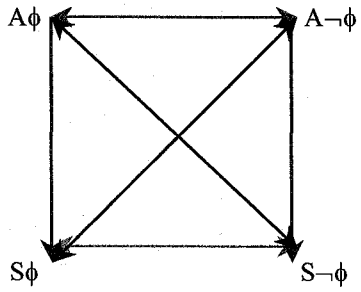
---

[95] Dans *Introduction à la logique standard* (p. 30 note 21) Vernant explique qu'il propose ce symbole "comme inverse de la négation" et fait référence à la symbolisation confondante de l'affirmation locutoire chez Lesniewski : ⊣ ϕ (qui symbolise chez Łukasiewicz le rejet, c'est-à-dire notre dénégation illocutoire). Le choix de ⊢ pour symboliser l'affirmation locutoire, c'est-à-dire notre supposition illocutoire, aurait pour autre avantage de justifier le symbolisme des opérateurs d'assertion et de rejet : l'assertion ⊢ symbolise la juxtaposition de la barre de jugement | et de la supposition de vérité ⌐ ; quant au rejet ou dénégation ⊣, il symbolise la juxtaposition de la négation ¬ et de la barre de jugement |, mais il a alors le tort de placer la négation avant la barre de jugement et d'assimiler alors le rejet à la non-assertion ¬A. Une juxtaposition dans l'ordre inverse ne serait pas plus judicieuse, puisqu'elle assimilerait la dénégation à l'assertion négative A¬.

[96] [Frege 1999], p. 24. Notre interprétation peut-être cavalière de la négation n'est pas étayée par l'histoire de la logique : le symbole ¬ a été introduit la première fois par Heyting en 1929 et représentait la négation *intuitionniste*. Cela étant, la ressemblance accidentelle entre les symboles ¬ et ⊤– est étonnamment cohérente de notre point de vue : si l'on regarde la négation idéographique, la supposition négative (S)¬ semble indépendante de la barre de contenu –.

logique est en réalité une explication des actes de langage représentés par des énoncés du langage logique.[97]

Sur la base de la distinction entre deux degrés d'engagement lors d'un acte de discours, et partant de l'idée novatrice selon laquelle tout énoncé des langues naturelles exprime un acte de discours comportant une force illocutoire, on peut réinterpréter la simple considération de Frege comme une supposition de vérité, de forme S et analogue à la modalité faible $\Diamond$; les fonctions de vérité $\llcorner\phi$ et $\neg\phi$ deviennent des modalités illocutoires faibles $S\phi$ et $S\neg\phi$, en complément des modalités illocutoires fortes $A\phi$ et $A\neg\phi$. On peut représenter ces modalités dans le carré des oppositions suivant :



Opérateurs :
$A\phi$: assertion de ϕ
$S\phi$: supposition de ϕ

Règles de dualité :
$A\phi \equiv \neg S\neg\phi$
$Cp \equiv \neg A\neg p$

Relations d'opposition :
$A\neg\phi$ et $S\phi$, $A\phi$ et $S\neg\phi$ sont contradictoires ;
$A\phi$ et $A\neg\phi$ sont contraires ;
$S\phi$ et $S\neg\phi$ sont subcontraires ;
$S\phi$ est subalterne à $A\phi$, $S\neg\phi$ est subalterne à $A\neg\phi$.

Sans entrer dans une axiomatique détaillée de $\mathcal{L}_{f.AS}$, une analogie sommaire peut déjà être faite avec les modalités *itérées*. Contrairement à tous les auteurs cités ici, l'itération des forces illocutoires semble plausible dans la mesure où des actes de discours sont souvent cités eux-

[97] [Kearns 2006], p. 4. En d'autres termes, l'énonciation représente l'unité de signification minimale dans $\mathcal{L}_{f.AS}$. C'est là une différence avec les logiques illocutoires décrite dans [Vernant 2003, 2005] : dans celles-ci, la modalité 'neutre' $\llcorner\phi$ désigne un énoncé locutoire ϕ réductible à son signe énonciatif, lequel constitue déjà une unité minimale de signification puisqu'il est vrai ou faux 'par lui-même', sans acte de discours à l'appui. L'avantage de l'approche purement illocutoire de Kearns est qu'elle permet entre autres l'analogie entre assertion et modalité forte, supposition et modalité faible. [Hintikka 1962] cite également (note 4 p. 6) un argument selon lequel ce serait effectivement l'acte d'énonciation ou déclaration [*statement*] qui constitue l'élément premier de la signification ; voir Y. Bar-Hillel, "New Light on the Liar", *Analysis* **18**(1957), pp. 1-6. En revanche, la sémantique de Kearns n'est pas relationnelle mais algébrique : elle ne donne pas de sens pour cette raison aux formules modales itérées, que l'auteur refuse de toute façon.

mêmes dans le contenu jugé : 'j'affirme que ϕ', 'oui, je le pense' sont deux exemples de jugements dans lesquels un signe d'engagement illocutoire est exprimé. 'Il est vrai qu'il est faux' ou 'il est faux qu'il est vrai', deux expressions ambiguës bien que pourvues de sens,[98] n'auront pas de contrepartie formelle tant qu'une itération de ce genre n'est pas admise. D'autre part, accepter l'itération des modalités illocutoires nous permettra de réviser plus loin la notion de dénégation et de lui attribuer une définition formelle précise.

On peut admettre la réduction des modalités itérées, conformément au schéma **4** : asserter que ϕ c'est asserter que l'on asserte ϕ, dans la mesure où une déclaration est un acte de discours conscient et susceptible d'être reproduit à volonté ('je suis innocent ; oui, j'affirme que je suis innocent'). Et inversement, d'où la loi d'équivalence $A\phi \equiv AA\phi$. Nous appuyons cette loi de réduction sur le principe de décitation ainsi que sur l'analogie entre asserter, croire et tenir pour vrai : si asserter ϕ revient à affirmer que ϕ est vraie, la théorie déflationniste de la vérité nous permet d'inférer l'énonciation $\ulcorner\phi\urcorner$ à partir de l'énonciation $\ulcorner A\phi\urcorner$; or puisque $\ulcorner\phi\urcorner$ et $\ulcorner A\phi\urcorner$ désignent ici des actes de discours (d'où l'usage des crochets), l'inférence de $\ulcorner A\phi\urcorner$ à partir de $\ulcorner\phi\urcorner$ donne la formule d'itération illocutoire $A\phi \rightarrow AA\phi$, déjà évoquée dans l'analyse pragmatique du paradoxe de Moore (ses versions doxastique et épistémique étant réduites en termes illocutoires).

Les autres analogies modales envisageables sont notamment les théorèmes B-valides et **S5**-valides : $(\phi \rightarrow AS\phi)$ et $(S\phi \rightarrow AS\phi)$, respectivement. La première formule pose le problème de la combinaison des formules locutoires et illocutoires au sein de $\mathcal{L}_{f.AS}$; si l'on remplace de nouveau ϕ par la supposition affirmative $S\phi$, **B** devient **5**. Plus importante, la contrepartie illocutoire de **S5** suppose en vertu de la dualité entre S et $\neg A\neg$ que toute personne qui n'asserte pas ϕ est susceptible d'asserter cette absence d'assertion : $\neg A\neg\phi \rightarrow A\neg A\neg\phi$, simplifiée sous la forme $\neg A\phi \rightarrow A\neg A\phi$ (après substitution de ϕ à $\neg\phi$). Cette dernière formule peut sembler acceptable dans un certain sens, celui dans lequel une non-assertion présuppose ou s'accompagne tout au moins d'une supposition de départ : étant donné que le fait de supposer ϕ implique *pragmatiquement* que le locuteur n'est pas en mesure d'asserter ϕ, alors $\neg A\phi \rightarrow A\neg A\phi$ est un théorème illocutoire valable sous cette condition, à savoir que celui qui ne s'engage pas sur ϕ a au moins une raison de ne pas s'engager. L'argument est le même ici que dans [Blanché 1968], lorsque l'auteur explique que le fait de *dire* 'je crois que ϕ' implique pragmatiquement que le locuteur ne sait pas si ϕ

[98] A l'instar des valuations algébriques (vrai, faux, indéterminé, impossible, inconnu, invérifié, etc.), les valuations itérées ci-dessus prêtent à confusion parce qu'elles donnent à la notion de vérité un sens à la fois ontologique (parlant d'un état de choses) et ontique (parlant d'un jugement *sur* un état de choses) ; comme nous l'avons précisé dans la section consacrée au principe de charité, notre acception de la notion de vérité est factuelle (ontologique) et concerne uniquement la relation 'de premier ordre' entre un énoncé et l'état de choses qu'il exprime ; pour cette raison, nous suivrons la suggestion de Kearns consistant à distinguer plusieurs noms de valuation en plus des valeurs de vérité et de fausseté : une valuation sémantique {T,F} exprime la valeur de vérité des *énoncés*, et une autre valuation sémantique valuation {+,–} (ou {1,0}) exprime la valeur de correction des *énonciations* : 'c'est le cas que cet énoncé est vrai' , ou 'ce n'est pas le cas que cet énoncé est vrai'. Cette distinction terminologique corrige l'erreur qu'il y a à parler de vérité ou de fausseté pour les énonciations, qui sont des jugements de second ordre portés sur des énoncés. Nous souscrivons ainsi à la thèse de [Suszko 1975] selon lequel il y a deux valeurs fondamentales en logique : la valeur *désignée* ou la (ou les) valeur(s) *non-désignée*(s), contrairement à l'impression laissée par les matrices multivalentes et leurs valeurs *algébriques* (voir Chapitre 1, note 56).

(contrairement à la thèse *locutoire* de l'entraînement : $K\phi \rightarrow B\phi$) ;[99] de même, ici, le fait de simplement supposer ϕ implique pragmatiquement que l'on n'asserte pas que ϕ. Sous cette interprétation, la formule $S\phi \rightarrow AS\phi$ ou $\neg A\phi \rightarrow A\neg A\phi$ est donc valable. Une telle présupposition de ce genre rappelle le traitement du paradoxe de Moore dans [Hintikka 1962], lorsque la non-assertion de ϕ est un fait explicitement *déclaré* par le locuteur ; dans ce cas de figure, le locuteur a évidemment conscience de son abstention, il suppose que $\neg\phi$ peut être le cas et est en mesure d'asserter qu'il n'asserte pas ϕ. En revanche, une telle interprétation du schéma **5** devient inacceptable si la non-assertion de ϕ ne présuppose pas la supposition conjointe de $\neg\phi$ et admet également une *absence* complète de supposition : $\neg A\phi \wedge \neg S\phi$. Dans ce cas de figure, le locuteur ne saisit ni ϕ ni $\neg\phi$ par la pensée et n'est donc pas en mesure d'asserter quoi que ce soit, contrairement à la règle de dualité modale suggérée plus haut : $\neg A\phi \equiv S\neg\phi$. En d'autres termes, l'analogue illocutoire de **S5** est intuitivement acceptable sous sa première forme $S\phi \rightarrow AS\phi$, lorsque l'absence d'assertion équivaut tout au moins à la présence d'une supposition, mais elle ne l'est pas sinon. La seconde éventualité est sous-entendue en logique modale épistémique, à travers le rejet de l'introspection négative : la formule $\neg K\phi \rightarrow K\neg K\phi$ est couramment rejetée parce qu'elle ne présuppose pas que l'agent doive avoir *conscience* de tout ce qu'il ne sait pas. Or dans le cadre d'une logique illocutoire, où chaque formule bien formée constitue une déclaration explicite, comment ne pas avoir conscience de ce que l'on n'asserte pas si toute formule doit faire l'objet d'une déclaration explicite ? Dans la mesure où les règles de formation de $\mathcal{L}_{f.AS}$ stipulent que seules les énonciations sont des formules bien formées, **S5** est correct puisque toute non-assertion de ϕ présuppose la supposition du 'contraire' $\neg\phi$. Une telle présupposition est-elle exagérée, et faudrait-il abandonner plutôt la relation de dualité entre A et S ? La difficulté est semblable au cas du concept de doute, évoqué au premier chapitre : l'équivalence $D\phi \equiv (\neg B\phi \wedge \neg B\neg\phi)$ était considérée comme insuffisante et fut complétée par $B(\neg B\phi \wedge \neg B\neg\phi)$, sans quoi il eût été possible sinon de dire d'un légume ou de tout être non pensant qu'il doute de ϕ sous prétexte qu'il ne tient aucune croyance au sujet de ϕ et $\neg\phi$.

Mais le problème posé par l'interprétation illocutoire des modalités épistémique ne s'arrête pas là. Même si l'on présuppose que tout acte illocutoire est explicite, une difficulté existe concernant le sens de la *négation illocutoire* : que signifie au juste la négation d'une assertion ou d'une considération ? En vertu de la dualité modale, toute énonciation est soit une assertion soit une simple supposition, ce qui revient à dire que le locuteur prend en considération tout énoncé quel qu'il soit : $\forall\phi(S\phi \vee S\neg\phi)$. Ce principe semble reproduire de prime abord le problème du principe de bivalence, mais c'est là une fausse impression : les intuitionnistes ont raison de dire qu'il serait absurde de porter une assertion sur tout énoncé, $\forall\phi(A\phi \vee A\neg\phi)$, mais ce n'est pas le cas dans $\mathcal{L}_{f.AS}$ puisque l'absence de *détermination* de la pensée (donner son assentiment ou son dissentiment à l'égard de ϕ) ne signifie pas la même chose que l'absence de pensée à l'égard de ϕ.[100]

---

[99] Voir section 2.2.4.4 : 'Je crois que p' implique pragmatiquement 'Je ne sais pas si p', en vertu de la règle d'économie informationnelle des actes de discours, ou maxime de quantité (Grice). Cette règle est incompatible avec la thèse d'entraînement $K\phi \rightarrow B\phi$: admettre $K\phi \rightarrow B\phi$ et $B\phi \rightarrow \neg K\phi$ conduirait à la conclusion absurde $K\phi \rightarrow \neg K\phi$.

[100] Nous aurions là la version illocutoire du principe de bivalence : pour tout énoncé ϕ, le locuteur asserte que ϕ est vraie ou que ϕ est fausse. Le cas de l'abstention avait été évoqué en section 3.2.3 comme un argument contre

Mais d'autre solutions sont peut-être envisageables pour défendre la contrepartie illocutoire de **S5** : est-il possible de comprendre le sens d'une non-assertion sans sa supposition conjointe ?

Une autre explication formelle consisterait à modifier la définition de la négation pour les cas de non-assertion. Une méthode pour les contextes de connaissance a été envisagée par [Englebretsen 1969], afin de distinguer l'ignorance de la non-connaissance. Peut-on dire d'une pierre qu'elle 'ignore', c'est-à-dire qu'elle ne sait pas si Nixon a été président des Etats Unis d'Amérique ? En apparence oui, mais on ne le peut pas en réalité dans la mesure où il est inconcevable, inapproprié ou insensé de dire d'une pierre qu'elle 'sait' ou qu'elle ne 'sait' pas si Nixon a été président. La difficulté est identique à celle des énoncés existentiels négatifs : lorsque le terme sujet d'un énoncé désigne un objet fictif ou non existant, alors ni l'affirmation de cet énoncé *ni sa négation* ne sont vrais. Par exemple, l'énoncé 'Pégase a des ailes' n'est pas vrai puisque aucun objet ne satisfait la propriété d'être Pégase et d'avoir des ailes ; or bien que cet énoncé ne soit pas vrai, il n'en devient pas faux pour autant car sa négation 'Pégase n'a pas d'ailes' ne donnera pas un énoncé vrai non plus. En d'autres termes, le problème de la non-assertion est le même que celui des énoncés dépourvus de sens (ni vrais ni faux) et se résume par une question commune : comment exprimer des énoncés dont on ne peut supposer ni la vérité ni la fausseté ?

Pour ce faire, l'auteur propose d'enrichir le symbolisme modal afin d'exprimer le cas des énoncés ni assertés ni supposés. En plus des deux types de négation modale de la logique moderne, c'est-à-dire la négation préposée $\neg A\phi$ et la négation postposée $A\neg\phi$, [Englebretsen 1969] propose une modalité du second ordre, à savoir un opérateur d'actualité @ qui s'ajoute à l'opérateur de connaissance : dans les expressions 'C'est le cas que $A\phi$' et 'C'est le cas que $\neg A\phi$', leurs traductions formelles $@A\phi$ et $@\neg A\phi$ sont mutuellement *contraires* et indiquent deux situations où le locuteur considère la possibilité de ϕ et se détermine ou non sur sa valeur de vérité ; quant aux expressions 'C'est le cas que $A\phi$' et 'Ce n'est pas le cas que $A\phi$', leurs traductions formelles $@A\phi$ et $\neg@A\phi$ expriment deux expressions mutuellement *contradictoires* où la deuxième formule indique une situation dans laquelle le locuteur ne considère même pas à la possibilité de vérité de ϕ. En d'autres termes, $\neg@\phi$ signifie que le locuteur n'a pas même conscience (n'a aucune idée) de ϕ. L'expression @ n'indique pas la propre opinion du locuteur et n'est pas prononcée par lui ; elle parle extérieurement de ce que le locuteur est susceptible d'asserter ou de ne pas asserter : l'actualité $@\phi$ signifie ainsi que le locuteur conçoit (c'est-à-dire suppose la possibilité de vérité de) ϕ mais ne s'engage pas sur sa vérité, tandis que la non-actualité $\neg@$ indique l'absence totale de quelque considération que ce soit ($S\phi$ ou $S\neg\phi$) de la part de ce locuteur.[101] La distinction entre @ et $\neg@$ formalise la distinction *médiévale* entre le sens exclusif de la négation propositionnelle (S n'est pas P : $@\neg\phi$) et le sens privatif de la négation nominale (S est non-P : $\neg@\phi$), une distinction que la logique moderne (fonctionnelle) ne serait pas capable de retranscrire sans introduire un

---

cette bivalence, mais nous avions précisé que celle-ci devait être considérée simplement comme une règle herméneutique de traduction radicale : toute traduction repose sur des cas d'assertion ostensibles.

[101] Si l'on représente par @ l'opérateur d'actualité du second ordre, alors on obtient les définitions suivantes : $@\phi =_{df} A\phi \vee A\neg\phi \vee S\phi \vee S\neg\phi$et $\neg@\phi =_{df} \neg A\phi \wedge \neg A\neg\phi \vee \neg S\phi \vee \neg S\neg\phi$. Par rapport à notre problème de départ, l'interprétation illocutoire de S5, on dira donc que la formule $\neg A\phi \rightarrow A\neg A\phi$ est valide si elle signifie $@\neg A\phi \rightarrow A\neg A\phi$ et qu'elle n'est pas valide si elle signifie $\neg@A\phi \rightarrow A\neg A\phi$.

discours du second ordre puisqu'elle recourt à la copule 'est' : 'S est non-P' signifie que la propriété d'être P ne *peut pas* être prédiquée de S (un canard est non-fumeur : $\neg@\phi$), tandis que 'S n'est pas P' signifie que S *pourrait* être prédiquée P mais ne l'est *pas* dans les faits (mon père ne fume pas : $@\neg\phi$).[102] Cette distinction issue de la logique traditionnelle est-elle pertinente, dans le cadre d'une logique modale illocutoire ? Dans la mesure où celle-ci ne porte que sur des actes de discours prononcés par un locuteur, la position tierce exprimée par @ n'exprime pas sa propre opinion et devrait donc être laissé de côté, bien qu'elle rende compte d'un cas de négation pourvu de sens et que des opérateurs du premier ordre ne sauraient formaliser : dans les cas d'énonciation 'actuels', ou actualisés, la négation *interne* porte sur le contenu de l'énoncé ϕ dans $A\neg\phi$, par opposition à la négation *externe* $\neg A\phi$; la troisième négation de la non-actualité ne se montre pas au sein de $\mathcal{L}_{f.AS}$, sinon par le biais de l'exclusion totale $\neg A\phi \wedge \neg A\neg\phi \wedge \neg S\phi \wedge \neg S\neg\phi$.

Sans parler davantage de cet usage plus complexe de la négation comme absence totale de considération, puisque nous n'en tiendrons plus compte dans ce qui suit (nous supposerons que toute formule de $\mathcal{L}_{f.AS}$ est au moins simplement considérée par un locuteur, ce qui revient à valider l'interprétation illocutoire du schéma **5**), Russell a fait référence par ailleurs à un type précis de négation illocutoire qui n'obéirait pas au carré des oppositions : la *dénégation* $D\phi$, qui exprime une attitude d'incroyance [*disbelief*] opposée à la croyance [*belief*]. En vertu de la dénégation, le locuteur considère (donc suppose) ϕ mais fait plus qu'afficher une des deux attitudes propositionnelles $A\neg\phi$ et $\neg A\phi$. Or si la dénégation semble en dire davantage que ce qu'il est permis d'exprimer par les modalités illocutoires de type $\Box$ et $\Diamond$, faut-il en conclure l'irréductibilité de certaines notions illocutoires au symbolisme modal, donc la preuve que ces notions psychologiques ne sont que partiellement formalisables ? D'ordinaire, l'affirmation locutoire ϕ est considérée comme l'*opposé* de la négation locutoire $\neg\phi$; mais est-ce une opposition de type contraire ou contradictoire ? Les énoncés locutoires $\llcorner\phi$ et $\neg\phi$ sont mutuellement *contradictoires* et, par analogie, les actes d'assertion et de dénégation devraient être considérés comme les formes contradictoires $A\phi$ et $\neg A\phi$. Mais si l'on suit l'interprétation illocutoire de [Kearns 1997], les énoncés locutoires deviennent des énonciations ou actes de discours illocutoires mutuellement *contraires* : $S\phi$ et $S\neg\phi$. Par analogie, l'affirmation illocutoire $A\phi$ n'aurait donc plus pour opposé la forme contradictoire $\neg A\phi$, mais la forme contraire $A\neg\phi$. Nier ϕ, est-ce asserter $\neg\phi$? Observons les propriétés de cette négation illocutoire, ou dénégation, et voyons s'il en va ainsi.

---

[102] [Englebretsen 1969] applique le cas de la négation par privation à la notion de connaissance. Nous proposons ci-dessous des traductions formelles entre crochets, où S et J désignent Socrate et Jones et ϕ l'énoncé 'Nixon a gagné' : "Il est faux que Socrate sait que Nixon a gagné [$\neg@K_S\phi$], et il est faux que Socrate ne sait pas que Nixon a gagné [$\neg@\neg K_S\phi$]. On pourrait objecter que Socrate ne sait effectivement pas que Nixon a gagné [$@\neg K_S\phi$]. Mais nul doute que ceci est un cas très différent de celui dans lequel Jones, qui pense que Humphrey a gagné, ne sait pas que Nixon a gagné [$@\neg K_J\phi$]. Jones est en situation (étant au courant de la compétition entre Nixon et Humphrey) de savoir ou de ne pas parvenir à savoir que Nixon a gagné [$@(K_J\phi \vee \neg K_J\phi$]. Mais ce n'est pas le cas pour Socrate, ni pour quiconque n'étant pas au courant de la compétition [$\neg@K\phi \rightarrow \neg S\phi$]. Il serait certainement faux de dire de Socrate qu'il a douté, simplement cru ou incru, etc. que Nixon gagnerait [$\neg@K_S\phi \rightarrow \neg(B_S\phi \vee B_S\neg\phi)$]. Gardons à l'esprit le fait que *Socrate sait que Nixon a gagné* [$@K_S\phi$] et *Socrate ne sait pas que Nixon a gagné* [$@K_S\neg\phi$] sont seulement contraires, et non contradictoires." ([Englebretsen 1969], pp. 581-582).

3.3.2 Négation et dénégation

De même que l'on peut itérer les signes lors d'un calcul d'arithmétique, on peut itérer les signes lors d'un calcul logique : dans les deux cas de figure, nous n'avons plus une simple fonction $f(x)$ appliquée à un argument x mais une fonction de fonction $g[(f)(x)]$. *Affirmer la négation* d'un énoncé, c'est le *nier* ; *nier l'affirmation* d'un énoncé, c'est le *nier* également. La raison de ces deux équivalences repose sur l'identité des actes d'affirmation et de supposition de vérité, d'une part, de négation et de supposition de fausseté d'autre part. Une façon de comprendre la relation algébrique entre un signe et son interprétation consiste ainsi à comparer les opérations logiques à des opérations arithmétiques : de même que l'on a les égalités $-(+) = +(-) = -$ et $-(-) = +$, l'analogie entre le signe positif + et le vrai 1, puis entre le signe négatif – et le faux 0 permet de dire que l'acte de déclarer la vérité de la fausseté d'un énoncé ou la fausseté de sa vérité revient à déclarer sa fausseté : $1(0) = 0(1) = 0$, et que déclarer la fausseté de la fausseté d'un énoncé revient à déclarer sa vérité $= 0(0) = 1$. La déclaration de vérité est l'opération consistant à interpréter une formule comme vraie, et la déclaration de fausseté est celle consistant l'interpréter comme fausse. La force illocutoire transmise lors de l'énonciation distinguera ensuite deux types d'actes parmi les déclarations ou affirmations : la supposition et l'assertion. Mais à quelle interprétation correspond la notion illocutoire de *dénégation* ? C'est ce que nous allons tenter de voir dans ce qui suit, en tâchant de distinguer plusieurs acceptions ambiguës de la négation. La distinction entre une opération effectuée par un agent (le *cogito* cartésien) et son résultat (la pensée ou *cogitatio* cartésienne) apportera un éclairage sur ce genre de concept.

Jusqu'à présent, nous avons retenu quatre forces illocutoires, c'est-à-dire quatre opérateurs analogues à ceux du carré des oppositions modales : l'opérateur d'assertion A analogue à $\Box$, et la supposition S analogue à $\Diamond$, ce qui donne en tout quatre formules de base définies selon la force illocutoire et la qualité de l'énoncé : Ap, $A\neg p$, Sp et $S\neg p$, où p est un signe propositionnel atomique.[103] Or [Vernant 2003] fait état chez Russell d'un cinquième élément apparemment irréductible aux quatre formes précédentes : la dénégation, exprimant l'état d'esprit du rejet.[104] En temps normal, chaque force illocutoire exprime une attitude psychologique distincte : $A\phi$ et $A\neg\phi$ expriment la croyance justifiée que ϕ ou $\neg\phi$; mais si ϕ et $\neg\phi$ sont des opposés locutoires, quel est l'opposé illocutoire de $A\phi$: $\neg A\phi$, $A\neg\phi$, ou aucun des deux ?[105]

---

[103] Puisque les atomes propositionnels de $\mathcal{L}_{f.AS}$ sont des formules illocutoires, les quatre formules indiquées ici sont des atomes. Or si les atomes sont des formes modales, $\mathcal{L}_{f.AS}$ crée un précédent dans les règles de formation des formules : l'introduction d'un connecteur logique ne transforme pas toujours une formule atomique en une formule moléculaire. En effet, la négation à $A\phi$ donne $\neg A\phi$ et $\neg A\phi$ équivaut à l'atome $S\neg\phi$, donc $\neg A\phi$ est un atome propositionnel de $\mathcal{L}_{f.AS}$.

[104] De même qu'il ne faut pas confondre l'affirmation illocutoire (ou assertion) $\vdash\phi$ avec l'opérateur vérifonctionnel ou locutoire de l'affirmation, symbolisé $\llcorner\phi$, il ne faut pas confondre le rejet illocutoire (ou dénégation) $\dashv\phi$ avec le connecteur vérifonctionnel ou locutoire du rejet, symbolisé $\phi\downarrow\psi$ et signifiant la négation de la disjonction ('ni ϕ ni ψ').

[105] La question a été posée par [Wertz 1973,1976], en référence à une contraposition pragmatique exposée par Austin dans *How to Do Things with Words* : si le fait que je dise que le chat est sur le tapis implique que je crois que le chat est sur le tapis, le fait que je ne croie pas que le chat est sur le tapis implique-t-il que le chat n'est pas sur le tapis, dans la langue anglaise ordinaire ? Il s'agit de savoir plus précisément si mon absence de croyance implique que je *dise* que le chat n'est pas sur le tapis, ou que je ne *dise pas* que le chat est sur le tapis : $(A\phi \rightarrow B\phi) \rightarrow (\neg B\phi \rightarrow \neg A\phi)$, ou $(A\phi \rightarrow B\phi) \rightarrow (B\neg\phi \rightarrow A\neg\phi)$. Le problème concerne le sens de la négation

Si l'on suit l'approche de Kearns, les forces illocutoires expriment deux degrés de croyance distincts : une assertion A et une supposition S, équivalant respectivement à la croyance forte ou conviction et à la croyance faible ou conjecture de [Lenzen 2003]. Si l'on suit par contre l'approche partiellement locutoire et illocutoire de Searle & Vanderveken et Vernant, la considération ou supposition n'est pas une modalité illocutoire analogue à la modalité faible $\Diamond$ mais la simple expression d'un contenu jugeable, de sorte qu'il n'y a alors qu'un degré de croyance exprimé par le jugement (l'assertion, ou conviction) ; de préférence à la relation duale que nous avons suggérée entre A et S, ces auteurs opposent à la croyance forte exprimée par A l'état d'esprit de l'incroyance [*disbelief*], que nous symboliserons provisoirement par I.[106] Comment s'articulent logiquement les opérateurs A, S et I ? Dans [Searle 1969], l'auteur déclare que la dénégation ou négation illocutoire de l'acte de discours 'Je promets de venir' n'est pas 'Je promets de ne pas venir', mais 'Je ne promets pas de venir', soit la modalité négative $\neg A\phi$. Mais encore une fois, la dualité établie entre $\neg A$ et $S\neg$ ne suffit pas ici car la dénégation de ϕ n'équivaut pas simplement à $S\neg\phi$. Que signifie l'incroyance, et le choix de $\neg A\phi$ est-il judicieux pour l'exprimer dans $\mathcal{L}_{f.AS}$? C'est également la négation externe que [Hintikka 1962] avait adoptée : la première version du paradoxe de Moore '$\phi \wedge \neg B\phi$' était présentée en effet comme un cas de *saying and disbelieving*.[107]

Il y a une ambiguïté sur l'interprétation de l'incroyance, partant de l'idée selon laquelle tout acte de discours exprime un engagement ou une attitude du locuteur à l'égard d'un signe propositionnel : que S soit considéré comme un *non*-engagement (selon Vernant) ou comme un engagement *faible* (selon notre position et celle de Kearns), il diffère en tout cas de l'état de *dés*engagement associé à la dénégation. Comment exprimer la différence entre engagement, non-engagement et désengagement ? Pour faire mieux comprendre l'expression de l'incroyance [*disbelief*] qu'est la dénégation, [Vernant 2003] propose tout d'abord une analogie religieuse : l'incroyance serait à l'athéisme ce que la non-croyance [*nonbelief*] est à l'agnosticisme.[108] En d'autres termes, l'athée incarnerait l'incroyant lorsqu'il asserte que Dieu n'existe pas ($A\neg\phi$), à la différence du simple agnostique qui ne s'engage pas sur la question de l'existence de Dieu et suspend son jugement ($\neg A\phi \wedge \neg A\neg\phi$). A partir de l'opposition entre assertion et dénégation, Vernant propose d'établir une dichotomie entre une classe d'engagements, composée des deux opérateurs pragmatiques que sont l'assertion et la dénégation, et une classe de non-engagement, composée de la 'modalité' neutre qu'est la considération ou supposition S. Pour exprimer cette dichotomie, [Vernant 2003] propose (p. 4) la ramification suivante des aspects locutoire et illocutoire correspondant à nos opérateurs illocutoires S et A :

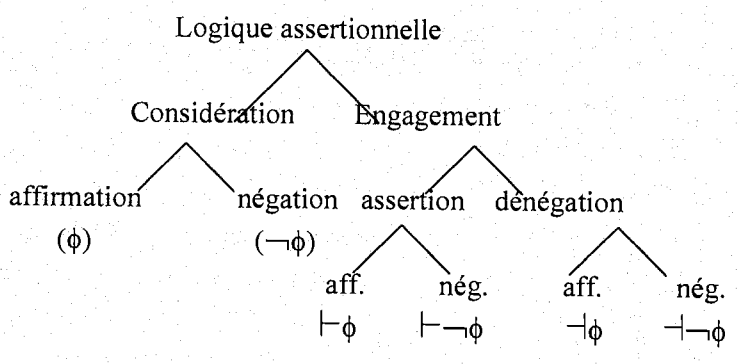
---

illocutoire, dont les propriétés ne sont pas exactement celles de la négation classique : dire 'je ne crois pas ϕ' suggère souvent que l'on croit que $\neg\phi$. Voir [Klima 1964].

[106] Vernant symbolise la dénégation par l'opérateur D, mais nous choisissons I pour la distinguer de l'opérateur du doute de [Hart 1980]. De même que 'untrue', 'disbelief' est une expression usuelle dans la langue anglaise mais n'a pas de contrepartie courante dans la langue française, d'où le caractère assez néologique de la notion d''incroyance' et la difficulté de s'entendre sur sa signification dans l'usage. Les notions de 'dénégation', de 'rejet' ou de 'refus' lui sont associées en général.

[107] Voir [Hintikka 1962], section 4.5, pp. 64-67. De même que dans [Black 1952], l'incroyance est exprimée par la forme modale $\neg\Box$. Dans [Wertz 1976], en revanche, l'incroyance est $\Box\neg$ et s'oppose à la non-croyance $\neg\Box$.

[108] "A la croyance de Dieu, la *foi*, s'oppose directement l'incroyance correspondant à l'*athéisme*, les deux se distinguant de l'agnosticisme qui correspond à un désengagement" ([Vernant 2003], p. 2 note 10.



Quelle est la relation logique entre la dénégation I, l'assertion négative A¬ et la non-assertion ¬A ? Alors que Vernant comparait plus haut l'incroyance à l'engagement négatif de l'athée, il procède par ailleurs à deux rectifications profondes. Dans un premier temps, l'auteur semble se contredire en reprenant la formalisation de l'incroyance par [Searle 1969], où la négation illocutoire n'est pas postposée mais antéposée au marqueur de force : l'incroyance de ϕ devient alors ¬Aϕ, et non l'attitude A¬ϕ de l'athée. Dans un second temps, Vernant précise que la négation illocutoire de [Searle & Vanderveken 1985] n'est plus la même que celle de [Searle 1969] : le tilde ~ est utilisé pour distinguer la négation locutoire ¬ de la négation illocutoire ~, si bien que l'incroyance d'un énoncé négatif n'a plus la forme ¬A¬ϕ mais ~A¬ϕ. D'après Vanderveken, la négation illocutoire ~ est comparable à la négation des intuitionnistes, ce qui permet de distinguer la dénégation ~A de la simple absence d'assertion ¬A :

> La négation illocutoire sous cet aspect ressemble à la négation intuitionniste puisqu'en logique intuitionniste, il est aussi non valide que P ∨ ~P et que ~~P → P.[109]

Toutefois, l'arborescence ci-dessus ne représente pas la dénégation Iϕ par la forme négative ~Aϕ et maintient la distinction entre les forces illocutoires A et I. Si Vernant choisit la forme ⊣ϕ (notre Iϕ) afin d'exprimer la dénégation, c'est parce que, pour reprendre Russell, l'incroyance ne signifie pas simplement l'absence d'assertion mais caractérise avant tout un acte ou état d'esprit positif :

> l'incroyance est un état aussi positif que la croyance.[110]

Elle ne signifie pas non plus un engagement sur ϕ de type A¬ϕ :

> L'incroyance est une nouvelle relation indéfinissable, impliquant simplement le rejet de la proposition incrue [*disbelieved*] et non pas consistant en l'acceptation de sa contradictoire,[111]

---

[109] [Searle & Vanderveken 1985], p. 154. Notre version modale de ce tiers exclu illocutoire donnerait Aϕ ∨ Iϕ : tout énoncé est soit asserté soit dénié ; celle de la double négation donnerait IIϕ → Aϕ : nier que l'on nie ϕ, c'est asserter ϕ. Dans l'article original cité, ~ et ¬ désignent respectivement la négation locutoire et illocutoire, mais nous les avons inversé en raison de notre symbolisme initial.

[110] B. Russell, *Human Knowledge*, cité in [Vernant 2005], p. 16 note 84.

[111] [Russell 1913], p. 142 ; nous soulignons. Partant que $=_{df}$ et ≡ sont interchangeables, Russell suggère ici qu'aucune loi bi-conditionnelle ne peut s'appliquer à Iϕ : ni Iϕ ≡ A¬ϕ, ni Iϕ ≡ ¬Aϕ, ni Iϕ ≡ S¬ϕ. L'ambivalence

concluant de fait que la dénégation exprime un état d'esprit intermédiaire entre ceux exprimés par $\neg A\phi$ et $A\neg\phi$. La même sorte de distinction entre $\neg\Box\phi$ et $\Box\neg\phi$ avait été évoquée par [Łukasiewicz 2000] pour reconstruire l'argumentation aristotélicienne du Principe Psychologique de Non-Contradiction (PPNC : personne ne peut croire qu'une même chose est et n'est pas) : la contradiction psychologique en question ne porte pas sur les croyances mais sur les énoncés crus, de sorte que la juste traduction modale du PPNC ne serait pas $B\phi \wedge \neg B\phi$ mais $B\phi \wedge B\neg\phi$.[112] Mais à la différence du Polonais, qui visiblement n'utilise pas encore ici son futur opérateur de rejet $\dashv$ (l'article sur le PPNC date de 1910), Russell distingue nettement les deux actes positifs que sont la croyance de $\neg\phi$ ou assertion négative $A\neg\phi$ et l'incroyance de ϕ ou dénégation $I\phi$.[113]

La dénégation est donc un acte de jugement qui ne se confond ni avec la simple absence de jugement, ni avec un jugement négatif. Pour Frege, en revanche, la négation désigne uniquement le contenu d'un énoncé et ne porte jamais sur la nature du jugement. Autrement dit, la négation ne relève pas du jugement et désigne uniquement selon lui une propriété locutoire du contenu jugeable :

> cette négation ne doit pas être mise au même rang que le jugement, ni être interprétée comme le pôle opposé du jugement. Il s'agit toujours, dans le jugement, de vérité.[114]

---

des concepts des langues naturelles se confirme avec l'expression douteuse 'incru' : si on la compare à 'insu', l'incroyance devrait donner $\neg B$ puisque 'à l'insu de *a*' indique que la personne *a* ne sait pas, $\neg K$; c'est le sens que [Gardiès 1979] prête à l'incroyance, contrairement à ici. Il en va de même avec le préfixe 'in' : 'invalide' s'oppose à 'non-valide', mais l'invalidité de ϕ signifie tantôt la validité de $\neg\phi$ tantôt sa simple satisfiabilité. Voir la section 1.3.1 et le carré des oppositions métalogiques, où $\Box\neg\phi$ est comparée à $W \models \neg\phi$ et $\neg\Box\phi$ à $w \models \neg\phi$.

[112] "Il faut reconnaître qu'en prouvant le principe psychologique de contradiction, Aristote raisonne de façon très prudente. Il est visiblement conscient que la coexistence dans le même esprit de deux convictions auxquelles correspondent des [énoncés] contradictoires ne constituerait pas de contradiction évidente. En effet, chaque conviction est un acte psychique *positif*. Par conséquent, si quelqu'un avait la conviction que quelque chose est et, en même temps, la conviction que la même chose n'est pas, il n'y aurait jamais de contradiction évidente. Une contradiction évidente apparaîtrait seulement si, dans le même esprit, une conviction à la fois existait et n'existait pas". ([Łukasiewicz 2000], p. 62; nous soulignons. Le PPNC affirme donc l'impossibilité de tenir conjointement des jugements *contraires* ($A\phi$ et $A\neg\phi$) portant sur des énoncés *contradictoires* (ϕ et $\neg\phi$).

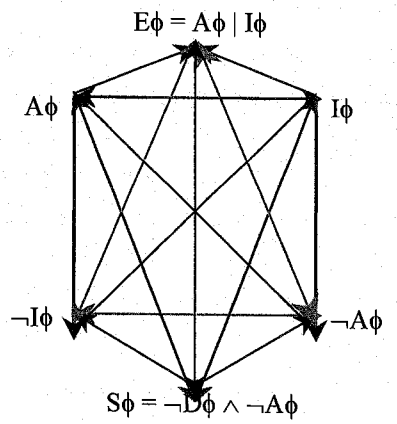
[113] A moins que Łukasiewicz n'assimile tout simplement la dénégation à l'assertion négative : $I\phi \equiv A\neg\phi$, de son point de vue? Il le semble, si l'on en juge par sa règle de rejet suivante (citée in [Vernant 2003], p. 3 note 14) : "si l'implication 'si α alors β' fait l'objet d'une assertion, mais que son conséquent soit rejeté, alors il faut nécessairement rejeter aussi l'antécédent α", soit $A(\alpha \rightarrow \beta) \rightarrow (I\beta \rightarrow I\alpha)$. Cette règle de rejet ayant l'apparence schématique d'une contraposition : $(\phi \rightarrow \psi) \rightarrow (\neg\psi \rightarrow \neg\phi)$, elle pourrait vouloir dire ceci : si la vérité de ϕ implique l'affirmation de la vérité de ψ, alors la fausseté de ψ implique la fausseté de ϕ. Or si le rejet exprime un *acte d'affirmation de* la fausseté, la 'fausseté de ϕ' ne devrait pas être symbolisée par $\neg\phi$ mais par $A\neg\phi$; donc la règle de rejet devient $(A\phi \rightarrow A\psi) \rightarrow (A\neg\phi \rightarrow A\neg\psi)$. Or ceci n'est pas une contraposition : si la règle de rejet est une contraposition, nous devrions l'exprimer sous la forme $(A\phi \rightarrow A\psi) \rightarrow (\neg A\psi \rightarrow \neg A\phi)$ et en conclure que $I\phi \equiv \neg A\phi$, et non $A\neg\phi$. Mais sur la base de l'analogie entre assertion = '*est* vrai' et rejet = '*est* faux', $I\phi$ prend le sens de l'assertion négative $A\neg\phi$. Le fait est que Łukasiewicz ne symbolisait pas l'affirmation de ϕ par un opérateur modal et l'assimilait au contenu propositionnel ϕ : ne pas affirmer la vérité de ϕ équivalait pour lui à affirmer sa fausseté, sans intermédiaire possible. Son rejet équivaut dans ce cas à la règle du modus tollens, qui exprime un *mode* d'affirmation : 'à partir des assertions de $\phi \rightarrow \psi$ et de $\neg\psi$, j'asserte $\neg\psi$. Sur cette règle de rejet, voir J. Łukasiewicz : *La syllogistique d'Aristote*, trad. F. Zaslawski, Paris, Colin (1972), en part. ch. 3, § 20, pp. 82-88.

[114] [Frege 1919b], p. 209. La dernière phrase signifie que tout jugement est de forme 'Il est vrai que ...', ce qui ne permet pas de rendre compte des attitudes propositionnelles autres que $A\phi$; contrairement aux commentaires

Si Frege se prononce ainsi, c'est parce qu'il réduit par l'absurde l'idée de dénégation à celle d'un jugement négatif synonyme de destruction ou élimination du jugement.[115] Sa position sur les jugements négatifs est similaire au traitement russellien des *faits négatifs* : de même qu'un fait négatif exprimé sous la forme $\neg\phi$ n'existerait pas en soi mais désignerait indirectement la fausseté du fait 'positif' ϕ (tout fait étant positif par principe), de même tout jugement négatif désignerait indirectement la fausseté du jugement 'positif' correspondant (tout jugement étant positif par principe).[116] Or c'est précisément cette réduction indirecte que Russell refuse d'appliquer à la notion d'incroyance et son expression de dénégation, d'où la préférence du symbole positif $\dashv$ ou I en lieu et place de $\sim\vdash$ ou $\sim$A. En conséquence, la dénégation de Russell n'équivaut ni à $A\neg\phi$, ni à $S\neg\phi$:

> Dénier une proposition n'est pas la même chose que d'asserter sa négation (...) quand nous dénions p, nous ne sommes pas concernés par une simple assomption, et il n'y a rien qui puisse être fait avec p qui soit logiquement équivalent à assumer non-p.[117]

A la place de notre carré des modalités précédent, composé des deux forces illocutoires A et S, Vernant tient compte de la dénégation D (notre I) et propose un hexagone des oppositions qui rappelle la logique du doute de [Hart 1980]. L'opérateur de non-engagement E joue ici le rôle de la contingence illocutoire ∇ et correspond à notre supposition S, opposé à un opérateur d'engagement E qui est l'interprétation illocutoire de la non-contingence $\triangle$. On obtient l'hexagone suivant :



Mais cet hexagone des modalités nécessite quelques aménagements qui semblent discutables, dans lequel deux notions distinctes, l'assertion et la dénégation, se partagent le

---

de Russell sur la notion de dénégation, la négation de Frege ne s'applique donc jamais à la barre de jugement mais uniquement au contenu jugeable.

[115] "La négation ne fait pas d'une non-pensée une pensée, pas plus qu'elle ne fait d'une pensée une non-pensée" ([Frege 1919b], p. 202). Cet argument est incorrect s'il assimile la dénégation à la non-pensée, alors que celle-ci est à situer quelque part entre la pensée négative $A\neg\phi$ et la non-pensée $\neg A\phi$.

[116] Contrairement à nous, Frege prétend que le rejet de ϕ et l'assertion de $\neg\phi$ ne font qu'un : "Le rejet de l'une et la reconnaissance de l'autre est un même acte. Il n'y a donc pas besoin d'un nom particulier ou d'un signe particulier pour le rejet [*Verwerfung*] d'une pensée". G. Frege, "Einleitung in die Logik", in *Ecrits Posthumes* (1999), p. 221 ; nous soulignons.

[117] [Russell 1973] p. 41.

rôle de l'opérateur fort □ : $A\phi$ et $I\phi$ sont considérés comme mutuellement contraires ; or la relation de contrariété modale habituelle oppose $\Box\phi$ et $\Box\neg\phi$, c'est-à-dire soit $A\phi$ et $A\neg\phi$, soit $I\phi$ et $I\neg\phi$. Mais rappelons que $I\phi$ n'est surtout pas équivalent à $A\neg\phi$, qui n'a d'ailleurs plus sa place dans cet hexagone de Vernant. De plus, une caractéristique singulière de l'hexagone des oppositions est qu'il admet les propriétés de paraconsistance et de paracomplétude : les lois de non-contradiction et du tiers exclu ne sont plus valides. Or [Vernant 2003] signale que la première ces deux lois est préservée dans sa logique illocutoire :

> 1°– la non-contradiction est respectée : pour chaque opérateur assertionnel (noté), on a $\neg(*p \wedge \neg*p)$. Par exemple, on ne peut pas à la fois asserter et ne pas asserter une proposition : $\neg(Ap \wedge \neg Ap)$. 2°– le tiers exclu ne vaut plus, car il est parfaitement possible de ne pas choisir entre asserter et dénier et d'adopter la *position neutre* du désengagement qu'est la seule considération. Ce qui joue alors le rôle de tiers exclu est ce que je propose de baptiser la 'loi de Russell' : $Dp \rightarrow A\neg p$ (...) qui constitue sans doute la première expression d'une loi proprement pragmatique. 3°– La double négation ne vaut plus, car si asserter implique ne pas dénier, ne pas dénier n'équivaut pas à asserter p. Il est en effet toujours possible d'adopter la position neutre de simple considération : $Ap \rightarrow \neg Dp$ et $\neg(\neg Dp \rightarrow Ap)$. Dès lors, ne pas asserter n'est pas dénier : $\neg(\neg Ap \equiv Dp)$, c'est une simple position de désengagement.[118]

Comment maintenir ces deux lois tout en admettant une représentation hexagonale des oppositions entre les opérateurs modaux ? Si l'hexagone ci-dessus a remplacé la forme contraire usuelle de $A\phi$, $A\neg\phi$, par la dénégation $I\phi$, c'est manifestement pour tenir compte de la 'loi de Russell' que l'auteur assimile au tiers exclu et pour reformuler ce dernier comme une opposition entre des termes *contraires* ($A\phi$ et $I\phi$). Mais si l'on admet ce déplacement conceptuel, pourquoi ne pas en faire autant pour le principe de non-contradiction et le reformuler dans un hexagone comme une opposition entre des termes *subcontraires* ($\neg I\phi$ et $\neg A\phi$) ? En d'autres termes, la non-contradiction désigne toujours une opposition entre des termes *contradictoires*, ici ($A\phi$ et $\neg A\phi$), tandis que la signification du tiers exclu a été modifiée dans sa relation d'opposition. Par ailleurs, comment l'équivalence est-elle établie entre le tiers exclu, la relation de contrariété symbolisée par une disjonction exclusive et la loi de Russell, symbolisée par un conditionnel ? Un raisonnement simple permet de montrer leur lien :

1.	$I\phi \rightarrow A\neg\phi$	prémisse
2.	$\neg A\neg\phi \rightarrow \neg I\phi$	1., contraposition
3.	$A\neg\phi \vee \neg I\phi$	2., déf. $\rightarrow/\vee$
4.	$A\phi \vee I\phi$	Sub $[\phi/\neg\phi]$

A partir de la loi de Russell, posée dans l'étape 1, la définition disjonctive du conditionnel permet d'obtenir dans l'étape 4 la relation de contrariété illustrée dans l'hexagone de Vernant,

[118] [Vernant 2003], p. 5. La 'loi de Russell' s'exprime ainsi : "Si p est déniée, non-p doit être assertée", Russell ajoutant qu'une telle loi ne remet cependant pas en cause le tiers exclu puisqu'elle n'appartient pas au domaine de la logique : "cette forme, il est vrai, est trop psychologique pour être ultime, mais le fait est qu'elle est signifiante et ne se réduit pas à une simple tautologie. Logiquement, la notion de dénégation d'une proposition p n'est pas pertinente ; seule la vérité de non-p concerne la logique." ([Russell 1973], p. 41).

à la différence notoire que la disjonction qui relie dans l'hexagone l'assertion et la dénégation n'est pas inclusive mais exclusive. Le raisonnement ci-dessus ne serait pas admis par Russell sous cette forme modale, le formalisme contenant des forces illocutoires qu'il excluait du symbolisme strictement logique. Vernant admet en revanche la logique illocutoire, et une ambiguïté plus pressante que le sens de la disjonction concerne le côté attribué dans son hexagone à la notion de dénégation : pourquoi Vernant lui attribue-t-il le côté contraire à l'assertion $A\phi$, et comment distinguer alors $I\phi$ de la forme également contraire $A\neg\phi$? S'il est impossible de définir la dénégation en termes primitifs d'assertion, alors l'assertion et la dénégation devront être considérés en effet comme des opérateurs indépendants et de nature distincte.[119] L'hexagone de Vernant donne-t-il une représentation correcte de la logique modale illocutoire, ou ne devrait-on pas procéder plutôt à un réajustement des opérateurs ainsi qu'à une révision de ladite 'loi de Russell' ?

L'analogie précédente de [Searle & Vanderveken 1985] entre dénégation et négation intuitionniste semble incliner dans ce sens et permettre de considérer $I\phi$ comme un terme définissable, contrairement à l'affirmation précédente de [Russell 1913] : si l'on adapte l'interprétation négative de la dénégation $I\phi = \sim\phi$, la traduction modale de la négation intuitionniste dans [Gödel 1933] donnait $\sim\phi = B\neg B\phi$ et donnerait ici la contrepartie illocutoire $A\neg A\phi$.[120] Prise dans ce sens, la dénégation signifie l'assertion d'une non-assertion ou la disposition du locuteur à déclarer qu'il ne s'engage pas sur ϕ. Un passage de [Vernant 2005] pourrait confirmer cette définition itérée de la dénégation :

> la dénégation n'est pas simple négation d'une assertion possible, mais position de rejet, manifestation de *refus* d'une proposition implicite si elle n'a pas été explicitement proposée.[121]

Pourrait-on définir finalement la dénégation en termes d'assertion, contrairement à l'idée russellienne d'une dénégation indéfinissable ? Si notre suggestion est correcte, alors la logique illocutoire n'est pas une logique modale comparable à la logique épistémique, composée de deux opérateurs indépendants de type K et B, mais une logique modale composée d'un unique opérateur primitif A. En revanche, l'analogie entre dénégation et négation intuitionniste devrait impliquer également que I est caractérisé par le système modal **S4** et n'admet pas **S5**. Est-ce le cas ?

D'un autre côté, le schéma **5** exprimé par la formule $\neg A\phi \rightarrow A\neg A\phi$ signifie désormais la même chose que $\neg A\phi \rightarrow I\phi$: toute non-assertion implique une dénégation. Si la dénégation se définit comme ce qui exclut l'assertion, alors cette formule est admissible et même valide,

---

[119] La relation entre $A\phi$ et $I\phi$ est du même type que celle entre $K\phi$ et $B\phi$: les deux paires de notions ne sont pas mutuellement définissables, même si l'on peut établir des relations logiques entre elles (la connaissance implique la croyance, par exemple). Le problème posé par le polygone de Vernant est qu'il traite la dénégation comme un acte de discours auquel ne correspond aucune attitude en particulier : si l'on admet qu'à l'assertion $A\phi$ correspond l'attitude de connaissance ou croyance justifiée $K\phi = B\phi \wedge J\phi$, on ne sait pas à quelle attitude correspond la dénégation.

[120] Les expressions 'incru' et 'insu' ne sont donc pas symétriques, dans la mesure où 'ϕ à l'insu de *a*' indique non seulement que *a* ne sait pas que ϕ mais aussi qu'il ne sait pas qu'il ne sait pas que ϕ. Par opposition, la définition ci-dessus de l'incroyance suggère que le locuteur *sait* ou a conscience du fait qu'il ne croit pas que ϕ. L'idée d'ignorance inconsciente véhiculée par 'insu' ne se retrouve donc pas dans 'incru' : l'incroyance est une non-croyance consciente.

[121] [Vernant 2005], p. 17

mais à condition que la non-assertion implique l'assertion de la non-assertion. La traduction modale de [Gödel 1933] avait pour but de montrer que la logique intuitionniste est comparable à **S4** et non à **S5**, mais notre interprétation de la non-assertion introduit toutefois une 'clause de supposition minimale' (ce qui n'est pas asserté est au moins supposé : $\neg A \rightarrow S$) qui modifie quelque peu les termes : la non-croyance de l'intuitionniste n'équivaut pas à la croyance d'une non-croyance, contrairement à notre interprétation illocutoire de la non-assertion ; celui qui n'asserte pas déclare sa non-assertion et, de ce fait, est disposé à asserter ce qu'il n'asserte pas. La version **S5** est donc maintenue, et la comparaison avec la négation intuitionniste a ses limites.

D'un autre côté, le problème de définition de la dénégation se pose davantage avec la double négation qu'avec le tiers exclu : même si l'on établit de bon droit la relation duale entre A et S, du fait que la dénégation désigne un acte de discours explicite, l'interprétation de la loi de double dénégation est non seulement ambiguë mais semble impliquer le rejet de **S4**. On obtient en effet la formule $II\phi \rightarrow A\phi$ qui, par substitution, produit la formule $ASSA\phi \rightarrow A\phi$. Une simplification de ce résultat peu intuitif exigerait une loi de réduction des opérateurs modaux, c'est-à-dire une formulation de conditions sémantiques adaptées à A, S et I. On peut concevoir à cette fin une sémantique relationnelle pour $\mathcal{L}_{f.AS}$, dans laquelle un monde possible représenterait un énoncé accessible à l'acte de discours de départ, c'est-à-dire susceptible d'être produit par le locuteur. Si $\mathcal{L}_{f.AS}$ est une logique modale de type **S4**, alors la relation R_i définie entre les mondes possibles illocutoire sera transitive ; si elle est de type **S5**, alors R_i sera euclidienne. Or on peut montrer que la version illocutoire de la loi de double négation, la loi de double dénégation équivalente à $ASSA\phi \rightarrow A\phi$, est valide dans **S5** mais ne l'est pas dans **S4** (il existe des contre-modèles dans lesquels la formule $ASSA\phi \wedge \neg A\phi$ est vraie). Faut-il en conclure que la caractérisation de l'assertion en logique modale illocutoire sera de type **K□T5**, et non **K□T45** ?[122] Plutôt que de rejeter la relation de transitivité pour R_i, nous allons proposer pour des raisons indépendantes une autre définition de la dénégation.

De préférence à notre formulation itérative de la dénégation, un passage de [Vernant 2003] pourrait plaider en faveur d'une logique illocutoire multimodale :

> Ce qui vaut pour l'assertion s'applique *a fortiori* à la dénégation. Celle-ci s'impose comme un *refus* d'une pro-position *antérieure*, explicite ou non. La dénégation apparaît ainsi comme *seconde* : d'où le préfixe 'dé' dans : dénier, désavouer, démentir, etc., où le préfixe 're' dans : refuser, renier, répudier, renoncer, etc., ou encore 'in' dans infirmer dans infirmer, incroyance, etc. On ne peut renier sa foi que si on l'avait antérieurement professée.[123]

---

[122] La distinction entre **□T** et T est essentielle, puisque le schéma **T** est syntaxiquement inadmissible : $A\phi \rightarrow \phi$ est dépourvue de sens, tandis que $A(A\phi \rightarrow \phi)$ est à la fois pourvue de sens et est valide. Alors que la propriété de R dans **T** est la réflexivité : $\forall w(wRw)$, la propriété de R_i dans **□T** est la réflexivité inversible : $\forall w \forall w^*[(wRw^*) \rightarrow (wRw)]$. Le schéma **T** implique le schéma **□T** (par la règle de nécessitation), donc tout ce qui est valide dans **□T** (la classe des structures où **□T** est valide) est valide également dans **T** (la classe des structures où **T** est valide), mais la réciproque n'est pas vraie. Or dans la mesure où les formules valides dans **T** sont dépourvues de sens dans notre logique assertorique, toutes les inférences de $\vdash\phi$ à $\vdash\Box\phi$ ou de $\vdash\Box\phi$ à $\vdash\phi$ sont syntaxiquement incorrecte.

[123] [Vernant 2003], p. 6. Dénégation et reniement semblent être des actes distincts : alors que la dénégation consiste à *annuler* une assertion antérieure en se désengageant à son égard, le reniement consiste plutôt à prendre le parti inverse et à passer d'une assertion $A\phi$ à son contraire : $A\neg\phi$. Pierre a renié trois fois le Christ en assertant

La référence ci-dessus au temps d'énonciation pourrait inciter à définir la dénégation en termes de modalités illocutoires et temporelles, d'autant qu'un traitement de ce genre avait été proposé pour le paradoxe de Moore dans [Bovens 1995]. Si l'on symbolise par P l'opérateur du temps passé, l'idée de désengagement qui accompagne la notion de dénégation pourrait être rendue comme l'idée selon laquelle un énoncé a été asserté par le locuteur et ne l'est plus au temps d'énonciation présent : $I\phi \equiv PA\phi \wedge \neg A\phi$. La méthode des combinaisons multimodales permettrait ensuite de fusionner les sémantiques des modalités temporelle et illocutoire et de calculer sur les expressions de dénégation ; mais ce n'est toutefois pas cette voie que nous suivrons ici : nous en resterons à l'idée de réduction des modalités au sein d'un système monomodal de l'assertion, dans lequel même la dénégation puisse être interprétée en termes de A.

Pour ce faire, une interprétation sémantique pourrait aider à clarifier le sens de la dénégation. C'est ce que propose [Vernant 2003], mais avec une valuation ambiguë. Une sémantique illocutoire a déjà été suggérée en fait par [Bochvar 1938], en partant de la distinction entre des opérateurs *internes* du langage objet et des opérateurs *externes* du métalangage. La description qu'en donne [Malinowski 1993] distingue les deux types d'opérateurs de la même manière que Frege distinguait les contenus jugeables des jugements : les opérateurs *internes* portent sur des énoncés tels que ϕ, les opérateurs *externes* portent sur des jugements tels que 'ϕ est vrai'.[124] Malgré des divergences entre cette approche et la nôtre (l'assertion de Bochvar correspond à la déclaration de fait de [Hanson 1952] et la valeur de l'énonciation $A\phi$ dépend donc de celle de l'énoncé ϕ, contrairement à notre lecture 'ontique' de l'assertion comme déclaration factuelle ou *portant sur* un fait), nous avons là une sémantique toute trouvée pour la logique illocutoire. Dans une sémantique modale de type [Gardiès 1979], où le couple de valeurs {vrai, faux} est attribué aux formules du 'monde réel' dans un 'monde proposé' (un monde imaginé ou construit au cours d'un dialogue), [Vernant 2003] pose ainsi les deux valuations suivantes : la vérité de $A\phi$ signifie que ϕ est vrai dans le monde proposé, et la vérité de $A\neg\phi$ signifie que ϕ est faux dans le monde proposé. En revanche, la valuation qu'il suggère pour la dénégation est contestable : la vérité de $I\phi$ signifierait que p est faux dans le monde proposé, et la fausseté de $I\phi$ signifierait que ϕ est faux également dans le monde proposé. Le problème est que la méthode de démonstration de Gardiès (dont [Vernant 2003] s'inspire) n'est pas indirecte (méthode apagogique de réfutation par l'absurde) mais directe (valuation dans un monde), et que cette valuation assigne une valeur de vérité à un monde proposé unique. Si l'on abandonne la sémantique relationnelle où les mondes proposés sont multiples (afin de distinguer $\Box$ et $\Diamond$), comment distinguer entre la

---

'Je ne connais pas cet homme' lors du procès. Une déclaration plus appropriée pour la dénégation serait 'Je n'affirme pas connaître cet homme' et exigerait que la non-assertion soit ouvertement déclarée : $A(\ulcorner\neg A\phi\urcorner)$.

[124] Lorsque l'on interprète les symboles de l'idéographie, $\vdash\phi$ se traduit entre autres par 'ϕ est vrai' et $\vdash\neg\phi$ par 'ϕ n'est pas vrai', c'est-à-dire 'ϕ est faux'. Sur la base de cette interprétation, [Greimann 2000] prétend que $\vdash$ n'est pas un opérateur pragmatique mais un opérateur de vérité indiquant la dénotation du contenu jugeable. Nous ne voyons aucune différence pour notre part entre ces deux aspects de la barre de jugement : puisque asserter et assigner une valeur de vérité sont synonymes pour nous, la vérité de ϕ et l'action d'asserter ϕ constituent les deux facettes indissociables d'un seul et même acte *ontique* de jugement.

valuation des assertions vraies et celle des suppositions vraies ? Il semble qu'on ne le puisse plus, dans la mesure où la vérité de $A\phi$ et celle de $S\phi$ signifiera dans les deux cas la vérité de p dans cet unique monde proposé. Que faut-il conclure du fait que la vérité et la fausseté de Ip impliquent la même valeur sémantique pour p : que l'objet d'une dénégation est toujours faux, que l'opérateur I est non-vérifonctionnel (même au sein d'une sémantique des mondes possibles) ?

Deux autres méthodes de valuation sont envisageables : soit l'on introduit une sémantique relationnelle trivalente de type $\{1,½,0\}$, dans laquelle on assignerait la valeur indéterminée ½ à chaque énoncé non-asserté ; soit l'on maintient une sémantique bivalente de type $\{1,0\}$, dans la lignée de [Hintikka 1962], et l'on interprète alors la non-assertion de ϕ comme équivalant à une supposition de $\neg\phi$ ($\neg A\phi \equiv S\neg\phi$). Dans le dernier cas, le maintien de la bivalence entraîne un recours à l'analogie structurelle entre $\Box$ et $\forall$, $\Diamond$ et $\exists$. En suivant cette dernière démarche, on peut démontrer la validité du théorème $A\phi \rightarrow \neg I\phi$ et la non-validité de sa converse $\neg I\phi \rightarrow A\phi$, par exemple, tout en marquant bien la distinction entre les valeurs sémantiques des énoncés (vrais ou faux, symbolisés par 1 et 0) et celles des énonciations (correctes ou incorrectes, symbolisées par + et –) : la valeur sémantique d'une assertion ou d'une supposition n'est pas une valeur de vérité, d'après nous, mais une valeur de correction ; en d'autres termes, nous ne parlerons pas de 'la vérité ou la fausseté de $A\phi$', mais de 'la correction ou incorrection de $A\phi$', déterminée en fonction de la vérité ou la fausseté de ϕ (dans des mondes possibles). Chaque valuation sera par ailleurs une fonction complète : la fonction de valuation classique (locutoire) et la fonction de valuation illocutoire assigneront une et une seule valeur sémantique à chaque formule de type ϕ (vérité ou fausseté) ou $\odot\phi$ (satisfaction ou échec).[125]

Une étude exhaustive de la logique modale illocutoire demanderait qu'une axiomatique accompagne la sémantique relationnelle ci-dessus, afin de mettre à l'épreuve les théorèmes obtenus. Nous les avons seulement suggérés, ici (**K**, $\Box$**T**, **4**, **5**), et la tâche de leur description est renvoyée à plus tard. Nous ne ferons qu'allusion aussi à un contre-exemple possible de notre version assertive de la dénégation, évoquée par [Vernant 2003] sous l'appellation de 'loi de Freud' : $A\phi \rightarrow I\phi$.[126] Tout comme le cas de la duperie de soi (*self-deception*) des logiciens

---

[125] Puisque nous rétablissons l'analogie $\Box/\forall$ et $\Diamond/\exists$ et la sémantique relationnelle de Hintikka, la valuation directe de Gardiès (dont use Vernant) est remplacée par la réfutation à l'absurde et nous proposons des règles sémantiques illocutoires analogues à celles de [Hintikka 1962], sur la base des substitutions [K/A] et [P/S] :

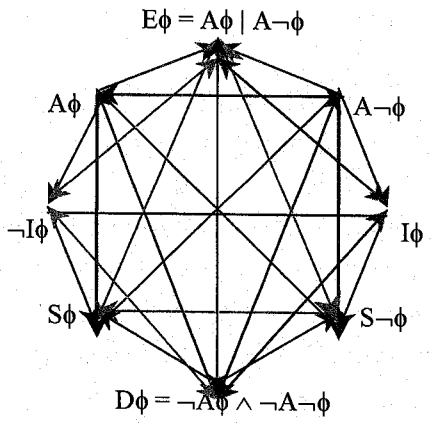
1.	$(A\phi \rightarrow \neg I\phi) \in w$	Prémisse
2.	$(A\phi \wedge I\phi) \in w$	1., contre-prémisse
3.	$A\phi \in w$	2., (C.$\wedge$)
4.	$\phi \in w_*$	3., (C.A*) (pour tout w* alterne à w)
5.	$I\phi \in w$	2., (C.$\wedge$)
6.	$A\neg A\phi \in w$	5., déf. I
7.	$\neg A\phi \in w_*$	6., (C.A*)
8.	$S\neg\phi \in w_*$	7., (C.$\neg$A)
9.	$\neg\phi \in w_*$	8., (C.S*) (pour au moins un w* alterne à w)

Les étapes 4 et 9 se contredisent, donc 2. est invalide en vertu de (C.$\neg$), donc $(A\phi \rightarrow \neg I\phi)$ est valide. La même démarche démontrerait la non-validité de la converse de 1 et de la loi de Russell : $I\phi \rightarrow A\neg\phi$ admet des contre-modèles.

[126] Voir [Vernant 2003], p. 7 : "'Vous me demandez qui peut être cette personne dans le rêve. Ma mère, ce *n'est pas* elle.' Nous rectifions : donc c'est sa mère", cité de S. Freud : "Die Verneinung", trad. française "La négation" in *Résultats, idées, problèmes*, II, Paris, PUF (1985), pp. 135-139. Cette forme de déni rappelle le cas

doxastiques, évoquée par certains comme un contre-exemple à la thèse d'entraînement $K\phi \rightarrow B\phi$, cette loi psychanalytique apparaît davantage comme une 'signification résiduelle' de la notion de dénégation et ne fait que montrer la polysémie des concepts des langues naturelles.

Au total, notre approche illocutoire des modalités épistémiques établit pour l'instant une distinction entre quatre forces illocutoire, toutes déductibles à une seule : l'assertion ($A\phi$, $A\neg\phi$), accompagnée de la supposition ($S\phi$, $S\neg\phi$) et de la dénégation ($A\neg A\phi$). Chacune de ces forces correspondant à un acte de discours dont le but est d'exprimer un état d'esprit particulier tel que la croyance. Nous avons insisté sur la différence entre deux interprétations de la saisie de la pensée ou considération de Frege, symbolisée par le trait horizontal – et que nous avons traduit par l'opérateur modal de supposition S : pour Vernant ainsi que Searle & Vanderveken, cette modalité est une force illocutoire 'neutre' équivalant à l'opérateur locutoire de l'affirmation ⊢ ; pour Kearns, elle constitue déjà une forme d'engagement minimale irréductible à un opérateur locutoire et comparable à la modalité faible $\Diamond$. Quant à la représentation de nos modalités illocutoires au sein d'un polygone des oppositions, nous avons contesté l'hexagone de Vernant pour la raison qu'il éliminait l'assertion négative $A\neg\phi$ et la remplaçait par la dénégation $I\phi$, au risque de créer une confusion indésirable entre $I\phi$ et $A\neg\phi$ lorsqu'il s'agit de définir la forme contraire à $A\phi$. Une autre représentation polygonale, un octogone modal, serait plus conforme à notre propre analyse des notions illocutoires et partirait d'une correction de la 'loi de Russell' (une version illocutoire du tiers exclu), sur la base de l'équivalence $I\phi \equiv A\neg A\phi$.



La nouvelle définition de la dénégation I pose toutefois problème à deux titres, au moins : l'octogone ci-dessus accorde une place à $A\neg\phi$ mais n'en accorde pas à $I\neg\phi$: la relation de contrariété entre $I\phi$ et $I\neg\phi$ n'apparaît donc nulle part. D'autre part, définir $I\phi$ par $A\neg A\phi$ n'exclut pas la possibilité de $A\neg\phi$ alors que la dénégation correspond à un acte de désengagement total exprimant le *doute*, c'est-à-dire un rejet de tout engagement vis-à-vis de ϕ *et de* $\neg\phi$: douter que ϕ ne signifie pas ici que l'on soit porté à croire $\neg\phi$, contrairement à la signification ordinaire de l'expression 'douter que', mais signifie que l'on ne sait pas si ϕ est

de *self-deception* (duperie de soi) de la logique doxastique, que les paraconsistants exploitaient afin de justifier leur système inconsistant.

vrai.[127] Etant donnée notre réduction de l'expression 'savoir si' en termes d'assertion, la dénégation correspond donc à un doute complet de forme $\neg A(\phi \vee \neg\phi) \equiv \neg\phi \wedge \neg A\neg\phi$. Une meilleure définition modale de la dénégation consisterait donc à suivre la voie empruntée par [Hart 1980] et à considérer la dénégation comme l'expression illocutoire de l'attitude de doute, à savoir : $I\phi =_{df} A(\neg A\phi \wedge \neg A\neg\phi)$. Par ailleurs, le doute D du polygone ci-dessus est trop faible et correspond à une autre attitude : l'*indifférence*, équivalant à l'absence complète de jugement ou ignorance à l'égard de ϕ (notre $\neg@A\phi$ précédent). La distinction entre l'indifférence et l'incroyance réside dans l'aspect positif de la dernière : dénier est un acte de *dés*engagement dans lequel le locuteur a un moins une *raison* de ne pas asserter, tandis qu'être indifférent exprime une simple absence d'engagement qui ne présume aucune réflexion préalable sur le contenu propositionnel. On obtient désormais une version de la dénégation D qui rend justice à l'aspect constructif ou actif que lui associait Russell. Les quatre notions illocutoires principales proposées précédemment doivent donc être réarrangées de la façon suivante. Le doute et la dénégation ne font plus qu'un, de sorte que l'ancien opérateur de dénégation I peut donc être assimilé à la notion D du doute de [Hart 1980] ; un opérateur d'ignorance est ajouté pour exprimer le sentiment d'indifférence :[128]

Attitude psychologique	**Acte de discours**	**Symbole**
Croyance forte	Assertion	$\vdash$, A
Croyance faible	Supposition	$-$, S
Incroyance = Doute	Dénégation, Ignorance	$\dashv$, D
Indifférence	$\varnothing$	I

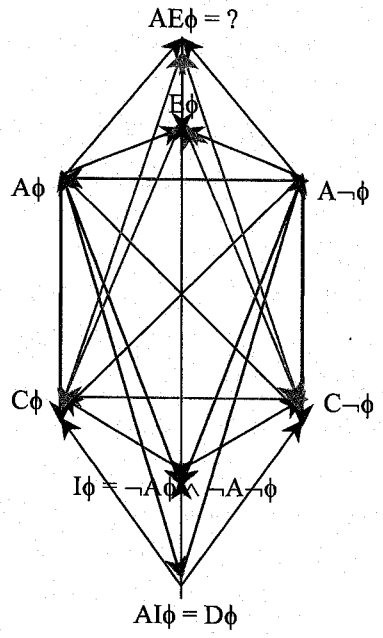
On obtient ainsi un opérateur de dénégation D plus complexe, mais qui devient la contrepartie illocutoire du doute et permet ainsi une symétrie généralisée entre plusieurs attitudes psychologiques et leur expression illocutoire. Le doute dont il est question ici n'est pas une simple hésitation, c'est un doute critique équivalant à l'état stable (actif) du sujet qui décide de suspendre son jugement faute de preuves suffisantes pour asserter quoi que ce soit.[129] Cette nouvelle définition de $D\phi$ par $A(\neg A\phi \wedge \neg A\neg\phi)$ est encore compatible avec la négation intuitionniste $\sim$, que [Gödel 1933] traduisait par $\Box\neg\Box\phi$ et que [Searle & Vanderveken 1985] proposaient de comparer à la dénégation illocutoire : la nouvelle version

[127] La différence entre 'savoir que' et 'savoir si' apparaît dans [Hintikka 1962], où l'auteur signale que l'expression 'il ne sait pas que ϕ' présuppose la vérité de ϕ : $\neg K\phi \wedge \phi$, contrairement à 'je ne sais pas si ϕ' : $\neg K\phi \wedge \neg K\neg\phi$. La même distinction apparaît dans la langue anglaise entre deux formes de doutes : 'to doubt that', qui présuppose la fausseté de ϕ : $B(\neg B\phi \wedge \neg\phi)$, contrairement à 'to doubt whether' : $B(\neg B\phi \wedge \neg B\neg\phi)$. La langue française semble ne contenir que la première version du doute, utilisant 'ignorer si' en lieu et place du barbarisme 'douter si'. Voir à ce sujet la traduction de [Hart 1980] dans l'Annexe 1.

[128] L'indifférence I correspondrait plus ou moins à la considération ou saisie de la pensée de Frege, c'est-à-dire l'opérateur C que nous associons pour notre part à une supposition ou croyance faible.

[129] Sur la relation entre doute et incroyance, voir notamment la traduction de [Russell 1913] dans l'Annexe 1. Notre doute $D\phi$ comme suspension de jugement *justifiée* semble être une force illocutoire située entre le simple doute vacillant de [Russell 1913] (p. 142) et la négation intuitionniste de [Searle & Vanderveken 1985] : à l'instar de la traduction de $\sim$ par [Gödel 1933], $\sim\phi$ signifie la preuve d'une absence de preuve de ϕ et paraît donc plus affirmatif encore que le doute critique. Gödel n'aurait pas dit qu'il 'doutait' de la complétude de l'arithmétique axiomatisée, mais qu'il était convaincu de son incomplétude.

est plus forte et *implique* la version I précédente : $A(\neg A\phi \wedge \neg A\neg\phi) \rightarrow A\neg A\phi$. Le résultat de ce nouvel ensemble de modalités illocutoires pourrait donner un polygone de forme hexagonale tel que celui-ci :



Bien qu'enrichi de deux nouveaux opérateurs AE et I, le polygone perd deux de ses côtés précédents mais certaines relations (AEϕ-Eϕ, Dϕ-Iϕ) sont 'recouvertes' par d'autres (AEϕ-Dϕ, Dϕ-AEϕ). Une représentation géométrique plus adéquate de cet hexagone exigerait de passer d'une figure de deux à trois dimensions, soit une sorte de gemme où les relations AEϕ-Eϕ et Dϕ-Iϕ seraient mises en évidence par une relation spatiale de profondeur. Hormis ce détail géométrique,[130] il resterait par ailleurs à trouver une contrepartie illocutoire à AEϕ : quelle expression naturelle lui associer, afin d'obtenir une opposition symétrique avec le doute Dϕ ? La restriction des formules bien formées dans $\mathcal{L}_{f.AS}$ à des actes illocutoires devant permettre par ailleurs de caractériser ce système illocutoire comme de type **K□T45**, une représentation polygonale devrait donner ce système modal fort. Mais le problème réside dans la représentation géométrique des modalités *itérées*, qui exigerait une analyse plus détaillée mais que nous ne préciserons pas davantage ici.

Ce système peut être mis à l'épreuve à partir d'une traduction des formules précédentes, parmi lesquelles les contreparties illocutoires du tiers exclu (loi de Russell) et de la double négation. On pourra vérifier que ces formules demeurent non valides dans $\mathcal{L}_{f.AS}$.

Un mot sur les attitudes d'assentiment et de dissentiment : si l'assentiment est un état d'esprit exprimé par l'assertion, par quel opérateur doit-on exprimer le dissentiment ? [Vernant 2003] donne une claire idée du type d'opposition qui les relie :

> Ockam distinguait déjà jugement (assentiment ou dissentiment) et simple appréhension qu'il nommait *proposition neutre* : 'Quelqu'un peut appréhender une proposition et cependant ne lui donner ni assentiment ni dissentiment, comme c'est patent avec les propositions neutres', cf. *Scriptum in librum primum Sententiarum, Ordinatio, prologus et distinctio prima*, I., Prol., qu.

[130] Pour un traitement des oppositions à trois dimensions, voir la référence donnée au Chapitre 1 note 56.

1, p. 16. On se souvient aussi que dans la *disputatio* traditionnelle trois attitudes étaient possibles : *concedo*, *nego*, *dubito*.[131]

Assentiment et dissentiment sont clairement conçus comme des jugements à part entière : l'assentiment à l'égard de ϕ correspond à l'assertion $A\phi$, c'est-à-dire l'opérateur d'affirmation externe de [Bochvar 1938] : 'ϕ est vrai'; le dissentiment à l'égard de ϕ correspond à l'assertion $A\neg\phi$, c'est-à-dire l'opérateur de négation externe de Bochvar : 'ϕ est faux'. Quant au doute, sentiment auquel Bochvar n'associe pas d'opérateur particulier et qui n'exprime pas une quelconque reconnaissance (de vérité ou de fausseté), il ne correspond ni à une assertion ni à notre dénégation. Or ce doute, si l'on suit notre définition de D en termes d'assertion $A(\neg A\phi \wedge \neg A\neg\phi)$, peut être assimilée à l'assentiment d'un non-assentiment ou à l'affirmation d'un doute. C'est là un état d'esprit peut-être difficile à distinguer du simple assentiment : comment distinguer une assertion simple des assertions itérées, lorsque l'on applique la méthode de traduction radicale basée sur le comportement observable de l'interprété ? Si une assertion itérée de type D est impossible à identifier parce que trop complexe, la charité logique semble mise en défaut ; si tel est le cas, notre logique modale illocutoire ci-dessus serait-elle en mesure d'affiner la méthode de traduction radicale et d'enrichir ou compléter la charité logique de Quine par d'autres types de comportement ? Pas plus que le cas de l'abstention, évoqué dans [Quine 1974], car la dénégation n'est pas un état d'esprit susceptible d'être exploité lors d'une traduction basée sur le comportement explicite; pour cette raison, seuls les comportements affirmatifs $A\phi$ (dire 'oui') et $A\neg\phi$ (dire 'non') peuvent être utilisés pour mener une traduction à bien et déterminer les significations des déclarations d'une langue indigène. En d'autres termes, la valeur ½ des sémantiques algébriques intuitionnistes exprime un état d'esprit inopérant pour Quine, et seules les forces illocutoires explicites satisfont la méthode de traduction radicale puisqu'elles symbolisent des actes d'énonciation publics. En un mot : l'ignorance I et la dénégation D ne sont pas exploitables pour le traducteur de Quine, et l'enrichissement des modalités épistémiques proposé ci-dessus ne constitue en rien un contre-argument au conservatisme logique de Quine, essentiellement pragmatique et fondé sur l'aspect déclaratif du langage.

En conclusion, les énoncés des langues naturelles correspondent symboliquement à une formule modale de type $A\phi$, où A désigne l'acte d'assertion ou de considérer comme vrai et ϕ une fonction propositionnelle. Comme son nom l'indique, cette fonction n'a pas de valeur de vérité en elle-même ; elle est susceptible de devenir vraie ou fausse mais doit être complétée par un acte illocutoire avant d'être qualifiable d'énoncé vrai ou faux. Par ailleurs, le fait de traiter tout énoncé comme une formule modale rend compte de la notion de 'complétive' attachée d'habitude à la notion de proposition : l'expression 'que ϕ' doit être complétée par une expression de type 'C'est le cas' ou 'Il est vrai' afin de représenter un jugement ($A\phi$), c'est-à-dire un énoncé vrai ou faux. Par conséquent, tout énoncé des langues naturelles exprime une assertion ; dans le cas contraire, le sens de l'énoncé doit être affaibli par d'autres expressions explicitement modales de type 'il est possible que' ou 'on suppose que' ($S\phi$).

[131] [Vernant 2003], p. 4 note 20.

Pour finir, traiter tout énoncé des langues naturelles comme l'expression d'une forme modale permet de rendre justice au principe du tiers exclu, mais à condition de lui donner une autre forme que la 'loi de Russell' $D\phi \rightarrow A\neg\phi$. Dans une perspective illocutoire des actes de discours, ce principe prend en effet la forme logique '$A\phi \vee \neg A\phi$' et non la forme rédhibitoire '$A\phi \vee A\neg\phi$'. Dans [Gödel 1933], la critique intuitionniste du tiers exclu reposait sur la seconde interprétation ; or celle-ci n'exprime pas une opposition entre deux formules contradictoires mais entre deux formules contraires. L'erreur commise semble être due à l'explication courante du tiers exclu en termes de valeurs de vérité : en vertu du principe de bivalence, toute *proposition* est soit vraie soit fausse, mais ce principe n'implique pas que tout *énoncé* est soit vrai soit faux. Le point de vue locutoire des énoncés oblige à rejeter le tiers exclu en cas d'indécision de la part du locuteur, puisque le tiers exclu signifie alors que, pour tout énoncé ϕ, soit ϕ soit sa contradictoire $\neg\phi$ est vrai. Mais d'un point de vue illocutoire, le tiers exclu signifie que, pour tout énoncé $A\phi$, soit $A\phi$ soit sa contradictoire $\neg A\phi$ est vrai. Le logicien intuitionniste s'accorderait avec cette interprétation illocutoire du tiers exclu, le seul grief portant sur le principe de bivalence ; si celui-ci signifie que tout énoncé est l'expression d'un assentiment ou d'un dissentiment, alors il n'est pas valide dans la mesure où la position tierce de l'abstention est possible pour les énoncés : l'absence d'assentiment et de dissentiment a été admise, puisqu'elle correspondait plus haut à la négation illocutoire ou dénégation.[132] Si en revanche la bivalence est un principe portant sur des *propositions*, c'est-à-dire l'entité idéale considérée comme la signification commune à plusieurs énoncés, alors ce principe ne nous intéresse pas puisqu'il concerne des entités fictives : '*être* vrai' ou '*être* faux' n'a pas de sens dans une perspective illocutoire des actes de discours, puisque seuls les actes de 'considérer comme vrai' et de 'considérer comme faux' y sont pris en compte. Etre vrai ou faux, c'est être vrai ou faux pour un ou plusieurs locuteurs, autrement dit, et l'accord des locuteurs sur la signification d'un énoncé peut varier selon la clarté des expressions qu'il contient.

En conséquence, le principe du tiers exclu peut être admis indépendamment du principe de bivalence, que celui-ci porte sur des énoncés ou sur des propositions. Si l'on redéfinit le tiers exclu comme le principe selon lequel toute formule bien formée ϕ est satisfaite ou insatisfaite, les actes d'assertion et de dénégation permettent de l'exprimer sans l'introduction de valeurs de vérité. On obtient ainsi un *tiers exclu illocutoire* valide dans $\mathcal{L}_{\text{f.AS}}$:

$$\vDash_{\mathcal{L}\text{f.AS}} A\phi \vee \neg A\phi,$$

quelque peu différent de la relation contraire entre l'assertion de ϕ et sa dénégation :

$$\vDash_{\mathcal{L}\text{f.AS}} A\phi \rightarrow \neg D\phi$$

[132] En règle générale, le principe de bivalence est conçu comme le principe métalogique selon lequel toute proposition est soit vraie soit fausse. Mais si toute valeur sémantique n'est pas nécessairement le vrai ou le faux, cela a-t-il un effet sur la portée de la bivalence ? L'hypothèse posée par [Suszko 1977] d'une distinction fondamentale entre deux valeurs logiques, le désigne et le non-désigné, permettrait de rétablir le principe de bivalence en redéfinissant le sens des 'valences', même en cas de formules indécidées : toute formule est considérée comme satisfaite (*désignée*) ou pas (*non-désignée*) ; sous cette interprétation, le tiers exclu illocutoire est valide puisqu'il signifie que tout énoncé déclaratif est soit satisfait (de valeur sémantique +) soit insatisfait (de valeur sémantique –). On obtient en effet la valuation : $v(A\phi) \vee v(\neg A\phi) = (+) \vee (-) = +$. L'ambiguïté de la bivalence et du tiers semble reposer sur une restriction des valeurs sémantiques aux valeurs de vérité, ce que nous avons contesté dans ce travail puisque ces valeurs ne concernent que la catégorie des énoncés et non celle fondamentale des énonciations (+, –).

Ce théorème illocutoire corrige la loi de Russell, en ce qu'il modifie le rapport de condition à conséquent et la portée de la négation : non pas 'si p est déniée alors non-p doit être assertée', mais 'si p est assertée alors p ne doit pas être déniée'.

Un traitement illocutoire du tiers exclu et de la négation en général pourrait être utilisé de façon à traiter d'autres paradoxes fameux : les paradoxes de l'implication matérielle, la bataille navale d'Aristote, le Menteur, l'exécution surprise, etc. Ce travail est remis à plus tard, et nous retiendrons simplement deux idées générales de cette section.

Premièrement, tout énoncé des langues naturelles devrait être symbolisé par une formule modale de type $\odot\phi$, où $\odot$ désigne l'assertion ou la supposition de vérité, et l'assignation d'une valeur est effectuée par le locuteur. Pour cette raison, le phénomène d'opacité référentielle est un problème contingent de compétence linguistique.

Deuxièmement, la notion de négation devrait être divisée entre sa version externe de la dénégation et sa version interne du dissentiment. De ce point de vue, les analyses proposées par la logique classique et la logique épistémique ne sont pas aussi distinctes qu'on les imagine d'habitude. En effet, ces deux logiques philosophiques portent sur les énoncés des langues naturelles et, pour cette raison, une logique modale illocutoire telle que $\mathcal{L}_{f.AS}$ semble en mesure de montrer leur *complémentarité* : la logique classique porte sur la vérité et la fausseté, la logique épistémique sur la reconnaissance de vérité et de fausseté ; or puisque, de notre point de vue, les attitudes épistémiques expriment des déclarations de vérité, une logique des déclarations peut être considérée comme la réunion en un seul langage formel des formules classiques et épistémiques.

L'argument de la non-vérifonctionnalité des opérateurs modaux nous a porté à croire que les énoncés déclaratifs et épistémiques sont de nature logique distincte ; ce qu'ils ne sont pas en réalité, dès lors que l'on considère leur appartenance à une même classe d'énoncés : les assertifs, et qu'on les caractérise par une même forme logique illocutoire. Sans être entré dans les détails de la logique illocutoire de [Searle & Vanderveken 1985], force pour nous de constater que les énoncés de croyance et de connaissance sont partagés entre plusieurs buts illocutoires : affirmer la réalité d'un fait (but assertif), user d'arguments d'autorité par une annonce de savoir (but déclaratif), ou encore exprimer son état d'esprit (but expressif). L'analyse de ces énoncés est donc complexe, ce qui explique la délicatesse de leur analyse logique et confirme l'appartenance des modalités épistémiques à plusieurs jeux de langages indépendants les uns des autres. Leur insertion dans une théorie des actes de discours, conjointement aux énoncés déclaratifs, a permis en tout cas de présenter la logique classique et la logique épistémique comme deux logiques philosophiques se chevauchant et portées sur un même mode de discours : le *mode assertorique*, au sein d'une large panoplie d'actes de discours. Le mérite de cette approche illocutoire a été partiellement confirmé ici, par sa capacité à traiter de plusieurs *puzzles* de la logique philosophique.

Conclusion

Il ne suffit pas de connaître les conditions de vérité d'une expression pour la comprendre, certes ; cela ne veut pas dire pour autant qu'aucune analyse logique du langage ordinaire ne puisse être effectuée. Contrairement à une idée répandue, selon laquelle la compréhension d'une expression dépendrait uniquement de la connaissance de sa signification, *c'est-à-dire* de ses conditions de vérité, le passage à une analyse illocutoire des actes de discours insiste sur la signification *dans l'usage*. La distinction est marquée de ce fait entre l'usage et la signification, dans le cadre d'une théorie de la communication : la compréhension, c'est la signification *plus* l'usage ; le pluralisme logique peut être conçu comme un effet de l'équation incorrecte 'compréhension = signification = conditions de vérité', qui consiste à modifier les propriétés des opérateurs logiques au moindre dysfonctionnement de leurs conditions de vérité lors d'un contexte de discours particulier.

Contre l'explosion du nombre des systèmes de logique philosophique, nous avons tenté de tirer deux leçons principales d'une analyse illocutoire des modalités épistémiques : qu'elle est réductible à une logique *assertorique*, à condition de traiter l'opérateur modal comme un élément primitif dans le processus de signification ; que la signification des opérateurs logiques n'est pas modifiée par les différents contextes d'usage, compte tenu de notre distinction précédente entre la signification d'une expression linguistique et sa compréhension au sein d'un acte de discours. La signification 'non-littérale' d'une expression n'oblige pas à transformer sa signification littérale et les règles des constantes logiques qu'elle contient, par conséquent ; une telle économie des opérations logiques permet d'éviter (ou de restreindre, tout au moins) le relativisme logique et l'impression désagréable d'inconstance des constantes logiques qui s'en dégage. C'est l'usage de ces constantes qui est inconstant, mais leur usage n'est pas leur signification ('de base', comme les appelle Hintikka), et la logique épistémique a servi de support pour illustrer cette argumentation générale. Avec pour mot d'ordre l'alignement des attitudes épistémiques sur le mode assertorique ; revenons sur les détails de cette analyse.

'C'est le ton qui fait la musique', dit-on parfois avec légèreté et sans trop prendre au sérieux cet adage populaire. Loin d'être léger, ce proverbe du *sens commun* insiste à juste titre sur le rôle joué par l'intonation de la voix dans la signification attribuée à un énoncé ordinaire. Il nous rappelle également que la signification transmise par un acte de discours n'est pas toujours explicitée par les mots et qu'il faut souvent plus qu'une interprétation littérale des signes écrits pour se comprendre au sein d'un discours. Sans aller jusqu'à cet aspect oral des langues, la grammaire naturelle indique la présence de différents *modes* de discours par le processus de flexion des radicaux : les préfixes et suffixes des noms communs, mais aussi la terminaison d'un verbe symbolisent des actes de signification particuliers qui peuvent être formalisés au sein des logiques modales philosophiques.

Le résultat de cette analyse formelle des langues naturelles est une combinaison des analyses formelles et informelles du discours : à la formalisation des modes s'ajoute une interprétation contextuelle des types de discours, dont la direction d'ajustement et les degrés

de force sont répertoriés par la logique illocutoire. Au total, l'analyse logique des modalités du discours s'accompagne d'une typologie des actes qui n'implique pas une révision des constantes logiques en fonction du contexte d'usage.

Une opération est effectuée sur un contenu de discours et produit un effet ; à travers ce processus, c'est la procédure algébrique d'application ou d'interprétation d'un ensemble de formules dans un modèle qui se retrouve au sein des logiques philosophiques. Nous avons tenté d'en faire l'illustration avec le cas des attitudes épistémiques. A partir d'une paraphrase des opérateurs modaux en termes d'attitudes propositionnelles, nous avons prétendu dans ce travail que la logique épistémique s'adresse aussi bien aux 'propositions' qu'aux 'attitudes propositionnelles' ; dans les deux cas, nous avons affaire à des *énoncés assertoriques* qui manifestent des actes d'énonciation et peuvent être formalisés comme des expressions modales gouvernées par un marqueur de force illocutoire. Plus, nous avons prétendu qu'un énoncé des langues naturelles manifeste toujours un certain acte d'énonciation de la part d'un locuteur, donc une attitude et une interprétation conjointes. Il n'y a donc pas de différence substantielle entre logique épistémique et logique assertorique (ou déclarative), dans la mesure où leurs formules partagent le même *meaning*, c'est-à-dire le même vouloir-dire ou sens déterminé par une direction d'ajustement (entre les expressions du langage et les choses du monde) ; telle est la conclusion 'philosophique' de ce travail. Mais avant de pouvoir avancer cette thèse, il aura fallu revisiter l'ensemble des problèmes et discussions philosophiques gravitant depuis 1962 autour de la logique épistémique.

Si l'on observe le chemin parcouru dans ce travail, l'étude des modalités épistémiques nous aura conduit du problème récurrent de la *non-vérifonctionnalité* des opérateurs modaux à la solution bien moins courante d'une théorie formelle des *actes de discours*. Le but de l'ensemble aura été d'unifier les énoncés des langues naturelles au sein d'une seule et même logique modale illocutoire, portant notre attention sur la catégorie des *modalités épistémiques* : il n'y a pas de différence catégorielle entre les *énoncés déclaratifs* tels que 'Socrate est sage', que l'on associe d'habitude à la logique non-modale classique ou standard, et les énoncés modaux épistémiques de type 'Je crois que Socrate est sage', tous deux ayant une même visée intentionnelle : celui de dire la vérité, avec succès ou non. [von Wright 1951] avait anticipé ce traitement uniforme en considérant les concepts de vérité et de fausseté comme des opérateurs à part entière, par analogie avec les quantificateurs existentiels. Dans tous les cas, le processus de signification dans une langue naturelle procède à partir de l'application complexe d'opérations sur des arguments, et c'est pour avoir négligé ce processus systématique que les énoncés affirmatifs et négatifs ont été opposés à tort au discours des modalités. Tout acte de discours revêt une forme modale ; la modalité 'montre' la projection d'une valeur sur une formule qui sinon n'en aurait pas et nécessite cette opération intentionnelle pour signifier quelque chose.

Pour cette raison, nous avons assimilé la totalité des énoncés des langues naturelles à des modalités de forme logique $\odot\phi$; prenant contre-pied avec le symbolisme non-modal de la logique classique, nous en avons conclu que la formule 'ϕ' n'est pas une proposition mais un signe propositionnel dépourvu de sens tant qu'aucune opération ne lui est appliquée à

l'occasion d'un acte de discours ; les modalités épistémiques n'en sont qu'un exemple parmi tant d'autres répertoriés en logique illocutoire. Car il y en a bien d'autres (promesse, question, regret, volonté, etc.), mais notre étude se sera limitée au cas des expressions épistémiques.

Rappelons les principaux résultats des trois chapitres :

Chapitre 1

a) La logique épistémique doit son apparition à la création des *modèles de Kripke* et de sa sémantique relationnelle ('des mondes possibles') : de préférence aux applications limitées des matrices algébriques, une sémantique relationnelle permet d'interpréter les expressions modales dans une pluralité de mondes capables de représenter le point de vue d'un sujet.

b) Le débat créé autour du statut ontologique des mondes possibles et de la quantification sur ces mondes ne concerne que peu ou prou la logique épistémique : celle-ci présente les individus possibles comme ce que les individus réels *pourraient* être en l'état de nos connaissance, mais cette variation des propriétés ne conduit pas forcément à une inflation ontologique (on peut faire varier les propriétés des habitants du monde réel sans inventer des êtres fictifs).

c) La difficulté philosophique qu'une logique modale épistémique peut mettre en évidence est le fossé qui sépare la *proposition* de l'*énoncé* : son mérite doit être de formaliser les moyens par lesquels deux agents en viennent à exprimer la même pensée. Le sens [*Sinn*] de Frege est l'incarnation idéale d'un *sens commun*, d'une signification supposée partagée par un ensemble d'expressions et qui transcenderait non seulement les locuteurs mais les langues naturelles. Contre la 'saisie' de ce sens commun, Quine reprochera par la suite la postulation d'entités mythiques détachées du contexte d'apprentissage des langues. Parce que ce sens 'incarne' la signification dans une totalité de mondes *métaphysiquement* possibles, il a été exclu de notre approche d'inspiration nominaliste et remplacé par une interprétation sociale et constructive des expressions naturelles. La communauté de sens n'est pas une entité idéale mais un produit social.

Chapitre 2

a) Si les principaux auteurs de la logique mathématique (Frege, Russell, Wittgenstein, Quine) ne pouvaient admettre l'idée d'une logique épistémique, c'est en raison de la confusion que celle-ci était susceptible de créer entre les domaines de la logique et de la *psychologie*. Les moyens formels limités dont ils disposaient à leur époque (à l'exception de Quine) explique en grande partie cette réticence vis-à-vis des modalités épistémiques : l'absence de sémantique relationnelle capable d'interpréter les opérateurs modaux, leur traitement de la vérité comme une relation bi-univoque entre le langage et le monde étaient le symptôme d'une certaine idée de la logique : la *logique comme langage universel*, par opposition au langage comme calcul. Le maintien de la vérité logique comme vérité dans un modèle unique et sans structure a constitué une sorte d'obstacle verbal contre l'émergence des logiques philosophiques.

b) La logique épistémique porte quoi qu'il en soit sur des expressions polysémiques, d'où une difficulté dans le choix du formalisme et du modèle appropriés : 'croire' et 'savoir'

semblent irréductibles à des constantes logiques univoques, d'où la réticence courante de nombreux philosophes à admettre le système de [Hintikka 1962]. Le meilleur moyen de paraphraser les modalités épistémiques a été, selon nous, l'explication *illocutoire* des théorèmes épistémiques : les paradoxes de l'*omniscience logique* (axiome **K**) et de *Moore* (axiome **4**) deviennent inoffensifs lorsque les énoncés sont interprétés comme des *énonciations*. Quant à la clause de vérité (axiome **T**), elle devient l'expression d'une condition de sincérité (de véracité) des assertions et n'oblige plus à passer par les fourches caudines du sceptique.

c) Ce même traitement illocutoire en termes de modalités 'assertoriques' a été appliqué au problème de l'*opacité référentielle* et de l'*identification* des individus dans les mondes possibles : il ne semble pas exister de critères objectifs capables de garantir l'identification réussie d'un objet ; l'attribution d'une attitude à autrui reste tributaire d'une attribution de compétence linguistique minimale de sa part, et le lien établi entre une expression et son objet est parfois méconnaissable d'un point de vue tiers. Le critère d'assertion explicite d'une identité entre deux objets a été réclamée pour éviter l'opacité référentielle, quitte à réduire l'intentionnalité à des attitudes *de dicto*.

Chapitre 3

a) Les concepts de vérité et de connaissance sont au cœur de la logique épistémique mais également au sein de deux théories sémantique : le *réalisme sémantique*, dans lequel la vérité est indépendante des conditions de connaissance ; *l'anti-réalisme sémantique*, dans lequel la vérité dépend de ces conditions et devient un synonyme d'assertabilité garantie. Chacune de ces théories se heurte à des problèmes de formalisation des concepts : contre le réaliste, le *paradoxe de Gettier* montre que la vérité n'est pas toujours synonyme de croyance vraie justifiée ; contre l'antiréaliste, le *paradoxe de Fitch* montre que la redéfinition de la vérité en termes de connaissabilité produit une nouvelle forme d'omniscience. La réduction des énoncés à des actes d'énonciation a permis de dissoudre ces difficultés conceptuelles.

b) Le *Principe de Charité* a été utilisé pour montrer le rôle central des attitudes épistémiques dans l'interprétation des énoncés : le lien *intentionnel* établi entre une phrase et sa signification est induit à partir d'un manuel de traduction minimal, et ce manuel repose sur des opérations logiques fondamentales. La tentation de réviser les constantes logiques a été repoussée, maintenant la logique épistémique au sein des *logiques enrichies* et *non déviantes*. Si la réussite d'une traduction passe par d'autres moyens d'analyse que la *théorie des fonctions de vérité*, c'est parce qu'il existe d'autres types de discours que l'assertion ; notre distinction entre la signification linguistique des expressions et leur usage dans un contexte particulier a été empruntée à la philosophie du langage de Grice, permettant d'expliquer l'échec d'une traduction vérifonctionnel sans impliquer l'abandon des fonctions de vérité.

c) L'orientation illocutoire des attitudes débouche sur une réduction des modalités épistémiques à des modalités *assertoriques* : tous les actes de discours visant à dire la vérité correspondent à des actes d'assertion, dont la forme logique a été étudiée sur la base de l'idéographie de Frege. Logique assertorique et logique épistémique correspondent ainsi à un même processus d'interprétation que la logique standard ne prend pas en compte, gradué selon deux *forces illocutoires* principales que sont l'*assertion* (le jugement catégorique) et la

supposition (le jugement hypothétique). Plusieurs notions peuvent être redéfinies dans le cadre d'une théorie des actes de discours, parmi lesquels la *dénégation* ; celle-ci a été distinguée de la négation locutoire, formalisée comme une contrepartie illocutoire de la *contingence nécessaire* et expliquée comme l'expression du doute.

La réduction de la logique épistémique à une logique des modalités assertoriques mérite une analyse plus poussée que celle menée ici : nous nous sommes tenus à une lecture approchée des travaux de Hintikka et de ses détracteurs, alors qu'une application plus avancée des actes illocutoires devrait en dire plus que nous n'en avons dit sur les concepts de connaissance et de croyance : leur but illocutoire, leur direction d'ajustement entre le langage et le monde, leur condition de succès, leur degré de force illocutoire ne vont pas de soi et ouvrent la voie à une distinction plus fine des subtilités du langage ordinaire. Non seulement la distinction méthodologique entre les philosophes du langage ordinaire (Ryle, Austin, Grice) et les philosophes formels (von Wright, Hintikka, Montague) est surmontée, mais une telle logique illocutoire permet de relever le défi posé par Wittgenstein au sujet des jeux de langage : il est possible de formaliser et rationaliser un tant soit peu l'ensemble de ces jeux, sans aller jusqu'à prétendre épuiser toute leur richesse ; dans la mesure où la logique philosophique propose d'appliquer *un* formalisme et *une* algèbre aux langues naturelles, nous avons considéré pour notre part cette logique illocutoire comme la direction la plus prometteuse à cet égard et la plus à même d'offrir ce que tout philosophe est en droit d'attendre de ses analyses : davantage d'unité et de généralité dans la construction théorique ; la réduction des modalités assertoriques à une catégories d'actes illocutoires parmi trois autres (voir [Searle & Vanderveken 1985]) plaide en ce sens.

Il reste à examiner les relations logiques entre les assertions et d'autres modalités du discours tels que les interrogatifs, mais aussi les exclamatifs ou les promissifs ; il reste également à vérifier la fécondité de cette analyse pragmatique par rapport aux programmes actuels des sémantiques de connaissance commune, de révision de croyance ou des jeux dialogiques. Mais il ne fait pas de doute que les effets pragmatiques d'une déclaration de connaissance sont intimement liés aux opérateurs d'annonce publique de van Benthem ainsi qu'aux jeux dialogiques.

Les médiévaux trouveraient en tout cas une explication plausible à leur problème des modes subalternes : une condition préparatoire des actes de discours est que l'on ne puisse signifier par eux que ce qui convient à leur usage ; sans être tout à fait artificielle, la logique illocutoire offre aux analyses formelles et informelles du langage ordinaire un compromis qui permet de rétablir une certaine unité entre les règles de la logique et de la communication. Puisse-t-elle satisfaire ainsi les besoins du langage ordinaire et régler un maximum de paradoxes sémantiques.

ANNEXE 1

Logique modale épistémique

The Collected Papers of Bertrand Russell, Vol. 7 : "Theory of Knowledge : The 1913 Manuscript", pp. 136-143

BERTRAND RUSSELL

CROYANCE, INCROYANCE ET DOUTE

Jusqu'ici, nous nous sommes intéressés aux attitudes non-dualistes à l'égard des objets, et aux propriétés non-dualistes des objets. Nous en arrivons finalement aux attitudes et aux propriétés dualistes : l'attitude dualiste de croyance et d'incroyance dans ce chapitre, et la propriété dualiste de vérité et de fausseté dans le suivant.

Lorsque je parle de "croyance", j'entends la même sorte de fait que celle que l'on qualifie d'ordinaire de "jugement". Je préfère le terme "croyance", parce qu'elle suggère avec bien plus de fermeté un événement daté particulier qui peut être étudié empiriquement par la psychologie. Le terme "jugement", en revanche, est employé généralement par les idéalistes, et il sert à brouiller la distinction entre psychologie et logique. Le jugement est fondamental en logique, d'après les idéalistes, mais c'est quelque chose qui ne pourrait pas subsister s'il n'y avait pas d'esprits, bien qu'il soit indépendant de cet esprit-ci ou cet esprit-là. Dans notre philosophie, nous ne pouvons pas donner une déclaration de ce point de vue qui sera de prime abord défendable. Mais nous pouvons peut-être l'illustrer en définissant un jugement comme le fait "que telle proposition est crue". Ce qui veut dire que le jugement "A parvient à B" sera le fait "Il y a un sujet qui croit que A parvient à B" (si c'est un fait). Je ne suggère pas que c'est là ce que les idéalistes entendent par jugement, mais que c'est l'approche la plus similaire à leur point de vue qui soit formellement possible pour nous. Sous cette forme, elle présuppose un examen de la croyance, et c'est donc sur la croyance que nous porterons désormais notre attention.

A un certain égard, le terme "croyance" est moins révélateur de points de vue corrects que le terme "jugement". Il est évident que le jugement comporte la sorte de chose qui est vraie ou fausse, et pas simplement des existants particuliers telles que des tables et des chaises. A l'inverse, on suppose très souvent que "croyance" peut porter sur le même type d'objets que la sensation. "Voir c'est croire" incarne une confusion dont un grand nombre de philosophes ont souffert. Cette confusion sous-tend par exemple l'interprétation par Hume de la Croyance,[1] qui commence par ces mots :

> L'idée d'un objet est une partie essentielle de la croyance que l'on en a, mais elle n'en constitue pas l'ensemble. Nous concevons un grand nombre de choses, auxquelles nous ne croyons pas ; **(137)**

Suite à quoi elle passe tout naturellement à la fameuse conclusion :

[1] *Traité*, Livre I, Partie III, Section VII.

> Comme la croyance ne fait rien d'autre que faire varier le mode par lequel nous concevons un objet, elle peut seulement conférer à nos idées une force et une vivacité supplémentaires. Une opinion ou croyance peut être définie avec le plus de précision comme ceci : UNE IDEE VIVACE LIEE A OU ASSOCIEE A UNE IMPRESSION PRESENTE.

Cette explication suppose que l'objet d'une croyance est le même que l'objet d'une idée. Les confusions qui ont amené à cette supposition sont nombreuses, mais puisqu'elle a eu des conséquences désastreuses en philosophie il est important qu'un nombre suffisant d'entre elles soit avancé.

En premier lieu, notre théorie admet qu'il y a un type de fait mental appelé "comprendre une proposition", qui ne comporte pas l'opposition de la croyance ou incroyance mais partage la neutralité de la simple accointance, et qui consiste toutefois dans un complexe dont la forme est exactement la même que la croyance. Lorsqu'un certain sujet comprend une proposition et lorsqu'il la croit, les deux complexes diffèrent selon nous uniquement par le fait que comprendre et croire sont des relations différentes : la croyance à la proposition résulte du fait de substituer la croyance à la compréhension dans le complexe qu'est la compréhension de cette proposition. Mais si notre théorie précédente sur ce point portait une quelconque ressemblance avec la vérité, comprendre une proposition est un complexe d'une forme totalement différente de l'accointance, excepté dans le cas de la compréhension d'une forme logique pure. A l'exception de ce seul cas, croyance et compréhension n'ont pas un objet unique, la "proposition", mais une pluralité d'objets, unies au sujet dans une relation *multiple*. Donc bien qu'il y ait une attitude neutre, à savoir la compréhension, qui donne des complexes de même forme que les croyances, cette attitude n'est toutefois pas fondamentalement différente de celle d'"avoir une idée d'un objet" qui, quel que soit le sens dans lequel nous puissions l'admettre comme fondamentale, doit être identifiée à l'accointance, que ce soit de la sensation, de la mémoire, de l'imagination ou (dans le cas d'un universel), de la conception.

Hume se défend contre l'accusation de confondre jugement et conception, dans une annotation à la définition susmentionnée de la croyance. Il dit :

> Nous pouvons prendre ici l'occasion d'observer une erreur très remarquable qui, étant fréquemment inculquée dans les écoles, est devenue une sorte de maxime établie et est universellement reçue par tous les logiciens. Cette erreur consiste dans la division vulgaire des actes de la compréhension entre *conception*, *jugement* et *raisonnement*, ainsi que dans les définitions que nous en donnons. La conception est définie comme la simple inspection d'une ou plusieurs idées ; le jugement, comme la séparation ou l'union d'idées différentes ; le raisonnement, comme la séparation ou l'union **(138)** d'idées différentes par l'interposition d'autres idées, lesquelles montrent la relation qu'elles entretiennent les unes avec les autres. Mais ces distinctions et définitions son fautives au sein d'articles très considérables. En *premier* lieu, il est loin d'être vrai que dans tout jugement que nous formons nous unissions deux idées différentes ; puisque dans la proposition qui suit : *Dieu est*, ou en fait dans toute autre qui concerne l'existence, l'idée d'existence n'est pas une idée distincte que nous unissons à celle de l'objet et qui soit capable de former une idée

composée par l'union ... Que nous considérions un objet particulier, ou bien plusieurs ; que nous nous arrêtions sur ces objets, ou que nous passions des uns aux autres ; et sous quelque forme ou ordre dans lequel nous les inspecterions, l'acte de l'esprit ne dépasse pas la simple conception.

L'exemple des propositions existentielles, que Hume invoque dans le passage ci-dessus, a été sans aucun doute un puissant allié de la théorie qui identifie jugement et conception. William James, dont le monisme neutre comporte selon nous la même identification, y a été conduit également, je pense, par la considération des jugements existentiels. Nous semblons juger que les objets des sens existent, et ne rien ajouter à ce qui a déjà été donné dans les sens en jugeant ainsi. Mais le fait est que la conception d'ensemble de l'existence est le résultat d'une confusion entre descriptions et noms propres authentiques. Il est dépourvu de sens de dire d'un *ceci* [*this*] effectivement donné, un objet d'accointance, qu'il "existe". Mais le même terme qui, a un moment donné, est utilisé en guise de nom propre pour un objet donné, peut être utilisé comme une description le moment d'après. Nous pouvons dire "ceci existe", ce qui signifie "l'objet de mon attention présente existe", ou "l'objet que je suis en train d'indiquer existe". Le mot "ceci" a cessé ici de fonctionner comme un nom propre et est devenu un terme descriptif, dans lequel un objet est décrit par ses propriétés, et la question qui peut être soulevée est celle de savoir s'il y a un tel objet, puisque l'on peut réaliser des descriptions auxquelles rien ne corresponde. Lorsque nous disons "le roi d'Angleterre existe", nous ne prononçons pas une tautologie, ni n'ajouter quoi que ce soit à "Le roi d'Angleterre" ; nous disons qu'il y a une personne qui peut être décrite ainsi. "Le Roi de France n'existe pas" est vraie, mais elle serait une contradiction si "existe" n'ajoutait rien à "le roi de France". En un mot, "A existe" ajoute quelque chose à "A" est une description et est dépourvu de sens si "A" est un nom propre authentique.

L'importance de ce résultat est très grande, en ce qui concerne notre question présente. Un très grand nombre de propositions peut être présenté sous une forme existentielle : "certains hommes sont sages" peut être présentée sous la forme "les hommes sages existent" ; "le fils de Philippe conquis la Perse" peut être présentée sous la forme "le fils de Philippe qui conquis la Perse a existé", et ainsi de suite. Les propositions qui ne peuvent pas être présentées sous cette forme peuvent être considérées comme niant [*denying*] l'existence : "tous les hommes sont mortels" peuvent être formulées par "les hommes immortels n'existent pas", et ainsi de suite. Tant que l'on pensait que l'"existence" évoquée dans ce genre de propositions est quelque chose qui peut être prédiqué significativement d'un particulier réel donné, il était **(139)** impossible de répondre à l'affirmation de Hume selon laquelle l'existence ajoute quelque chose au sujet, et il était par conséquent impossible de montrer en quoi les propositions diffèrent des concepts, et en quoi la croyance diffère de l'accointance.

Est étroitement liée à cette question de l'existence la question de la relation du sujet grammatical avec l'analyse réelle d'une proposition. Dans "les hommes immortels n'existent pas", il n'y a pas une entité "les hommes immortels" dont la non-existence serait prédiquée. "Les hommes immortels" est simplement un constituant grammatical de l'expression, il ne correspond à aucun constituant réel de la proposition, dont une expression plus correcte est "tout ce qui est humain n'est pas immortel". Et cela s'applique à un grand nombre de sujets

grammaticaux, présentables sous la forme de concepts complexes. Dans ce genre de cas, la complexité est *propositionnelle* en temps normal, c'est-à-dire que les propositions qui ont ces sujets à titre grammatical sont en réalité moléculaires et impliquent une séparation de ce qui apparaît comme les parties d'un seul sujet complexe. Ainsi, il est tout à fait trompeur de supposer que, partout où un sujet grammatical est formé par la combinaison de noms de concepts, il doive y avoir un complexe formé par la combinaison des concepts dont les noms figurent dans le sujet grammatical. Cette supposition trompeuse a étendu jusqu'à l'excès la croyance générale à des concepts complexes parmi les philosophes, et elle a obscurci la partie jouée par les relations reliantes dans la formation des complexes, ainsi que la différence fondamentale entre concept et proposition.

Au cours de l'examen de la croyance, Hume mentionne le fait de comprendre des propositions qui ne sont pas crues. Il dit :

> Supposons une personne présente avec moi qui émet des propositions, auxquelles je ne donne pas mon assentiment, *que* César *est mort dans son lit*, *que l'argent est plus fusible que le plomb*, *ou que le mercure est plus lourd que l'or* ; il est évident que, malgré mon incrédulité, je comprends clairement sa signification, et je forme la totalité des mêmes idées que lui forme. Mon imagination est dotée des mêmes pouvoirs que la sienne ; il n'est pas non plus possible pour lui de concevoir une idée que je ne puisse concevoir ; ni d'en conjoindre aucune que ne je ne puisse conjoindre.

Sur la seule base de ce passage, on pourrait penser que Hume distingue entre conjoindre des idées et concevoir une nouvelle idée complexe, mais ce serait là une erreur d'interprétation due au contexte apparent. Mais ce qui ressort du passage ci-dessus, c'est que Hume conçoit la pensée comme conjoignant les *idées* d'objets, alors que ce qui rend une pensée *vraie* est une conjonction des *objets*. Cela donne bien sûr un moyen bref et simple de définir la fausseté, et de faire la distinction entre les propositions et le fait qui les rend vraies. Pour nous, du fait de notre rejet des "idées" comme entité tierce [*tertium quid*] entre sujet et objet, aucune explication de ce genre n'est possible. Lorsque nous jugeons que le mercure est plus lourd que l'or, *mercure*, *plus lourd* et *or* doivent **(140)** être les constituants de l'événement que représente notre acte de jugement [*judging*] : nous ne pouvons pas dire que nous mettons notre idée de mercure dans une certaine relation avec notre idée d'or, mais que le mercure et l'or ne se trouvent pas eux-mêmes dans la relation "correspondante". Il est curieux que les auteurs qui ne croyaient pas aux idées n'aient pas été troublées concernant cette correspondance de relations. La relation entre mon idée de mercure et mon idée d'or ne peut pas être "plus lourde", puisque mes idée ne sont pas censées avoir de poids. L'*idée* de "plus lourd" non plus, puisque ce n'est pas une relation. Ce doit être par conséquent une sorte de nouvelle relation, en un sens reliée à "plus lourd" et subsistant entre mes idées, mais pas nécessairement présente à ma conscience lorsque je juge. Mais cela est évidemment absurde. Mon acte de jugement consiste à l'évidence dans mon acte de croire qu'il y a une relation entre les objets réels, *mercure* et *or*, non dans le fait qu'il y ait une relation effective entre mes idées de ces deux objets. Ainsi, la nature d'ensemble de la croyance doit être nécessairement mal comprise par ceux qui supposent qu'elle consiste dans une relation entre des "idées"

plutôt que dans la croyance d'une relation entre des objets. *Quelque chose* de subjectif doit être inséparablement lié à la croyance et de manière telle qu'il empêche de la considérer comme une relation duale avec un objet particulier puisque, si nous la considérions comme telle, la fausseté deviendrait inexplicable. Mais les partisans des "idées" introduisent l'élément subjectif au mauvais endroit. Il intervient avec le sujet lui-même, et il est plus important que dans le cas de l'accointance parce que, du fait que la croyance est une relation multiple, elle n'implique pas *un* seul objet qui corresponde à la croyance et ne comporte pas le sujet. Mais les divers objets qui sont les *constituants* de la croyance ne tout aussi dépourvus de référence au sujet que l'objet unique dans le cas de l'accointance –excepté bien sûr lorsqu'il s'avère que la croyance porte sur le sujet.

Ce que William James dit à propos de la croyance dans sa *Psychologie* est à peu près la même chose que ce qui est dit par Hume, de notre point de vue. Au début de son examen,[2] il dit :

> Tout le monde connaît la différence entre imaginer une chose et croire à son existence, entre supposer une proposition et acquiescer à sa vérité. Dans le cas de l'acquiescement ou de la croyance, l'objet n'est pas seulement appréhendé par l'esprit, il est considéré comme ayant une réalité. La croyance est donc l'état mental ou fonction d'avoir connaissance de [*cognizing*] la réalité ... *Dans sa nature interne, la croyance, ou sens de la réalité, est une sorte de sentiment plus liée aux émotions que n'importe quelle autre.*

Ce qui a déjà été dit concernant les propositions existentielles s'applique tel quel à la première partie de la citation ci-dessus. D'après notre théorie, "avoir connaissance de la réalité" décrirait plus justement l'accointance que la croyance. La croyance fausse est-elle une prise de connaissance d'une irréalité ? N'y a-t-il pas des croyances qui **(141)** ne sont en aucune façon existentielles ? La conception d'ensemble de l'"existence" n'est-elle pas inapplicable aux objets dont nous sommes précisément le plus certains, à savoir ceux donnés dans l'accointance présente ? Et les objets imaginés ne sont-ils pas aussi "quelque chose", avec une "réalité" à ces objets d'un type approprié ? Je ne m'étendrai pas une nouvelle fois sur ces questions de nouveau, et passerai immédiatement à l'*émotion* de la croyance.

En ce qui concerne l'"émotion" de la croyance, la seule chose qu'il faille comprendre ici c'est que, aussi réelle et importante soit-elle en qualité de fait psychique, elle ne concerne pas l'épistémologie et ne doit être signalée que pour éviter les confusions qui pourraient résulter de son intrusion inopinée, tout comme un étranger indésirable dont la photographie serait transmise aux autorités à la frontière. Il y a une émotion de conviction, susceptible d'un grand nombre de degrés et survenant avec les jugements qui maintiennent notre assentiment inébranlable, ou avec les perceptions qui mettent fin à un doute. Mais cette émotion, bien qu'elle accompagne souvent les jugements, ne les constitue en aucune façon ; à vrai dire, elle peut être exactement la même dans le cas de deux jugements différents. Une personne portée sur le patriotisme ressentira exactement la même émotion de conviction en concevant la croyance que son pays est le meilleur au monde qu'en concevant la croyance que son école ou son club est le meilleur du pays. L'émotion n'est pas une relation aux objets de la croyance

[2] Chapitre XXI, "The Perception of Reality", Vol. II, p. 283.

mais un fait mental vif, causé peut-être par la croyance mais totalement distincte de celle-ci. Et il semblerait que son intensité ne soit pas réellement proportionnelle à notre certitude mais à l'énergie avec laquelle nous repoussons le doute. Personne ne ressent grand-chose de cette émotion lorsqu'il s'agit d'observer les faits d'une table de multiplication, parce que le doute à leur encontre n'est pas conçu comme étant possible. Mais les croyances religieuses et politiques, précisément parce qu'elles soient déniées [*denied*] ou mises en doute par un grand nombre, éveillent une conviction de la plus grande ferveur. Mais de telles croyances ne sont pas celles qu'un philosophe devrait prendre pour modèle.

L'incapacité à faire une distinction suffisamment radicale entre croyance et représentation [*presentation*] a provoqué des difficultés dans la théorie de la croyance erronée [*erroneous*], et elle a conduit à des tentatives d'introduire le dualisme du vrai et du faux parmi les objets. D'aucuns supposent qu'il y a des objets réels et des objets irréels, ou des objets vrais et des objets fictifs. Ces distinctions sont toujours dues à l'intrusion de la croyance sous diverses apparences ou, tout au moins, à la compréhension des propositions. Une des apparences les plus trompeuses est le sujet grammatical : la montagne d'or , ou le carré rond, sont conçues comme des objets parce qu'ils peuvent être des sujets grammaticaux ; ils sont déclarés par conséquent comme étant des objets "irréels". Tout cela repose sur une mauvaise analyse des propositions, comme on l'a déjà expliqué. Le dualisme du vrai et du faux, ainsi que la totalité de ses distinctions afférentes, présuppose des propositions et ne survient pas tant que nous nous limitons à l'accointance, excepté peut-être dans le cas des formes logiques abstraites ; et même alors, il n'y a pas de dualisme proprement dit, puisque la fausseté est logiquement impossible dans ces cas de figure. Mais à l'inverse, lorsque nous traitons de la croyance nous devons toujours porter à l'esprit le dualisme du vrai et du faux, et éviter les théories basées **(142)** sur l'idée excessivement enthousiaste selon laquelle toutes les croyances sont vraies. Ce test très simple condamne un grand nombre de théories de la croyance.

Comme nous l'avons dit plus tôt, l'analyse de la croyance doit être précisément analogue à l'analyse de la compréhension d'une proposition. Les mêmes constituants entrent dans un complexe de même forme, lorsqu'une proposition est crue et lorsqu'elle est comprise. La seule différence est que l'acte de croire [*believing*] est une relation, et l'acte de comprendre[*understanding*] en est une autre. Il apparaît immédiatement que les mêmes arguments s'appliquent précisément aux deux cas, et que toute modification qui s'avérerait nécessaire dans la théorie de l'une serait tout aussi nécessaire dans la théorie de l'autre. Le fait que la différence réside seulement dans les relations est rendu évident par le critère que nous avons appliqué pour distinguer la sensation et la mémoire, par exemple : l'acte de comprendre une proposition donnée est différent de l'acte de la croire, mais il n'y a aucune différence toutefois en ce qui concerne les objets concernés dans les deux cas.

Nous devons considérer maintenant une nouvelle relation, à savoir l'*incroyance*. La question principale ici est celle de savoir si l'incroyance est une relation inanalysable, ou si elle constitue vraiment la croyance à la contradictoire. Lorsque j'incrois[1] que A et B sont similaires, est-ce que je crois que A et B ne sont pas similaires, ou suis-je dans une nouvelle relation, de même forme que les croyances et avec les mêmes objets que lorsque je crois que A et B sont similaires ? Il est difficile de trouver des arguments sur ce point, et il semblerait que nous devions compter sur une inspection de l'esprit [*inspection*].[2] Sur la base d'une

inspection de l'esprit, je soutiens, quoique sans grande certitude, que l'incroyance est une nouvelle relation inanalysable impliquant simplement le rejet de la proposition incrue, et ne consistant pas dans l'acceptation de sa contradictoire.

Cette conclusion peut être quelque peu renforcée si l'on considère le *doute*. Il est assez évident, je pense, que nous pouvons avoir un grand nombre de degrés d'incertitude concernant une proposition. Nous pouvons être presque sûrs qu'elle est vraie, balancés entre sa vérité et sa fausseté, être presque sûrs qu'elle est fausse ; et un nombre indéterminé d'autres sentiments semble possible entre ces attitudes. Si tous ces sentiments sont des relations indéfinissables à une proposition, ce que je crois, il semblerait normal de situer l'incroyance totale en bas de l'échelle, pour faire contrepoids avec la croyance au sommet. Il serait absurde de prétendre qu'un tel argument est décisif, mais lorsqu'il s'accorde avec l'inspection directe on peut admettre qu'il compte pour quelque chose.

Le terme *doute* n'est peut-être pas le meilleur terme pour décrire les attitudes intermédiaires entre la croyance complète et l'incroyance complète. "Doute" suggère une hésitation, une croyance et incroyance alternée qui n'est pas l'attitude que j'entends caractériser. J'entends une attitude qui peut être parfaitement stable, dans laquelle figure un élément d'incertitude –ce que nous devrions décrire en disant "Je le pense, mais je ne me sens pas sûr". Il y a bien entendu des degrés d'incroyance ainsi que des degrés de croyance ; nous pouvons les combiner dans une échelle de "certitudes", où la pleine croyance et la pleine incroyance se trouvent tous deux au sommet, tandis que la suspension de jugement complète se trouve tout en bas. Mais même la suspension de jugement complète appartient à la croyance et l'incroyance plutôt **(143)** qu'à la simple compréhension d'une proposition. C'est une attitude déterminée par rapport à la croyance et l'incroyance, et elle représente le résultat d'une tentative de décider entre les deux. L'incertitude complète est manifestement un état d'esprit différent de celle d'une personne à laquelle l'attitude de répondre par oui ou par non ne s'est pas imposée d'elle-même, de même que l'incapacité à répondre à une question qui a été posée est différente du fait de ne pas considérer la question comme une question qui doive trouver une réponse si possible.

La croyance peut donc avoir des degrés de certitude, de zéro à la conviction absolue ; et si l'incroyance diffère de la croyance à la contradictoire, comme ce semble être le cas, alors l'incroyance comporte également des degrés de certitude correspondant exactement à celles de la croyance. Il faut dire que les relations qui diffèrent par degrés sont des relations différentes ; par conséquent, chaque degré de croyance est une relation nouvelle différant de chaque autre degré ; et si l'incroyance est indépendante, alors il y a également un grand nombre de relations d'incroyance correspondant aux différents degrés. Toutes ces relations donnent lieu à des complexes de même forme et, par conséquent, aucun problème d'analyse logique nouveau ne ressort de leur multiplicité.

NOTES DU TRADUCTEUR

[1] Néologisme de circonstance, 'to disbelieve' n'a pas de contrepartie française basée sur la racine 'croire' ; 'rejeter', 'dénier' ou 'refuser' seraient à peu près synonymes, mais l'analyse

exige une distinction prudente entre ces termes et nous maintiendrons pour cette raison le néologisme ci-dessus. [Vernant 2003] a justifié son choix en raison de sa connotation religieuse, même si cette raison nous a semblé tout sauf judicieuse en section 3.3 (voir note 128). L'explication de Russell qui suit le confirme.

[2] Nous précisons 'inspection de l'*esprit*', pour insister sur l'aspect psychologique ou introspectif de l'argument ci-dessus de Russell : l'inspection n'indique pas ici un raisonnement abstrait mais une réflexion proprement dite sur nos usages *intuitifs* de la notion d'incroyance. La notion d'introspection aurait pu être utilisée pour traduire 'inspection', mais l'équivalent anglais existe et les deux ne sont peut-être pas interchangeables en tout point.

International Logic Review **21**(juin 1980), pp. 31-45

ALLAN M. HART*

POUR UNE LOGIQUE DU DOUTE**

L'existence de théories sémantiques bien développées de la logique modale a conduit certains philosophes, notamment Hintikka, à la conjecture selon laquelle de telles structures pourraient être appliquées à l'étude de modalités non-aléthiques, sous réserve des modifications appropriées. En particulier, il est suggéré que la logique de certaines attitudes propositionnelles tout au moins peuvent être étudiées ainsi [1]. Comme on le sait, le *Knowledge and Belief* de Hintikka est une tentative de satisfaire partiellement cette suggestion [2].

Dans ce qui suit, j'aimerais tenter une formalisation de la logique du doute. La stratégie en vue de cette tentative, ainsi que les techniques utilisées dans sa réalisation, prolongeront celles employées par Hintikka dans [2]. En particulier, la procédure des systèmes modèles sera employée ici.[1] De plus, mon analyse de la relation entre la logique du doute et les logiques de la connaissance et de la croyance procèdera par le biais d'une acceptation explicite des systèmes épistémique et doxastique de Hintikka. En revanche, je ne chercherai pas à étendre mon analyse pour y inclure l'appareil quantificationnel. La raison de restreindre mes efforts de la sorte est avant tout économique. L'inclusion d'une analyse quantificationnelle du doute rendrait cet article excessivement long et ce sans contribuer au but de clarification de la notion de doute, je le crains. Qui plus est, il semble peu probable qu'une extension de ce genre engendre des problèmes nouveaux qui ne soient pas déjà parus lors de tentatives de mélanger les quantificateurs à d'autres attitudes propositionnelles. Ces problèmes sont évidemment formidables, pour sûr.

Du point de vue étymologique, le 'doute' trouve son origine dans le latin 'dubitare', lequel est lui-même un fréquentatif[2] de 'dubare' –être entre deux avis. 'Dubare' dérive probablement de l'adjectif 'dubus', construit à partir de 'duo' –deux [3]. In *Modern English Usage*, H.W. Fowler suggère que l'idiome privilégié est ['*doubts whether*']et non ['*doubts that*'] [4]. En outre, le *Oxford English Dictionary* nous informe du fait que 'whether' sert le plus **(32)** souvent à introduire des alternatives mutuellement exclusives [5].[3] Cette ascendance linguistique du 'doute' est plus ou moins parallèle à l'interprétation philosophique traditionnelle du doute comme suspension du jugement ou de la croyance, ou comme un refus

---

* Professeur au Département de Philosophie, Université du Michigan Central, Mount Pleasant, Etats-Unis.

** Une version antérieure de cet article fut déposée lors des conférences de l'Académie des Arts, Sciences et Lettres du Michigan en mars 1978. Je suis redevable à un grand nombre de suggestions précieuses lors de la préparation de cet article. J'aimerais mentionner en particulier le Dr. Charles Bruce ainsi que le Professeur Herbert E. Hendry. Ce fut au cours de discussions avec le premier que la faisabilité de (D2) émergea. Je suis particulièrement redevable au dernier pour des suggestions, critiques et secours dans pratiquement chaque partie de cet article. Aucun des deux n'est responsables de quelque erreur que ce soit, bien entendu ; cette responsabilité est entièrement la mienne.

[*withholding*] d'assentiment. Dans son *Three Dialogues Between Hylas and Philonous*, Berkeley suggère par exemple la caractérisation suivante du doute :

Philonous Et si le fait de douter consistait à acquiescer à la partie affirmative ou à la partie négative d'une question ?
Hylas Ni l'une ni l'autre ; car quiconque comprend l'anglais ne peut pas ignorer que douter signifie une mise en suspens entre les deux. [7]

Conformément à la suggestion de Berkeley, nous pourrions procéder en définissant le 'doute' de la manière suivante : '*a* doute que p'[4] s'explique par '*a* ne croit pas que p et *a* ne croit pas que non-p' (*a* peut se comprendre comme le nom d'une personne, un pronom personnel, voire peut-être une description définie désignant un être humain, et p est une complétive [*clause*] indépendante.[5] La représentation formelle de cette définition au sein du système doxastique de Hintikka s'accomplit maintenant en deux étapes : (i) On augmente le vocabulaire de ce système en introduisant l'opérateur qui suit : 'D'. 'D_a' a pour but d'être la contrepartie formelle de l'expression 'a doute que'. (ii) Une fois le vocabulaire augmenté, les règle de formation sont complétées de la manière habituelle. En conséquence, notre définition devient celle-ci :

(D1) $D_a p =_{df.} (\neg B_a p \wedge \neg B_a \neg p)$[1 6]

Le caractère inacceptable de (D1) devient évident sur la base des réflexions qui suivent. Alors que

(1) $D_a p \rightarrow (\neg B_a p \wedge \neg B_a \neg p)$

semblerait incontestable, sa converse

(2) $(\neg B_a p \wedge \neg B_a \neg p) \rightarrow D_a p$

semblerait sujette à des contre-exemples. Descartes ne croyait certainement pas que les électrons ont des charges négatives, et il ne croyait certainement pas non plus qu'ils en ont ; mais la chose la plus certaine, c'est qu'il n'a pas douté qu'ils en eussent. Le fait de supprimer la condition que le référent de '*a*' soit un être humain permet des contre-exemples à (D1) encore plus éloquents. Il va de soi que **(33)** le Taj Mahal pourrait satisfaire l'antécédent de (2), mais de même qu'il ne parvient évidemment pas à satisfaire son conséquent. La cause de nos difficultés devient évidente dès que nous passons en revue les remarques de Berkeley. Berkeley définissait le 'doute' (en gros) comme une *suspension* de la croyance. Ce que la partie droite de (D1) représente n'est pas tant une *suspension* de croyance qu'une simple *déficience* ou *absence* de croyance. A l'évidence, ce qu'on nous demande est un moyen

[1] Pour une analyse similaire de la notion 'refuser l'assentiment', voir "A system of epistemic logic", de Roderick Chisholm et Herbert Keim, *Ratio* **14**(2) (1972), pp. 99-115, en part. 102.

d'indiquer que l'absence de croyance chez *a* est une chose à laquelle *a* consent lui-même/elle-même et n'est pas simplement le résultat de l'ignorance de *a*. Et avec peut-être moins de clarté, la définition du 'doute' de Berkeley peut donc être représentée adéquatement par :

(D2) $D_a p =_{df.} B_a(\neg B_a p \wedge \neg B_a \neg p)$.[7]

La lourdeur de cette définition peut être réduite quelque peu lorsque nous considérons une transcription de (D2) quelque peu différente (mais équivalente), à savoir

(3) $D_a p =_{df.} B_a(C_a p \wedge C_a \neg p)$.[8]

Or on peut montrer facilement que (3) donne

(4) $D_a p \equiv (B_a C_a p \wedge B_a C_a \neg p)$

La partie droite de (4) peut se lire '*a* croit que p est compatible avec tout ce que *a* croit et *a* croit que non-p est compatible avec tout ce que *a* croit'.[9] Or ce que cet énoncé décrit, me semble-t-il, c'est une situation dans laquelle *a* croit qu'il n'y a rien dans le stock de ses croyances qui lui permette de choisir entre p et non-p. Une situation de ce genre est clairement, semblerait-il, une situation dans laquelle il serait correct d'attribuer le doute à *a* ; c'est, je suppose, une situation dans laquelle *a* a *suspendu* ses croyances en ce qui concerne p.

Pour autant que je puisse dire, (D2) est parfaitement adéquat en vue de la tâche consistant à saisir nos intuitions pré-analytiques concernant au moins un sens de 'doute'. Comme pour toute définition, toutefois, son caractère adéquat ne devrait pas être jugé dans les seuls termes de succès ou d'échec de ce genre mais également en termes d'acceptabilité des conséquences que son acceptation apporte. Comme nous le verrons, ces conséquences ne s'opposent pas à nos intuitions, et la plupart des temps, toutes peut-être leur sont conformes.

Avant de procéder à une exploration des conséquences de (D2), une remarque de précaution s'impose. La tentative a consisté jusqu'ici à apporter une reconstruction acceptable d'un sens précis du 'doute'. Le sens du 'doute' en question est un sens qui a traditionnellement plu aux philosophes lors de l'examen du scepticisme, etc. et qui ne doit pas être confondu avec le sens du 'doute' représenté par le scénario suivant : supposons que quelqu'un déclare que le soleil ne se lèvera pas demain. Supposons par ailleurs que moi, croyant que la déclaration de l'individu est absurde, réponde 'je doute que ce soit vrai'. Ce scénario illustre, je pense, un sens assez commun du 'doute' qui n'est pas reflété par (D2). La meilleure façon peut-être de distinguer cet usage du 'doute' **(34)** de celui représenté par (D2) serait de le munir de l'expression ['*doubts that*'], réservant de fait ['*doubts whether*'] à (D2).[10] La distinction est cruciale, et il faut la garder à l'esprit si l'on veut que les résultats obtenus par la suite ne soient pas mal interprétés.[11] Comme le scénario le laisse penser, 'doute que' trouve peut-être sa meilleure expression en termes d''incroyance' [*disbelief*] plutôt qu'en termes de *suspension* de la croyance. L'incapacité à distinguer ces deux sens du 'doute' peut provoquer et a provoqué une bonne part de confusion. Dans son examen du principe cartésien

du 'doute méthodologique', le jésuite Bourdin interprète par exemple l'usage par Descartes du terme 'doute' comme un cas d'incroyance concrète. Il écrit :

> Si [c'est le cas, et si] la déclaration que 2 et 3 font 5 est douteuse, je croirai, et ce conformément à ce que votre principe asserte, qu'elle est fausse. En conséquence de quoi j'affirmerai l'opposé et asserterai ceci : que 2 et 3 font bien 5. [8]

La réponse assez froissée de Descartes à cette objection, alors qu'elle se montre injuste envers Bourdin en raison même de l'imprudence de Descartes dans son usage des termes 'doute', 'douteux', etc. dans les *Méditations*, indique que le sens du 'doute' qu'il avait à l'esprit n'était pas celui de l'incroyance concrète, mais celui caractérisé par la 'suspension de jugement' ou le 'refus de l'assentiment'.[2] [12]

Etant donné l'acceptation de (D2), nous sommes (espérons-le) en situation de fournir un traitement plus systématique et moins informel de '$D_a p$'. Les conditions C (i.e. les critères de consistance) de '$D_a p$' peuvent être obtenues assez facilement à partir de (D2) et des conditions C de '$B_a p$'. Elles peuvent être inventoriées comme suit :

(C.d*) Si '$D_a p$' $\in \mu$ et si μ appartient à un système modèle Ω, alors il y a dans Ω au moins une alternative doxastique μ^* à μ (par rapport à *a*) telle que $p \in \mu^*$.[13]

(C.DD*) Si '$D_a p$' $\in \mu$ et si μ^* est une alternative doxastique à μ (par rapport à *a*) dans un système modèle, alors '$D_a p$' $\in \mu^*$.[14]

(C.D¬) Si '$D_a p$' $\in \mu$, alors '$D_a \neg p$' $\in \mu$.[15]

(C.DB) Si '$D_a p$' $\in \mu$, alors '$B_a(\neg B_a p \wedge \neg B_a \neg p)$' $\in \mu$.[16]

(C.¬D*) Si '$\neg D_a p$' $\in \mu$ et si μ appartient à un système modèle Ω, alors il y a dans Ω au moins une alternative doxastique μ^* à μ (par rapport à *a*) telle que soit p est un membre de chaque alternative à μ^* (par rapport à *a*) soit '$\neg p$' est un membre de chaque alternative à μ^* (par rapport à *a*).[17]

(C.d*)–(C.¬D*) suffisent pour ce que nous voulons faire, tant que (35) l'appareil de la quantification n'est pas inclus dans la portée de notre discussion. L'ajout de toute autre condition serait soit non désirée soit injustifiée, me semble-t-il, au sens où les conditions en question s'avéreraient soit être dérivables d'une des conditions ci-dessus ou plus, soit donner des conséquences inacceptables. En fait, on peut s'apercevoir facilement que (C.d*), (C.DD*), (C.D¬) et (C.¬D*) sont dérivables des conditions C de 'B' augmentées de (C.DB) et, en ce sens, elles sont éliminables. Lors de l'inventaire des conditions C ci-dessus, nous devons évidemment garder à l'esprit les divers systèmes que nous examinons. Dans un système où 'D' apparaît comme seul opérateur non-vérifonctionnel primitif, (C.DB) n'apparaîtrait pas, mais (C.d*), (C.DD*), (C.D¬) et (C.¬D*) apparaîtraient tous. Il faut donc considérer les résultats qui vont être explorés maintenant (ainsi que ceux prononcés précédemment) comme valables au sein d'un système composé de l'union des systèmes épistémique et doxastique de Hintikka, augmentée du vocabulaire et des règles de formation appropriées ainsi que de (C.DB), (C.d*), (C.DD*), (C.D¬) et (C.¬D*).

---

[2] Voir [8], pp. 265-267.

L'acceptabilité de nos conditions C, et donc celle de (D2), ne peuvent être pleinement testées qu'en examinant les conséquences que leur acceptation apporte. En particulier, il serait souhaitable que toutes les conditions placées sur 'D' soient telles qu'elles donnent l'indéfendabilité doxastique[18] de

(5) $p \wedge D_a p$

On peut, je pense, facilement concevoir (5) comme une traduction de la variante du 'paradoxe de Moore' qui suit :[3]

(6) p, mais je doute que p.

Le paradoxe de Moore, faut-il le rappeler, consiste dans l'étrangeté du fait de prononcer un énoncé tel que

(7) p, mais je ne crois pas que p.

On rappellera également que la solution de Hintikka au paradoxe de Moore consiste à démontrer l'indéfendabilité doxastique de (7). Mais nul doute que si le fait de prononcer (7) est en un certain sens étrange, il en va de même pour une énonciation de (6). En conséquence, une démonstration de l'indéfendabilité doxastique de (5) semblerait imminente. C'est :

(1) $B_a(p \wedge D_a p) \in \mu$	Hypothèse	
(2) $(p \wedge D_a p) \in \mu^*$	de (1) par (C.b*)	
(3) $D_a p \in \mu^*$	de (2) par (C.$\wedge$)	**(36)**
(4) $D_a \neg p \in \mu^*$	de (3) par (C.D$\neg$)	
(5) $\neg p \in \mu^{**}$	de (4) par (C.d*)	
(6) $B_a(p \wedge D_a p) \in \mu^{**}$	de (1) par (C.BB*)	
(7) $p \wedge D_a p \in \mu^{**}$	de (6) par (C.B*)	
(8) $p \in \mu^{**}$	de (7) par (C.$\wedge$)	

(5) et (8) violent (C.$\neg$), et la preuve est complète.

Une caractéristique unique et souvent critiquée des systèmes doxastique et épistémique de Hintikka est la suivante : chaque fois que l'on peut montrer que p implique virtuellement q, on peut montrer que

(8) $K_a p \rightarrow K_a q$

et

[3] Le paradoxe de Moore n'est pas un paradoxe du tout, selon certaines interprétations. Une meilleure description est peut-être celle fournie par Hintikka, à savoir "Moore's problem of saying and disbelieving" ; voir p. 64 de [2].[19]

(9) $B_a p \rightarrow B_a q$

sont toutes deux autosuffisantes [*self-sustaining*]. Il convient de noter à ce sujet que le comportement de 'D' est assez différent de celui de 'K' ou 'B'.

(10) $D_a p \rightarrow D_a q$

n'est pas auto-suffisante, en général, quand bien même p impliquerait q virtuellement.[20] Loin d'être contre-intuitif, toutefois, ce résultat est conforme à nos intuitions, me semble-t-il. Une personne qui doute que p est le cas n'est tenue en temps normal par aucune contrainte de rationalité à douter des conséquences de p, quand bien même ces conséquences seraient 'évidentes'. En effet, puisque chaque fois que q est autosuffisante 'p → q' sera autosuffisante également, il serait irrationnel pour la personne en question de douter de quelques-unes, pour le moins, des conséquences de ce dont elle doute effectivement.[21] Il ne faudrait pas supposer non plus qu'une personne doive nécessairement douter de ce que ce dont elle doute est une conséquence. Une fois encore, puisque chaque fois que q est indéfendable 'q → p' est autosuffisante, il serait irrationnel pour la personne en question de douter de quelques-unes, pour le moins, des choses de ce que ce dont elle doute est une conséquence.[4] [22] La tentation de supposer qu'une personne doive douter des choses de ce que dont elle doute est une conséquence provient, selon moi, d'une incapacité à distinguer le doute de l'incroyance. En effet, on peut facilement montrer au sein du système doxastique de Hintikka que chaque fois que 'p → q' est autosuffisante et que *a* rejette [*disbelieves*] q, *a* rejette p également.[23] La tentation de supposer qu'une personne doive douter de toute déclaration indéfendable provient, selon moi, de la même incapacité. L'incapacité en question est peut-être (**37**) une incapacité à apprécier la distinction faite plus tôt entre 'doute que p est le cas' [*doubts whether*] et 'doute que p soit le cas' [*doubts that*].[24]

D'autres différences entre le comportement de 'D' et celui de 'K' et 'B' surgissent dès que nous examinons la question de la distribution et de la factorisation [*collection*]. Alors que 'K' et 'B' se distribuent tous les deux sur '∧', '→' et '≡' et se factorisent sur '∧', '∨' et '≡', 'D' ne se distribue et ne se factorise sur aucun de ces opérateurs. Des similarités entre le comportement de 'D' et celui de 'K' et 'B' surviennent lorsque apparaît la question de savoir dans quel sens, si tant est qu'il y en ait un, on peut dire du doute qu'il est 'auto-suggéré' [*self-intimating*]. Dans [2], Hintikka prétend que '$K_a p$' est virtuellement équivalente à '$K_a K_a p$', et que '$B_a p$' implique virtuellement '$B_a B_a p$'.[25] Les deux résultats sont démontrables logiquement, dit Hintikka, et ne dépendent pas de l'interprétation de la connaissance ou de la croyance comme états mentaux.[26] A ce sujet, l'analogue du doute n'est évidemment pas la conjecture que '$D_a p$' impliquerait virtuellement ou serait virtuellement équivalent à '$D_a D_a p$', mais plutôt la conjecture que '$D_a p$' impliquerait virtuellement ou serait virtuellement équivalent soit à '$K_a D_a p$', soit à '$B_a D_a p$'. Ainsi qu'on peut facilement le démontrer, la bonne conjecture est que '$D_a p$' implique virtuellement '$B_a D_a p$'. La tentation de supposer que '$D_a p$'

[4] Comme on pourrait le supposer, je définis la notion de conséquence de la manière suivante : soit p une formule et ϕ un ensemble de formules. Alors p est une conséquence de ϕ ssi $\phi \cup \{\text{'}\neg p\text{'}\}$ est indéfendable.

implique virtuellement 'K_aD_ap' provient selon moi de la même origine que la tentation de supposer que '$B_ap \rightarrow K_aB_ap$' est autosuffisante. L'origine en question est donc l'argument de l'introspection, déclare Hintikka. En gros, l'argument de l'introspection affirme qu'une personne peut connaître et connaît ses propres états mentaux, où 'état mental' est compris de manière à inclure tout ce que nous incluons en temps normal sous l'appellation des 'attitudes propositionnelles'. Une critique de cette ligne d'argumentation a été assez vigoureuse, bien sûr, et Hintikka a produit en particulier sa propre critique. D'après Hintikka, une acceptation de l'argument de l'introspection obligerait à l'acceptation de (C.BB*ep), laquelle garantirait à son tour l'autosuffisance de '$B_ap \rightarrow K_aB_ap$'. Mais (C.BB*ep) ne doit pas être acceptée, déclare Hintikka, puisque son acceptation aurait pour effet de rendre indéfendable le fait de supposer que toute personne abandonnerait ses croyances actuelles s'il avait plus d'information qu'il n'en a maintenant.[5] [27] Pour des raisons exactement analogues, toute condition qui garantirait l'autosuffisance de '$D_ap \rightarrow K_aD_ap$' devrait être rejetée également. Il ne semblerait pas y avoir de raison de supposer qu'une personne ne puisse jamais résoudre ses doutes.[28]

Jusqu'à présent, je n'ai pas discuté les problèmes intéressants qui surviennent lorsque 'D_a' est itérée avec elle-même. En particulier, la question de savoir si le doute peut être mis en doute ou non apparaît.[6] Certains philosophes ont déclaré que le doute **(38)** ne peut pas être mis en doute, soutenant habituellement leur affirmation en plaidant soit pour l'inconsistance (indéfendabilité) de 'D_aD_ap', soit pour son caractère dépourvu de sens.[29] La plupart des arguments en faveur de la première solution sembleraient dépendre d'un recours à l'argument de l'introspection (i.e. d'un recours à l'autosuffisance de '$D_ap \rightarrow K_aD_ap$'), affirmation pour laquelle nous avons déjà trouvé des raisons de rejeter. Les arguments en faveur de la seconde solution sembleraient puiser leur force d'arguments plus généraux disant qu'aucun terme modal, ou aucun des termes d'attitudes propositionnelles ne peuvent être itérés de manière pourvue de sens. L'étrangeté d'énoncés contenant des occurrences itérées de 'doute que' [*doubts whether*], si tant est qu'ils soient curieux, ne trouve pas sa source selon moi dans l'indéfendabilité de 'D_aD_ap' ; elle se trouve plutôt dans la relation entre 'D_aD_ap' et 'D_ap', ainsi que je le suggérerai brièvement. De plus, il me semble qu'il n'y a pas plus de raison de supposer que '*a* doute que *a* doute que p' est insensé qu'il n'y en a de supposer que '*a* sait que *a* sait que p' ou '*a* croit que *a* croit que p' le sont.

Comme je l'ai indiqué précédemment, l'étrangeté présumée de 'D_aD_ap' provient de sa relation à 'D_ap' elle-même. La relation en question est simplement celle de l'indéfendabilité de '$D_ap \wedge D_aD_ap$' ou, ce qui est peut-être encore mieux, l'autosuffisance de

(11) $D_ap \rightarrow \neg D_aD_ap$

On peut prouver l'autosuffisance de (11) assez facilement, de la manière suivante :

(1) $D_ap \in \mu$ Hypothèse

[5] Voir l'examen par Hintikka de (C.BB*ep) dans [2] pp. 49-53. Voir également son examen de l'argument de l'introspection dans [2] pp. 53-56.
[6] Pour un traitement antérieur de cette question, voir "Can doubt be doubted ?" de Sibajiban, *Mind* **69**(1960), pp. 84-87.

(2) $\neg\neg D_a D_a p \in \mu$	Hypothèse
(3) $D_a D_a p \in \mu$	de (2) par (C.¬¬)
(4) $D_a{\sim}D_a p \in \mu$	de (3) par (C.D¬)
(5) $\neg D_a p \in \mu^*$	de (4) par (C.d*)
(6) $D_a p \in \mu^*$	de (1) par (C.DD*)[30]

(5) et (6) violent (C.¬), et la preuve est complète.

La démonstration de l'autosuffisance de (11) constitue, je crois, un résultat assez remarquable (une généralisation de ce résultat est fournie dans l'appendice B). Il semblerait en particulier que, sur la base de l'autosuffisance de (11), un argument assez puissant pourrait être présenté qui démontre l'indéfendabilité du scepticisme, où le scepticisme est considéré comme la recommandation que toute affirmation non-logique doit mise en doute. En effet, ce que (11) asserte est simplement ceci que chaque fois qu'une personne *a* doute qu'une déclaration p donnée soit vraie, il y a une autre déclaration, à savoir '*a* doute que p', dont *a* ne peut pas douter. Des formes de scepticisme moins drastiques sont à l'abri de ce type d'attaque directe, bien sûr.[31]

Quelques remarques finales méritent l'attention. Etant donnée l'acceptation de (D2), la question de savoir si 'B' peut être définie **(39)** en termes de 'D' (plus des connecteurs vérifonctionnels, bien sûr) survient naturellement. Il s'avère que non (voir la preuve dans l'appendice C). Ce résultat, prévisible peut-être, n'est pas sans un certain intérêt intuitif. En particulier, il semblerait qu'il apporte son soutien à l'affirmation disant que, en un certain sens, la croyance est préalable au doute d'un point de vue conceptuel ; ou, en guise d'échange, que le doute 'présuppose' la croyance. Les résultats portant sur le nombre de modalités distinctes exprimables au sein d'un système de logique modale donné ont constitué un souci des logiciens modaux. Dans l'appendice D, une preuve est fournie concernant le nombre de 'modalités D' distinctes. **(40)**

Appendice A :

Système KBD :

I. Vocabulaire :
   1. Un stock infini de variables propositionnelles.
   2. Un stock infini de symboles d'individus.
   3. Les connecteurs propositionnels : '¬' et '→'.
   4. Les connecteurs épistémiques/doxastiques : 'K', 'B' et 'D'.
   5. Les indicateurs de regroupement.[32]

II. Règles de formation :
   (i) Les variables propositionnelles sont des formules.
   (ii) Si p est une formule, alors '$\neg p$', '$K_a p$', '$B_a p$' et '$D_a p$' sont des formules.
   (iii) Si p et q sont des formules, alors '$(p \rightarrow q)$' est une formule.

III. Conditions C

(C.¬¬) Si '¬¬p' ∈ μ, alors p ∈ μ* [33]

(C.→) Si '(p → q)' ∈ μ, alors soit '¬p' ∈ μ soit q ∈ μ.

(C.¬→) Si '¬(p → q)' ∈ μ, alors p ∈ μ et '¬q' ∈ μ.

(C.¬) Si p ∈ μ, alors '¬p' ∉ μ

(C.K) Si 'K_ap' ∈ μ, alors p ∈ μ.

(C.KK*) Si 'K_ap' ∈ μ et si μ* est une alternative épistémique à μ (par rapport à *a*) dans un système modèle, alors 'K_ap' ∈ μ*.

(C.¬K*) Si '¬K_ap' ∈ μ et si μ appartient à un système modèle Ω, alors il y a dans Ω au moins une alternative épistémique μ* à μ (par rapport à *a*) telle que '~p' ∈ μ*.

(C.BB*) Si 'B_ap' ∈ μ et si μ* est une alternative doxastique à μ (par rapport à *a*) dans un système modèle, alors 'B_ap' ∈ μ*.

(C.¬B*) Si 'B_ap' ∈ μ et si μ appartient à un système modèle Ω, alors il y a dans Ω au moins une alternative doxastique μ* à μ (par rapport à *a*) telle que '¬p' ∈ μ*.

(C.b*) Si 'B_ap' ∈ μ et si μ appartient à un système modèle Ω, alors il y a dans Ω au moins une alternative doxastique μ* à μ (par rapport à *a*) telle que p ∈ μ*.

(C.B*) Si 'B_ap' ∈ μ et si μ* est une alternative doxastique à μ (par rapport à *a*) dans un système modèle, alors 'p' ∈ μ[34]

(C.d*) Si 'D_ap' ∈ μ et si μ appartient à un système modèle Ω, alors il y a dans Ω au moins une alternative doxastique μ* (41) à μ (par rapport à *a*) telle que p ∈ μ*.[35]

(C.DD*) Si 'D_ap' ∈ μ et si μ* est une alternative doxastique à μ (par rapport à *a*) dans un système modèle, alors 'D_ap' ∈ μ*.

(C.D¬) Si 'D_ap' ∈ μ, alors 'D_a¬p' ∈ μ.

(C.¬D*) Si '¬D_ap' ∈ μ et si μ appartient à un système modèle Ω, alors il y a dans Ω au moins une alternative doxastique μ* à μ (par rapport à *a*) telle que soit p est un membre de chaque alternative doxastique à μ* (par rapport à *a*), soit '¬p' est un membre de chaque alternative doxastique à μ* (par rapport à *a*).[36]

(C.DB) Si 'D_ap' ∈ μ, alors 'B_a¬B_ap' ∈ μ et '$B_a B_a$¬p' ∈ μ[37]

[Note : (C.¬K*) et (C.¬B*) sont des règles dérivées, bien entendu].

Le système KBD peut être divisé en trois sous-systèmes K, B et D. Le système K est le système qui résulte de la suppression du vocabulaire approprié, des règles de formation et des conditions C de 'B' et 'D'. Le système B est le système qui résulte de la suppression du vocabulaire approprié, des règles de formation et des conditions C de 'K' et 'D'. Le système D est le système qui résulte de la suppression du vocabulaire approprié, des règles de formation et des conditions C de 'B' et 'K' (ainsi que de (C.DB)).*

Dans l'énonciation du vocabulaire ci-dessus, des règles de formation et des conditions C, 'p' et 'a' sont utilisées de façon métalinguistique.[38] **(42)**

Appendice B

Considérons l'introduction dans KBD du procédé de notation suivant : soit '$D^n{}_a p$' (n est un entier positif) le résultat du fait d'écrire p et n occurrences de 'D_a' à la gauche de p. Par conséquent, '$D^2{}_a p$' s'abrège '$D_a D_a p$', '$D^3{}_a p$' s'abrège '$D_a D_a D_a p$', etc. Nous voulons montrer que '$D^n{}_a p \rightarrow \neg D^m{}_a p$' est autosuffisante dans KBD (tant que $m \neq n$). Pour le prouver, nous procédons par induction sur m de la manière suivante. Nous représentons m par n + 1, n + 2, ..., n + k. Nous montrons d'abord que l'affirmation vaut pour m = n + 1 :

(1) $D^n{}_a p \in \mu$	Hypothèse
(2) $\neg\neg D^{n+1}{}_a p \in \mu$	Hypothèse
(3) $D^{n+1}{}_a p \in \mu$	de (2) par (C.¬¬)
(4) $D^n{}_a \neg D^n{}_a p \in \mu$	de (3) par (C.D¬)
(5) $\neg D^n{}_a p \in \mu^*$	de (4) par (C.d*)
(6) $D^n{}_a p \in \mu^*$	de (1) par (C.DD*)

(5) et (6) violent (C.¬), et la preuve est complète. (Notez que '$D^{n+1}{}_a p$' s'abrège '$D_a D^n{}_a p$').

Notre hypothèse d'induction est que '$D^n{}_a p \rightarrow \neg D^{n+k}{}_a p$' est autosuffisante. Donc si nous pouvons montrer que '$D^n{}_a p \rightarrow \sim D^{n+k+1}{}_a p$' est autosuffisante, l'induction sera complète. Pour le montrer, nous remarquerons tout d'abord que '$B_a(D^n{}_a p \rightarrow \neg D^{n+k}{}_a p)$' est autosuffisante.[39] Nous complétons l'induction comme suit :

(1) $D^n{}_a p \in \mu$	Hypothèse
(2) $\neg\neg D^{n+k+1}{}_a p \in \mu$	Hypothèse
(3) $B_a(D^n{}_a p \rightarrow \neg D^{n+k} p) \in \mu$	Hypothèse
(4) $D^{n+k+1}{}_a p \in \mu$	de (2) par (C.¬¬)
(5) $D^{n+k}{}_a p \in \mu$	hypothèse inductive
(5) $D^n{}_a p \rightarrow \neg D^{n+k}{}_a p \in \mu^*$	de (3) par (C.b*)
(6) $D^n{}_a p \in \mu^*$	de (5) par (C.d*)
(7) $\neg D^n{}_a p \in \mu^*$ ou $\neg D^{n+k}{}_a p \in \mu^*$	de (6) par (C.→)

(5), (7) et (8) violent (C.~), et la preuve est complète.[40]

[Notez que '$D^{n+k+1}{}_a p$' s'abrège '$D_a D^{n+k}{}_a p$'].

Pour montrer que '$D^n{}_a p \rightarrow \neg D^m{}_a p$' est autosuffisante dans KBD lorsque $m < n$, nous remarquons que sa contraposition est virtuellement équivalente à '$D^m{}_a p \rightarrow \neg D^n{}_a p$'.[41]
Ces résultats complètent donc la preuve. L'argument démontrant que cette dernière formule est autosuffisante est pour l'essentiel celle donnée plus haut. **(43)**

Appendice C

Prouver : 'B' est primitif par rapport à 'D'.

Considérons le système qui résulte du fait d'accroître le système B par l'introduction de 'D' telle que définie par (D2). Appelons BD ce système. Nous voulons montrer qu'aucune formule de BD écrite uniquement en termes de 'D', '$\neg$', '$\rightarrow$' et p n'est virtuellement équivalente à '$B_a p$' dans BD. Pour le montrer, nous remarquons tout d'abord que BD peut être représenté axiomatiquement par les schémas d'axiomes et règles d'inférence qui suivent :[42]

(1) $p \rightarrow (q \rightarrow p)$

(2) $(p \rightarrow (q \rightarrow r)) \rightarrow ((p \rightarrow q) \rightarrow (p \rightarrow r)$

(3) $(\neg p \rightarrow \neg q) \rightarrow (q \rightarrow p)$

(4) $B_a(p \rightarrow q) \rightarrow (B_a p \rightarrow B_a q)$

(5) $B_a p \rightarrow B_a B_a p$

(6) $D_a p \rightarrow B_a(\neg B_a p \wedge \neg B_a \neg p)$

(7) $B_a \neg(\neg B_a \rightarrow B_a) \rightarrow D_a p$

Les schémas d'axiomes de BD restants sont donnés par la stipulation suivante : si A est un schème d'axiome de BD, alors de même pour '$B_a A$'.[43]

Considérons maintenant le système axiomatique qui résulte de l'addition du schéma d'axiome qui suit :

(8) $\neg B_a \neg p \rightarrow B_a p$.[44]

Appelons BD+ ce système. Pour montrer que BD+ est consistant, nous donnons simplement à '$\neg$' et '$\rightarrow$' leurs interprétations normales tout en donnant à 'B_a' l'interprétation 'Il est vrai que' et à 'D_a' l'interprétation 'Il est faux que'. Il va de soi qu'en vertu de cette interprétation, les schémas d'axiome ci-dessus sont tous des tautologies. Or on peut s'apercevoir aisément que BD+ contient BD mais n'est pas équivalent à BD. Par conséquent, s'il y a une formule de BD qui peut être construite uniquement à partir de 'D', '$\neg$', '$\rightarrow$' et p, par exemple '$\ldots D_a p \ldots$', et qui est virtuellement équivalente à '$B_a p$', alors '$B_a p \rightarrow \ldots D_a p \ldots$' et '$\ldots D_a p \ldots \rightarrow B_a p$' seront des théorèmes de BD, donc de BD+. Mais remarquez que '$D_a p$' est indéfendable dans BD+, donc '$\neg D_a p$' est autosuffisante dans BD+.[45]

A partir de là, on peut s'apercevoir facilement que toute formule de BD+ construite uniquement à partir de 'D', '¬', '→' et p sera virtuellement équivalente dans BD+ à l'une des formules suivantes :

(i) p
(ii) '¬p'
(iii) 'D_ap'
(iv) '¬D_ap' **(44)**

Mais on peut s'apercevoir facilement que 'B_ap' n'est virtuellement équivalente dans BD+ à aucune des formules (i)-(iv). Par conséquent, il s'ensuit qu'aucune formule de forme 'B_ap → ...D_ap...' ou '...D_ap...→ B_ap' n'est un théorème de BD+, ni de BD. **(45)**

Appendice D

Entendons par modalité D toute formule constructible uniquement à partir de p, '¬' et 'D'. Nous voulons montrer que toutes les modalités D sont *virtuellement équivalentes* à (⇔) l'une des formules

(i) p
(ii) '¬p'
(iii) '$D^m{}_a$p'
(iv) '¬$D^m{}_a$p'

Preuve : soit q une modalité D quelconque. Supposons à titre d'hypothèse inductive que toute formule contenant moins d'occurrences de '¬' et de 'D_a' que q ⇔ une des formules de (i)-(iv). Nous avons trois cas de figure :

Cas 1. q = p : l'affirmation va de soi.

Cas 2. q = ¬r : en vertu de l'hypothèse inductive, r ⇔ une des formules de (i)-(iv), disons r*. Puisque r ⇔ r*, ¬r ⇔ ¬r*, i.e. q ⇔ ¬r*. Puisque la négation de chacune des formules de (i)-(iv) ⇔ une des formules de (i)-(iv), il s'ensuit que q ⇔ une des formules de (i)-(iv).

Cas 3. q = D_ar : remarquez tout d'abord que 'D_ap' ⇔ 'D_a¬p'. En vertu de l'hypothèse inductive, r ⇔ une des formules de (i)-(iv), disons r*. Puisque r ⇔ r*, 'D_ar' ⇔ 'D_ar*', i.e. q ⇔ 'D_ar*'. Puisque la préfixation de 'D_a' à chacune des formules de (i)-(iv) ⇔ une des formules de (i)-(iv), il s'ensuit que q ⇔ une des formules de (i)-(iv).

RÉFÉRENCES

[1] J. Hintikka, "Semantics for Propositional Attitudes", réimprimé in *Reference and Modality*, éd. par L. Linsky, Oxford Univ. Press, 1971, pp. 145-167, en part. Sec. IV, pp. 150-151.

[2] J. Hintikka, *Knowledge and Belief*, Cornell Univ. Press, 1962.

[3] E. Partridge, *Origins: A Short Etymological Dictionary of Modern English*, Macmillan, 1958, p. 164.

[4] H.W. Fowler, *Modern English Usage*, Oxford Univ. Press, 1944, pp. 121-122.

[5] Oxford English Dictionary, vol. XII, V-Z, Oxford Univ. Press, 1933 (réimprimé en 1961); cf. "whether", en part. "whether or not".

[6] *The Encyclopedia of Philosophy*, vol. 2, éd. par Paul Edwards Macmillan, 1967 ; cf. "doubt", de H.C. Frankfurt.

[7] G. Berkeley, *Three Dialogues Between Hylas and Philonous*, Library of Liberal Arts, 1954, p. 11.

[8] Cf. les objections de Bourdin in *The Philosophical Works of Descartes*, vol. II, traduit par E.S. Haldaine et G.R.T. Ross, Dover, 1930, p. 264.

NOTES DU TRADUCTEUR

[1] 'Système modèle' désigne chez Hintikka la procédure sémantique des mondes possibles (ou sémantique relationnelle) : un système modèle est un ensemble de mondes possibles, c'est-à-dire une structure de Kripke W interprétée au sein d'un modèle $\mathcal{M} = \langle v, W, R \rangle$. L'approche héritée de Hintikka n'étant ni axiomatique ni algébrique, la sémantique en termes de matrices logiques sera donc absente de cet article.

[2] 'Fréquentatif' se dit d'un verbe qui marque une action fréquemment répétée (ex. : *clignoter*, *cachoter*, *criailler*).

[3] Ces alternatives exclusives correspondent à des *contradictions* : ϕ ou non-ϕ. On retrouve cette idée d'alternative ouverte dans l'expression 'savoir si' : 'knows whether' signifie 'savoir si' et suggère une alternative encore possible entre la vérité de ϕ ou de non-ϕ, tandis que 'knows that' exclut cette alternative : savoir que ϕ/non-ϕ signifie que ϕ/non-ϕ est vraie, auquel cas l'alternative est close.

[4] Par opposition aux combinaisons multiples du concept de savoir (+ que, si, qui, quand, où, comment), le 'doute' de la langue française ne possède qu'une seule construction possible : 'doute *que*', qui exprime à peu près l'indétermination ou alternative ouverte.
A noter que l'acte de discours, i.e. le fait de dire 'douter que ϕ' suggère ou implique (au sens d'une implicature de Grice) que l'on a tendance à croire que ϕ est fausse ; autrement dit, l'alternative est encore possible (que p soit vraie) mais l'expression indique une certaine inclination du locuteur vers la négative. Quoi qu'il en soit, cet aspect pragmatique du discours

n'altérera pas la notion sémantique de doute, qui exprime une alternative égale entre croyance et incroyance. Hintikka dirait que cet effet pragmatique est sans doute dû à l'utilisation de la particule démonstrative 'que' : conformément au principe d'économie de Grice, celui qui dit 'je doute que ϕ' marque son intention explicite de porter le doute sur (la vérité de) ϕ, en sorte qu'il indique implicitement sa préférence consistant à croire plutôt que non-ϕ (que ϕ est fausse). La même chose est dite au sujet du concept de connaissance (cf. [Hintikka 1962], Partie 1) : alors que dire 'je ne sais pas si ϕ' laisse l'alternative ouverte entre la vérité de ϕ ou de non-ϕ, dire 'je ne sais pas que ϕ' semble absurde à dire [*self-defeating*] puisque la particule 'que' suggère ici la vérité de son complément. Conformément au paradoxe de Moore, la référence au temps devient pertinente puisqu'elle distingue l'énoncé 'absurde' 'je ne sais pas que ϕ' de l'énoncé 'je ne savais pas que ϕ' (suggérant que je connais désormais la vérité de ϕ). Paradoxalement, la langue française exprime cette indétermination *via* la particule unique 'que' utilisée d'habitude pour la *détermination*, alors que la langue anglaise reste fidèle à la particule d'indétermination 'whether' dans 'doubts whether'. Pour distinguer en langue française les sens alternatif et non-alternatif du doute, la solution n'est donc pas grammaticale mais semble plutôt d'ordre orthographique, ou *tropique* (voir section 1.3.3). En effet, le sens alternatif (qui nous occupera ici) du doute de ϕ donne la forme conjuguée à l'indicatif présent : 'je doute que ϕ *est* le cas', tandis que le sens non-alternatif ('pragmatique') se conjugue sous la forme subjonctive : 'je doute que ϕ *soit* le cas'. La solution est forcée, toutefois : le doute alternatif s'exprime couramment non pas 'je *doute que* ϕ est le cas', mais bien plutôt par 'je *me demande si* ϕ est le cas', qui réintroduit par ailleurs l'alternative en utilisant de nouveau la particule pronominale 'si'. Etant donné que cet article parle du 'doute', nous ne remplacerons pas malgré tout le verbe 'douter' par 'se demander', même si la traduction vernaculaire de ce dernier semble bien plus acceptable pour le doute alternatif.

De même que pour 'croire si', 'doute si' est un *solécisme*. Sur le cas du solécisme 'croire si' et son explication en termes sémantique d'incapacité pour le concept de croyance à satisfaire la clause de succès (i.e. l'opérateur 'B' n'est pas **T**-valide), voir [Hintikka 1989], en particulier l'article : "Different Constructions in Terms of the Basic Epistemological Terms" (pp. 1-25).

[5] Nous traduisons 'clause' par '(proposition) complétive', qui désigne toute proposition grammaticale connectée à une proposition principale par l'intermédiaire d'un pronom relatif ('que', ici). La définition sémantique du doute est donc la contrepartie doxastique de l'expression 'ne pas savoir si' (ou 'se demander si') : '*a* ne sait pas si' signifie que *a* ne sait que ϕ et ne sait pas que non-ϕ. Par dualité, 'savoir si' signifie donc la négation de 'douter que', i.e. 'ne pas douter que' : si *a* sait si ϕ, alors il ne doute pas de la valeur sémantique de ϕ et ce n'est pas le cas que (*a* doute que ϕ et *a* doute que non-ϕ).

Pour résumer formellement, et sans empiéter sur l'analyse à suivre de l'auteur, nous avons donc déjà les quelques relations suivantes :

Savoir si ϕ	$K_a\phi \vee K_a\neg\phi$
Ne pas savoir si ϕ	$\neg(K_a\phi \vee K_a \neg\phi) \equiv \neg K_a\phi \wedge \neg K_a\phi$
Douter que ϕ = ne pas savoir si ϕ	$D_a\phi \equiv \neg K_a\phi \wedge \neg K_a\neg\phi$
Ne pas douter que ϕ = savoir si ϕ	$\neg D_a\phi \equiv \neg(\neg K_a\phi \wedge \neg K_a\neg\phi)$

$\equiv K_a\phi \vee K_a\neg\phi$

Douter que p implique ne pas douter que non-ϕ $D_a\phi \rightarrow \neg D_a\neg\phi \equiv \neg(D_a\phi \wedge \neg D_a\neg\phi)$

La dernière propriété du doute correspond à la propriété de *consistance* du système D de Lewis : le doute semble être donc un opérateur modal isomorphe à celui de croyance.

[6] La définition syntaxique de l'opérateur de doute est donc isomorphe à la modalité de la contingence ou 'possibilité bilatérale' ∇ : $\nabla\phi =_{df} \neg\Box\phi \wedge \neg\Box\neg\phi$. La dualité des opérateurs modaux et la commutation permettent de simplifier la définition : $\nabla p =_{df.} \Diamond\phi \wedge \Diamond\neg\phi$. Le doute est un opérateur non primitif, puisqu'il se définit en termes primitifs de croyance ; il correspond ainsi à la forme Y dans un hexagone de Blanché, où A et E constitueraient les notions respectives de croyance et d'incroyance.

[7] Cette version distingue quelque peu la contingence pure du doute. Par analogie en retour, une application de (D2) à ∇ ne signifie plus l'alternative simple mais la nécessité de l'alternative, i.e. $\Box(\Diamond\phi \wedge \Diamond\sim\phi)$.

[8] (3) est équivalente à (D2) en vertu de la dualité des opérateurs modaux ($\neg B = C\neg$), puis par double négation et commutativité.
Cette itération des opérateurs modaux sert ainsi à distinguer les variables a susceptibles de ne pas croire et de douter, grâce à l'opérateur de croyance compatible qui est la contrepartie modale faible $\Diamond$ de la croyance (opérateur fort $\Box$). La distinction est comparable à celle que nous avons observée dans un article d'Englebretsen relatif aux *contraires* ; l'auteur y avait expliqué qu'une pierre n'est ni voyante ni non-voyante, puisqu'elle n'est peut être privée d'une propriété qu'il n'appartient pas à son essence de posséder. Si l'on remplace $\neg B$ par la négation propositionnelle '~' et 'non' par la négation des termes prédicats, on obtenait ainsi : $\neg Fa$ et $\neg$non-Fa, de structure similaire à $\neg B_a\phi \wedge \neg B_a\neg\phi$. La solution employée ici par Hart pour distinguer les variables susceptibles de douter des autres aurait peut-être une contrepartie pour la privation : on pourrait introduire des modalités pour différencier la négation suivant l'entité à laquelle elle s'applique (énoncé, prédicat, terme de prédicat).

[9] Etant donné le contexte moderne de la présente discussion, il va de soi que la négation de termes prédicats est réductible à la négation de prédicats (est non-ϕ = n'est pas ϕ) et à la négation propositionnelle. Nous travaillons dans une analyse frégéenne de type fonction-argument, bien sûr.

[10] Le problème de traduction est clairement anglo-anglais, ici, puisque la langue française ne dispose pas de distinction pronominale de ce genre en réserve. Dans une note précédente, une distinction orthographique a été suggérée entre 'est le cas' et 'soit le cas' pour marquer la différence de sens du doute en langue française.
La suggestion faite ici par Hart recoupe deux remarques faites précédemment : premièrement, le sens pragmatique de ce 'doute' différent de (D2) (notons-le : D') exclut l'alternative ouverte entre ϕ et non-ϕ, il suggère que a croit que non-ϕ : $D'_a\phi =_{df.} B_a\neg\phi$; deuxièmement, cette suggestion est conforme à la signification des pronoms de la langue anglaise, soulignée

par Hintikka dans les premières sections de son *Knowledge and Belief* et dans les *Intentions de l'Intentionalité* : + *whether* exprime une alternative ouverte entre ϕ et non-ϕ, + *that* exprime la conviction établie que ϕ ou non-ϕ est vraie. L'absence d'une telle distinction pronominale dans la langue française explique sans doute l'ambiguïté de la notion française de 'doute'.

[11] Au total, le doute alternatif D_a correspond au doute 'sceptique', sans engagement sur la vérité de l'un des membres de l'alternative et conformément à l'attitude neutre de Pyrrhon ; par opposition, le doute engagé D'_a correspond au doute 'nihiliste', qui caractérise l'attitude consistant à croire que tout est faux (à 'disbelieve' ou 'incroire' toute proposition quelle qu'elle soit, au sens où Hart entend la notion d'incroyance et qui n'est pas le nôtre : pour nous comme pour [Vernant 2003], l'incroyance n'est précisément pas la croyance de la fausseté). Le doute nihiliste paraît incohérent à ce titre, dans la mesure où il revient à croire la fausseté de 'tout et son contraire' (de ϕ et non-ϕ, pour tout ϕ). Quant à la distinction entre 'doute destructeur' (sceptique) et 'doute méthodique' (constructif, ou cartésien), il semble qu'elle concerne que la seule forme 'D_a' et se manifeste dans les *conséquences* du doute : le doute constructif (cartésien) implique que certaines conséquences de ce doute donnent des croyances positives, i.e. que le cartésien est conduit à croire certaines choses en conséquence de son doute ; le doute destructeur (pyrrhonien) implique que toutes les conséquences de ce doute sont des absences de croyance, i.e. que le sceptique est conduit à ne croire à rien. Cela dit, peut-on dire que, dans les *Méditations Métaphysiques*, la croyance au cogito ou d'autres résultats positifs de la méditation de Descartes sont des *conséquences* de son doute méthodique de départ ? Si oui, ces conséquences sont-elles logiques, psychologiques, épistémologiques, etc. ?

[12] Dans la langue française, il semble que la forme adjectivale 'douteux' connote également le sens de l'incroyance selon Hart : *dire* qu'une chose est douteuse présuppose que l'auteur de la déclaration pense qu'elle est fausse, en effet, mais ne dispose pas de preuves suffisantes pour justifier son attitude : il le croit sans le savoir, autrement dit.

[13] Puisque la notion de doute est une notion dérivée et définie en termes primitifs de croyance, elle se définit sémantiquement en termes d'alternative *doxastique* ; il en va autrement pour les notions primitives de connaissance (K/P) et de croyance (B/C), définies quant à elles en termes respectifs d'alternative épistémique et doxastique. La définition de la notion primitive de croyance est conforme à la réduction que l'on trouve en sémantique relationnelle, en termes de quantification sur des mondes :

$B_a\phi \in w \in W =_{df.} \forall w^* Rw, \phi \in w^*$

$C_a\phi \in w \in W =_{df.} \exists w^* Rw, \phi \in \mu^*$

On peut expliquer (C.d*) par l'absurde : en admettant '$D_a\phi' \in \mu$, si ϕ *n'appartenait pas* à *au moins une* alternative à w alors ϕ n'appartiendrait à *aucune* de ces alternatives, et donc $\neg\phi$ serait le cas pour toute alternative (par rapport à *a*). Or si $\neg p$ est le cas pour toute alternative doxastique, cela veut dire que *a* croit que $\neg p$, i.e. $B_a\neg\phi$. Or $D_a\phi$ implique $\neg B_a\neg\phi$; donc si '$D_a\phi' \in w$ alors ϕ appartient à moins une alternative w* à w. En fait, la négation de la

condition (C.d*) correspond à l'incroyance telle que Hart la caractérise, c'est-à-dire l'analogue doxastique de l'impossibilité $\Box\neg$ et que le doute exclut (en tant qu'analogue doxastique de la contingence ∇). 'D' l'exclut *a fortiori* sous sa forme (D2) : ce qui est vrai de la contingence l'est d'autant plus de la contingence nécessaire.

[14] Contrepartie 'dubitative' du système de **S4** de Lewis, dont le schéma caractéristique est la validité ($\Box\phi \rightarrow \Box\Box\phi$) et donne l'introspection positive en logique modale épistémique.

[15] Cette règle montre clairement que D fonctionne comme la notion de *contingence* : si ϕ est contingente, alors il est possible que ϕ *et* que non-ϕ, i.e. $\nabla\phi = \nabla\neg\phi$. En effet : $\Box\nabla\neg\phi \equiv \Box(\Diamond\neg\phi \wedge \Diamond\neg\neg\phi) \equiv \Box(\Diamond\neg\phi \wedge \Diamond\phi) \equiv \Box(\Diamond\phi \wedge \Diamond\neg\phi) = \Box\nabla\phi$.

[16] Cette règle explicite la règle précédente : parallèlement à l'analyse de la contingence, douter de ϕ signifie que l'on croit (qu'il est possible) que l'on ne croit ni ϕ (que ϕ n'est pas nécessaire) ni non-ϕ (que non-ϕ n'est pas nécessaire). Attention à ne pas simplifier $\neg B\neg\phi$ par $B\phi$! Cette opération de dualité ne fonctionne qu'entre les opérateurs fort et faible correspondants : $\neg B\neg\phi = C\phi$, et la croyance est uniquement un opérateur fort. Sans cette précaution, le doute signifierait que l'on ne croit et que l'on croit p à la fois, ce qui est contraire aux conditions C de *consistance*. (Voir la note précédente sur le critère de consistance pour les opérateurs de croyance et de doute : $B/C\neg\phi = \neg B/C\phi$, pour tout ϕ.

[17] En vertu de (4), $\neg D$ donne en effet ceci :

$$\begin{aligned}\neg D\phi &= \neg(BC\phi \wedge BC\neg\phi)\\ &= \neg BC\phi \vee \neg BC\neg\phi\\ &= C\neg C\phi \vee C\neg C\neg\phi\\ &= CB\neg\phi \vee CB\phi\end{aligned}$$

C représente l'alternative doxastique w*, en sorte que soit $B\neg\phi$ soit $B\phi$ dans w* ; or en vertu de la définition quantifiée de B, $\neg D\phi$ signifie donc que, dans *toute* alternative doxastique à w*, on a soit ϕ soit $\neg\phi$.

[18] *Indéfendabilité* est un néologisme métalogique introduit par [Hintikka 1962] comme contrepartie épistémique de l'antilogie : étant donné que les théorèmes épistémiques de son système ne sont vérifiés que dans des conditions idéales, leur validité n'est pas effective mais valable uniquement pour des agents épistémiques *a* idéalement rationnels (les 'Logically Perspicuous Knowers', comme Lemmon les appellera). En conséquence, Hintikka a introduit la notion d''indéfendabilité' de p comme signifiant la validité *pour ces agents idéaux.* L'auteur a proposé ainsi les adaptations métalogiques suivantes : validité = autosuffisance [*self-sustenance*], invalidité/simple satisfiabilité = défendabilité, antilogie = indéfendabilité, implication = implication virtuelle.
Quant à l'*indéfendabilité doxastique* de p, elle correspond à l'indéfendabilité simple de $B\phi$. Hintikka a introduit cette complication dans le contexte du paradoxe de Moore qui suit : alors que la formule '$B_a\phi \wedge \neg\phi$' est défendable en tant que telle, le problème de la formule de

Moore ne vient pas de son indéfendabilité en tant que telle mais de son indéfendabilité doxastique, i.e. de l'indéfendabilité de $B_a(B_a\phi \wedge \neg\phi)$.

[19] La seconde variante du paradoxe de Moore est : $B\phi \wedge \neg\phi$, i.e. $D\neg\phi \wedge \neg\phi$ et non $\neg D\phi \wedge \neg\phi$. En effet, en vertu de (D2) $\neg D\phi \wedge \neg\phi$ donne ceci : $[\neg B(\neg B\phi \wedge \neg B\neg\phi)] \wedge \phi = (\neg B\neg B\phi \vee \neg B\neg B\neg\phi) \wedge \phi = (CB\phi \vee CB\neg\phi) \wedge \phi$. Or bien que $CB\phi \wedge \neg\phi$ soit indéfendable, elle ne peut pas être déduite de la disjonction précédente : en effet, $(P \wedge R)$ n'est pas une conséquence logique de $(P \vee Q) \wedge R$ (l'élimination de la disjonction n'est pas une règle de déduction naturelle). En revanche, $D\neg\phi \wedge \phi$ donne ceci : $BB\phi \wedge BB\neg\phi \wedge \phi$, dont une conséquence logique est en effet la formule indéfendable $BBp \wedge \neg p$. Mais on peut simplifier davantage encore cette seconde variante : en vertu de (C.D¬), $D\phi$ et $D\neg\phi$ sont équivalents, donc $B\phi \wedge \neg\phi = D\neg\phi \wedge \neg\phi = D\phi \wedge \neg\phi$. En termes aléthiques, cela veut dire que $\Box\phi \wedge \neg\phi$ équivaut à $\Box\nabla\phi \wedge \neg\phi$.

[20] Avant de passer à l'explication intuitive de ce résultat, on peut expliquer (10) du point de vue syntaxique (axiomatique) et sémantique.
Syntaxiquement : le détachement des modalités $\Box\phi$ et $\Box\psi$ à partir de l'implication valide $\phi \rightarrow \psi$ est permis en logique modale grâce à un critère des logiques modales dites *normales* ou *régulières*, i.e. la règle de nécessitation : si $\vdash A$, alors $\vdash\Box A$. Puisque A est ici l'implication virtuelle $\phi \rightarrow \psi$, on a donc $\vdash(\phi \rightarrow \psi)$, donc $\vdash\Box(\phi \rightarrow \psi)$. En vertu d'un autre critère des logiques modales normales, i.e. la clôture modale sous implication logique (= implication matérielle toujours vraie) ou distributivité de la modalité sur l'implication logique, de $\Box(\phi \rightarrow \psi)$ on obtient : $(\Box\phi \rightarrow \Box\psi)$. Or comme l'auteur va l'expliquer plus bas, la règle de nécessitation n'est *pas* applicable lorsque $\Box = D$: en vertu d'une contrainte de rationalité valable pour tout contexte modal, aucune personne ne peut douter rationnellement de vérités logiques, ce qui signifie que la logique du doute n'est pas une logique normale. Cette particularité de la logique du doute est due, me semble-t-il, à l'aspect 'négatif' de l'opérateur de doute D, défini en termes de $\neg B$: étant donné que l'opinion du doute porte sur des propositions contingentes, donc susceptibles d'être fausses, elle a tendance à 's'opposer' ou à 'exclure' les cas de propositions logiques, supposées connues et prises pour certaines. Par opposition, les opérateurs 'positifs' de connaissance et de croyance ont tendance à 'confirmer' ou à 'inclure' les cas de propositions logiques.
Sémantiquement : puisque le doute s'exprime en termes d'alternative doxastique, $D_a\phi \in w$ ssi il existe *au moins une* alternative doxastique w* à w telle que $p \in w^*$. Or la définition de D en termes d'alternative *existentielle* la distingue des définitions de K et B en termes d'alternative *universelle* : $D_a(\phi \rightarrow \psi) \in w$ signifie que $(\phi \rightarrow \psi) \in w^*$, mais $D_a\phi$ et $D_a\psi$ n'appartiennent pas forcément à *la même* alternative w* à w. Cela est dû à la nature non-distributive de l'existentielle, structurellement analogue à l'opérateur modal faible : $\exists(\phi \rightarrow \psi)$ n'implique pas $(\exists\phi \rightarrow \exists\psi)$, $\Diamond(\phi \rightarrow \psi)$ n'implique pas $(\Diamond\phi \rightarrow \Diamond\psi)$, de même que $D_a(\phi \rightarrow \psi)$ n'implique pas $(D_a\phi \rightarrow D_a\psi)$. Autrement dit, les alternatives à $D_a\phi$ et $D_a\psi$ peuvent être disjointes, i.e. $w \cap w^* = \varnothing$.

[21] Si ψ est autosuffisante, ψ est une instance de vérité logique et *a* ne peut pas douter de ces vérités, en vertu d'une contrainte de rationalité semblable à celle des logiques épistémique et doxastique. En effet, si l'on note T la classe des vérités logiques, alors en vertu d'un des paradoxes de l'implication matérielle on a $(\phi \rightarrow T) = T$ pour tout ϕ (*verum sequitur ad quodlibet* : le vrai s'ensuit de n'importe quoi), y compris pour les ϕ dont *a* doute. Donc pour certaines propositions ψ : $D_a\phi \wedge (\phi \rightarrow \psi) \wedge \neg D_a\psi$ est satisfiable.

[22] Si ψ est indéfendable (noté : $\perp$), alors en vertu d'un des paradoxes de l'implication matérielle on a $(\perp \rightarrow \phi) = T$ pour tout ϕ (*ex falso sequitur quodlibet* : du faux il s'ensuit n'importe quoi). La contrainte de rationalité voulant que *a* ne doute jamais des vérités logiques T, on a donc, pour certaines propositions ψ : $D_a\phi \wedge (\psi \rightarrow \phi) \wedge \neg D_a\psi$. Les deux cas que l'auteur évoque ici, $\phi = T$ et $\psi = \perp$, prouvent ainsi la non-régularité de la logique modale du doute.

[23] L'expression anglaise 'to disbelieve' se traduit parfois par le barbarisme 'incroire' afin de conserver la racine 'croire', ainsi que dans [Vernant 2003]. On pourrait remplacer ici cette notion indélicate par 'croire la fausseté de' ; mais en guise de contrepartie, il faudrait 'charger positivement' la croyance positive de forme neutre 'croire que' par l'expression 'croire la vérité de', ce qui alourdirait les phrases. En guise d'alternative, on choisira la contrepartie illocutoire de l'incroyance, i.e. le rejet ou assertion inversée $\dashv$: l'incroyance de Hart (voir note 11), 'je crois que ϕ est fausse', sera traduite ici par 'je rejette ϕ'.
La preuve du rejet de ϕ s'obtient par contraposition : si '$\phi \rightarrow \psi$' est autosuffisante, alors $\vdash(\phi \rightarrow \psi)$, donc par contraposition on obtient : $\vdash(\neg\psi \rightarrow \neg\phi)$. Etant donné que la logique doxastique de Hintikka est *normale*, la nécessitation et la clôture sous implication logique y sont valables. Donc de $\vdash(\neg\psi \rightarrow \phi)$ on obtient successivement : $\vdash B_a(\neg\psi \rightarrow \phi)$, puis $\vdash(B_a\neg\psi \rightarrow B_a\neg\phi)$. Donc on obtient : $[B_a(\neg\psi \rightarrow \phi) \wedge B_a\neg\psi] \rightarrow B_a\neg\phi$.

[24] Puisqu'un critère de rationalité général veut que toute personne *a* accepte (B_a) les vérités logiques, toute personne rejette ($B_a\neg$) donc leurs négations (antilogies ou contradictions) : $\forall\phi$ $((\phi = \perp) \rightarrow \vdash B_a\neg\phi)$. Autrement, si $B_a\neg\phi \in w$ alors $\neg\phi \in w^*$ pour *toute* alternative w* à w (par rapport à *a*); en revanche, si $D_a\neg\phi \in w$ alors $\neg\phi \in w^*$ pour *au moins une* alternative w* à w, mais pas forcément pour *toutes* !
La différence entre le doute et l'incroyance correspond une fois encore à la distinction aléthique entre possibilité et contingence : de même que la nécessité de $\neg\phi$ implique la possibilité unilatérale de $\neg\phi$ mais exclut la possibilité bilatérale (contingence) de $\neg\phi$, l'incroyance de $\neg\phi$ (pour un agent *a* quelconque) implique la compatibilité de $\neg\phi$ avec l'ensemble des croyances de *a* mais exclut le doute de *a* sur $\neg\phi$. Le critère de rationalité interdit donc le doute sur toute proposition ϕ telle que $\phi = \perp$.

[25] Hintikka emprunte la notion de 'self-intimation' à Ryle, qu'il utilise à titre de justification psychologique standard de l'introspection positive (cf. [2], p. 57).
L'implication virtuelle pour la croyance est à sens unique : en effet, Hintikka montre dans [2] (cf. Partie 5) que $B_a\phi \rightarrow B_aB_a\phi$ est autosuffisant mais que la converse n'est pas toujours vraie.

[26] L'idée d'un accès privilégié à soi de ses propres états mentaux pose problème dans le cas du concept de connaissance : celui-ci ne semble pas aussi psychologique que les états de croyance et de doute, puisqu'il exige dans sa définition une clause objective de vérité (cf. $K_a\phi \rightarrow \phi$).
L'auteur parle ici de démonstration 'logique' de l'itération modale. Il faudrait peut-être nuancer ce qualificatif : Hintikka rejette dans [2] l'argument psychologique de l'introspection en lui préférant un sens logique ou 'quasi-performatif' (*dire* que l'on sait implique que l'on sait que l'on sait quelque chose) du concept de connaissance.

[27] (C.BB*ep) est définie dans [2] p. 50 : Si '$B_a\phi' \in w$ et si w* est une alternative *épistémique* à w (par rapport à a) dans W, alors '$B_a\phi' \in w^*$. En termes informels, cette condition signifie que l'on ne croit que ce que l'on sait, ainsi que Hintikka l'explique lorsque, dans [2] p. 51, il déclare que (C.BB*ep) équivaut à la condition suivante : (C.BK) Si '$B_a\phi' \in w$, alors '$K_aB_a\phi' \in w$. Or puisque la connaissance implique l'abandon de certaines croyances injustifiées, savoir que l'on croit est donc incompatible avec certaines de nos croyances effectives : "Or il est sans doute possible que l'on en sache plus [dans w*] que l'on en sait maintenant [dans w] et que, dans le même temps, on ait moins de croyances que l'on a maintenant. En effet, il est très probable que l'on aurait abandonné certaines de nos croyances si l'on avait davantage d'infirmation. Mais pareil abandon de croyances est précisément ce qui est interdit par (C.BB*ep)".
A noter que cette interprétation introduit une distinction de temps entre l'antécédent et le conséquent, alors qu'une des clauses de départ de [2] est l'unité de temps des inférences, dans le but d'ignorer le problème des changements d'avis dans l'analyse des concepts de connaissance et de croyance. Cette clause n'est-elle pas violée dans le but de réfuter (C.BB*ep), ici ?

[28] Puisque le doute se définit en termes de croyances et que les croyances sont révisables (falsifiables) par un gain de connaissance, (C.DK) représente une condition rejetée sur la base du rejet de sa contrepartie doxastique (C.BK).

[29] Douter de ses propres doutes pourrait exprimer un argument contre le doute méthodologique de Descartes, dont le but était d'amener par élimination jusqu'à une connaissance indubitable : le *cogito*. Cette attitude pourrait s'exprimer sous la forme : le doute portant sur mes perceptions, sur la vérité des affirmations mathématiques, sur ma propre existence, etc., n'est-elle pas une illusion ? Plusieurs auteurs ont prétendu (parmi lesquels des phénoménologues) que la mise entre parenthèses de mes facultés sensibles était impossible et que la conclusion du cogito ne pouvait être obtenue sans les facultés précédentes mises en doute : ces auteurs doutent donc non pas de ce dont Descartes a douté, mais de la possibilité même de douter de ces facultés.

[30] Erratum : dans l'article original, l'auteur avait écrit que (6) est déduite par (C.DD*) de (5), et non de (1).

[31] Le type de scepticisme considéré ici à l'abri concerne notamment l'attitude de Montaigne : 'il y a au moins une chose que je sais, c'est que je ne sais rien' ne serait pas une affirmation défendue par un sceptique radical. L'introspection positive est liée ici à la question de la nature itérative du doute, liaison dont témoignent certains articles portant sur le lien entre introspection positive et scepticisme.

[32] I.e. les parenthèses, utilisées comme des signes de ponctuation pour distinguer les portées des connecteurs : '(' et ')'. Les points sont omis, puisqu'ils ne sont plus usités de nos jours.

[33] L'auteur ne se demande pas pourquoi, à la suite de Hintikka, il écrit certaines propositions entre guillemets et d'autres pas. Seules les formules déclaratives *atomiques* de type 'p' ne sont pas entre guillemets ; dans le cadre de la logique épistémique KB de Hintikka, deux auteurs, Chisholm et Hocutt, avaient prétendu que les guillemets indiquaient la *mention* d'une proposition et introduisaient un aspect pragmatique dans l'interprétation des conditions. Cela concorde avec le rejet chez Hintikka de l'introspection et sa défense 'quasi-performative' de l'introspection positive.

[34] Erratum : l'auteur a écrit (à tort) '$B_a\phi$' $\in$ w*, qui donne la condition précédente (C.BB*).

[35] En vertu de la définition sémantique 'existentielle' du doute (cf. l'analogie avec la contingence), on peut dire que (C.D*) = (C.d*), alors que (C.B*) et (C.K*) impliquent mais ne sont pas équivalents à (C.b*) et (C.k*), respectivement.

[36] Cette condition est analogue à la *non-contingence* de ϕ : si ϕ n'est pas contingente dans w, alors soit ϕ est soit nécessaire (vraie dans toute alternative w* à w), soit impossible (fausse dans chaque alternative w* à w). A noter que, dans le cas de la contingence de type aléthique, la modalité n'est pas une attitude propositionnelle, ce qui signifie que les alternatives n'y sont pas subjectives mais objectives et posées indépendamment de toute variable d'individu *a*.

[37] Puisque l'opérateur B se distribue sur '$\wedge$' (voir plus haut), (C.DB) correspond à la version distributive de (D2).

[38] Cette précision semble répondre à notre note ci-dessus concernant l'usage des guillemets : la proposition atomique 'p' ne serait pas entourée de guillemets parce qu'elle serait utilisée de façon métalinguistique. Mais pour quelle raison : pour éviter le paradoxe de l'auto-référence dans l'énonciation des conditions, étant donné que 'ϕ' est synonyme de 'ϕ est vraie' ? Si tel est le cas, faut-il en conclure que toute proposition complexe, modale ou non-modale, est préservée du paradoxe ?

[39] Cette observation n'est pas gratuite : le fait que la formule soit dans la portée de 'B' va permettre d'appliquer (C.b*) et d'obtenir la contradiction attendue pour compléter la preuve.

[40] Il y a trois termes, ici, puisque (8) est une alternative issue de (C.→) : en vertu de $(D^n{}_a\phi \rightarrow \neg D^{n+k}{}_a\phi) \in w^*$, soit $(\neg D^n{}_a\phi) \in w^*$ soit $(\neg D^{n+k}{}_a\phi) \in w^*$. Si l'on choisit la première solution $(\neg D^n{}_a\phi) \in w^*$, alors ce sont (7) et (8) violent (C.¬); si l'on choisit la seconde solution $(\neg D^{n+k}{}_a\phi) \in w^*$, alors ce sont (5) et (8) violent (C.¬). Donc (6) se décompose en deux sous-conditions indéfendables.

[41] Lorsque m < n, la formule ci-dessus n'aurait pas été autosuffisante si KD avait inclus parmi ses conditions C la contrepartie 'dubitative' de la clause de succès épistémique (C.K), à savoir:
(C.D) Si '$K_a\phi$' $\in \mu$, alors $p \in \mu$.
Or l'auteur rejette (C.D) par contraposition sur la formule prouvée plus haut, i.e. $(D^m{}_a\phi \rightarrow \neg D^n{}_a\phi)$: $\vdash(A \rightarrow \neg B)$, donc $\vdash(B \rightarrow \neg A)$. Soit $A = D^m{}_a\phi$ et $B = D^n{}_a\phi$, alors : $\vdash(D^n{}_a\phi \rightarrow \neg D^m{}_a p)$.
On aurait pu le montrer sinon par *analogie* : dans la mesure où $D\phi$ se comporte comme la contingence nécessaire $\Box\nabla\phi$ et que, même après la simplification permise de $\Box\nabla\phi$ en $\nabla\phi$, $\nabla\phi \in w$ n'implique pas $\phi \in w$ (ce qui est contingent n'est pas toujours vrai). Dans ce cas, quelle condition Hart emploie-t-il pour passer de (4) à (5) ? Malheureusement, il ne précise pas les conditions de la preuve ci-dessus dans l'article.

[42] 'Verum sequitur ad quodlibet' : du vrai il s'ensuit n'importe quoi ; Distributivité de '→' ; Contraposition ; Distributivité de 'B' ; Itération de 'B' : (C.BB*) ; Implication de (CDB), en vertu de la traduction de 'Déf.' par '≡' ; Converse de (6) + interdéfinition des connecteurs $\neg A \wedge \neg B = \neg(\neg A \rightarrow B)$.

[43] Contrepartie doxastique de la règle de nécessitation des logiques modales *normales* : pour tout ϕ, si $\vdash\phi$ alors $\vdash B_a\phi$.

[44] (8) est la converse de la propriété de consistance du système modal D : $\Box\phi \rightarrow \Diamond\phi$, dont la contrepartie doxastique est : $(B_a\phi \rightarrow C_a\phi) \equiv (B_a\phi \rightarrow \neg B\neg_a\phi)$. Donc $(B_a\phi \equiv \neg B_a\neg\phi)$, en sorte que 'B' satisfait la propriété du système **Q** (= **D** + (8)). En termes sémantiques de structures de Kripke, cette équivalence signifie que si $\Diamond\phi \in w$ dans une structure W, alors il n'y qu'une seule alternative w* à w dans W : 'B' satisfait la condition de fonctionnalité. (8) implique ainsi que 'B' est **Q**-valide, en vertu du théorème principal du système modal **Q** : $\Diamond\phi \equiv \Box\phi$. **Q** a été examiné notamment par Prior dans son *Time and Modality* (1957), dans lequel l'auteur interprétait les modalités en termes de futur et de passé. Dans ce contexte temporel, la propriété **Q** était considérée comme une condition nécessaire en vue de satisfaire la formule de Barcan et l'hypothèse de sempiternalité qui en résulte.

[45] Comment l'auteur justifie-t-il l'indéfendabilité de '$D_a\phi$', i.e. $(D_a\phi \rightarrow \bot)$ dans BD+ ? Celle-ci semble reposer sur l'interprétation contradictoire de 'D' en termes de 'B' : l'interprétation de 'B' par 'est vrai' implique que '$D_a\phi$' signifie 'il est vrai que ni ϕ ni non-ϕ ne sont vraies', laquelle viole la condition (C.¬). Donc si '$D_a\phi$' $\rightarrow \bot$, alors '$\neg D_a\phi$' $\rightarrow (\neg\bot = \top)$.

ANNEXE 2

Logique illocutoire

JOHN. T. KEARNS

DEDUCTION ET COGNITION

1. PSYCHOLOGIE La logique s'intéresse au raisonnement et aux arguments, mais la psychologie également. Cela dit, les psychologues étudient la façon dont les gens raisonnent effectivement, tandis que les logiciens se concentrent sur la façon dont les gens devraient raisonner. Les psychologues aiment concevoir des problèmes de raisonnement pour les sujets, souvent des étudiants de leurs propres universités. Ils aiment déterminer les genres de conclusions que leurs sujets infèreront à partir de différentes sortes d'information, le temps que cela prend, ainsi que découvrir les facteurs qui influencent les conclusions auxquelles les sujets parviennent.

Les psychologues signalent que les gens raisonner souvent de manière incorrecte. Philip Johnson-Laird, par exemple, a construit sa carrière en montrant à quel point les gens raisonnent mal, accompagnée de théories expliquant pourquoi les gens raisonnent comme ils le font. D'après les psychologues, le principe d'inférence du *modus tollens* pose des difficultés pour nous, si bien que nous sommes souvent incapables d'inférer la conclusion lorsque nous avons les prémisses qui conviennent pour ce schéma. Dans [Johnson-Laird 1999], il est déclaré que le désastre de Tchernobyl s'est produit parce que les autorités n'ont pas été capables de faire une simple inférence basée sur le principe du *modus tollens*.

Gottlob Frege, que l'on peut créditer à juste titre d'avoir inventé les théories fondamentales de la logique moderne, faisait une nette distinction entre les considérations psychologiques et les considérations logiques. Il pensait que le *psychologisme* n'avait sa place ni en logique ni en mathématiques. Le philosophe Kant avait injecté à tort des considérations psychologiques dans les mathématiques, en particulier dans l'arithmétique, et Frege entreprit le projet de montrer que les mathématiques sont de caractère logique et non psychologique.[1] Par opposition à la psychologie, la logique se soucie de formuler et démontrer des vérités générales abstraites qui vaudraient dans n'importe quel "monde".

2. ONTOLOGIE Si Frege n'a pas inventé la logique moderne à lui seul, il fait partie du très petit nombre des inventeurs de la logique moderne. Ce fut Aristote qui inventa à lui seul la logique comme discipline intellectuelle. Du temps d'Aristote jusqu'au milieu du dix-neuvième siècle, la logique était conçue assez largement comme contenant des procédures inductives et des arguments inductifs. Cette perspective rapproche la logique de la psychologie, car elle voit la logique traiter des activités humaines et des productions humaines. Mais la logique est une science normative, toutefois, tandis que la psychologie ne l'est pas ou ne l'est pas de la même la manière, tout au moins.

Nous entendons maintenant une théorie logique comme ayant trois composantes : (1) un langage spécialisé ou artificiel, (2) une explication sémantique de ce langage, et (3) un système déductif en vue de la codification de certains éléments logiquement distinctifs dans le langage. Frege ne reconnaissait pas pour sa part l'explication sémantique comme faisant partie de la théorie logique car il se concentrait uniquement sur le langage et le système déductif, utilisant le système déductif pour présenter et développer le langage. En revanche, les travaux de Tarski ont produit une révolution intellectuelle dans la logique moderne et fait de la sémantique une partie légitime et nécessaire d'une théorie logique pleinement développée. Ce supplément à la théorie logique n'a pas changé le caractère fondamental de la logique telle que Frege et d'autres la concevaient,[2] mais il constitua un développement important et influent de la théorie logique.

(2)

Avant la période moderne, avant les travaux de Boole et Frege, la logique avait un caractère principalement épistémologique, ou *épistémique*. Dans la période moderne, le centre de gravité de la logique a été déplacé dans une direction ontologique, ou *ontique*.[3] Frege s'est occupé à développer un langage artificiel qui reflète de façon concise la composition de la réalité, ainsi qu'à établir les lois logiques qui ont un caractère énonciatif [*sentential*] comparable aux lois de la physique. Il développa des systèmes déductifs qui réduisaient ou éliminait les recours à l'intuition, et fournit des démonstrations rigoureuses des lois de la logique. Mais il s'intéressait à l'argumentation correcte, non à l'étude des arguments. Pour Frege, la logique n'est pas plus une étude de la connaissance et de notre manière d'y parvenir que la physique n'est une étude de ces choses.

Lorsque nous édifions un langage artificiel, spécifions les conditions de vérité des énoncés du langage et examinons les conséquences vériconditionnelles des énoncés et ensembles d'énoncés, nous effectuons une étude ontologique. Nous nous intéressons aux choses qu'il y a dans le monde, ainsi qu'aux façons dont des choses sont reflétées et représentées par le langage artificiel. Il est important dans une étude de ce genre de concevoir un système déductif qui soit adéquat pour extraire les conséquences d'un ensemble d'énoncés, et pour déterminer à la fois dans quelle mesure un système déductif particulier est adéquat et dans quelle mesure un système déductif peut l'être en général. Mais l'essentiel porte sur les énoncés et les conditions de vérité, non sur le raisonnement et sa contribution à la connaissance.

Lorsque l'on enseigne la logique aux lycéens, et même aux étudiants des universités, le cours peut avoir des buts épistémiques. Il peut être destiné à permettre aux étudiants de reconnaître et analyser des arguments, à leur permettre d'évaluer ces arguments, ou à leur permettre de produire des arguments pour étayer des conclusions de leur propre cru. On peut faire bon usage des systèmes de déduction naturelle dans un cours de ce genre, car ces systèmes capturent des principes du raisonnement déductif qui sont intuitifs et figurent effectivement dans notre pratique déductive. Ces systèmes peuvent être utilisés pour apporter une compréhension au raisonnement déductivement correct. Mais il y a encore un anicroche entre la théorie logique considérée comme étude de l'ontologie et le but d'une compréhension

des arguments et de leur contribution à la connaissance. Car bien que les étudiants apprennent à reconnaître et à faire des arguments déductivement corrects, ils n'étudient pas des arguments.

3. EPISTEMOLOGIE La dimension ontologique de la logique est un objet légitime de l'étude logique. Le moment où la logique fut reconsidérée pour incorporer l'ontologie constitua une avancée très importante. Mais cette avancée ne doit pas ni ne devrait nous amener à abandonner la dimension épistémologique de la logique. Pour comprendre le rapport entre logique et cognition, il nous faut mettre l'accent sur l'épistémologie plutôt que sur l'ontologie. Mes travaux *sur la logique illocutoire* montrent de quelle façon nous pouvons intégrer à la fois l'ontique et l'épistémique dans une théorie logique.

La logique illocutoire a été inventée et inaugurée en qualité de thème distinctif par John Searle et Daniel Vanderveken. Ce domaine est présenté dans [Vanderveken 1985]. Cela dit, leur compréhension du domaine est quelque peu différente de la mienne, et il n'y a pas tant de recoupement entre les thèmes qu'ils étudient et les résultat que j'ai obtenus. J'expliquerai la logique illocutoire dans ma propre perspective.

(3)

Le plus commode est de présenter et expliquer la logique illocutoire au sein d'un cadre conceptuel des *actes de discours* ou *actes de langage*. J'entends la réalité linguistique fondamentale comme étant constituée par des actes de discours/actes de langage. Un acte de langage est un acte pourvu de sens effectué en utilisant une expression. Une personne peut effectuer un acte de langage en parlant à voix haute, en écrivant ou en pensant par des mots. (Bien que j'utilise les expressions 'acte de langage' et 'acte de discours' de manière interchangeable, le terme 'discours' porte la suggestion que l'on parle à voix haute. Cela fait probablement de 'acte de langage' une expression plus appropriée.) La personne qui lit ou écoute en le comprenant quelqu'un qui parle effectue également des actes de langage.

De mon point de vue, les actes de langage sont les porteurs premiers des caractéristiques sémantiques tels que la signification et la vérité. Les expressions écrites et orales sont les porteurs des caractéristiques syntaxiques, et elles peuvent être considérées elles-mêmes comme des objets syntaxiques. Bien que ce soit les actes de langage qui sont pourvus de sens, diverses expressions sont utilisées conventionnellement pour effectuer des actes dotés de significations particulières ; les significations couramment assignées aux expressions sont les significations d'actes qu'elles ont conventionnellement pour habitude d'effectuer. Les intentions de l'utilisateur du langage déterminent les significations de ses actes de langage ; mais il est traditionnel de viser les significations associées conventionnellement aux expressions que l'on utilise.

La caractéristique sémantique fondamentale d'un acte linguistique est sa *structure sémantique*. Celle-ci est déterminée par les caractères sémantiques des actes qui le composent, ainsi que par leur organisation en vue de former des actes complexes. Cette conception du langage et des actes de langage apporte une distinction nette entre syntaxique et sémantique,

mais elle obscurcit la distinction entre sémantique et pragmatique. La structure sémantique incorporera ce que l'on considère couramment comme des caractéristiques pragmatiques. Par exemple, la force illocutoire sera un élément de la structure sémantique. La structure syntaxique d'une expression pose des contraintes sur la structure sémantique d'un acte de langage effectué avec cette expression, et donc fournit quelques "indications" sur la structure sémantique. Mais une expression complexe unique, par exemple un énoncé, peut souvent être utilisée (à différentes occasions) pour effectuer des actes pourvus de structures sémantiques différentes.

Bien qu'il y ait une énorme diversité d'actes de langage, les actes effectués avec des énoncés, ou *actes énonciatifs*, sont d'une importance particulière en logique. Historiquement parlant, les actes énonciatifs qui sont vrais ou faux ont reçu la plus grande attention : ce sont les *actes propositionnels*. Puisque l'expression 'acte propositionnel' est quelque peu gênante à répéter plusieurs fois, j'appellerai ces actes des affirmations [*statements*]. C'est là un usage stipulé du terme 'affirmation', et il est différent de son usage normal signifiant quelque chose comme l'*assertion*.

Certains actes énonciatifs sont effectués avec une certaine force illocutoire et constituent des *actes illocutoires*. Des exemples sont les promesses, sont avertissements, les assertions, les déclarations, et les requêtes. Les affirmations peuvent elles-mêmes être utilisées avec toute une variété de forces illocutoires. Un *argument* entendu comme un acte de langage a des actes illocutoires en guise de composantes. L'argumentateur passe d'actes de prémisse à un acte de conclusion que ceux-ci sont censés étayer.

(4)

4. UN SYSTEME DE LOGIQUE ILLOCUTOIRE Si nous abordons un système de logique dans une perspective d'actes de discours, un langage logique artificiel n'est pas réellement un langage. En effet, nous ne parlons pas, n'écrivons pas ou ne pensons pas avec des expressions du langage logique. En échange, des expressions du langage logique servent de *représentations* d'actes de langage. Et l'explication sémantique du langage logique constitue en réalité une explication des actes de langage représentée par des énoncés du langage logique. Pour intégrer l'épistémologie dans la théorie logique, nous étendons un système de logique standard pour obtenir un *système illocutoire*. Cela s'accomplit en faisant trois choses :

(i) Le langage artificiel est enrichi d'expressions indicatrices de force illocutoire, ou *opérateurs illocutoires*.

(ii) L'explication sémantique des conditions de vérité est complétée par une explication des *engagements rationnels* engendrés par la réalisation d'actes illocutoires. Asserter ceci ou dénier cela engagera une personne à faire d'autres assertions et dénégations [*denials*]; la même chose vaut pour le fait de supposer des affirmations vraies ou fausses.

(iii) Le système déductif est modifié par une prise en considération des opérateurs illocutoires et de la force illocutoire.

Nous esquisserons un système de logique illocutoire pour le langage L, lequel langage contient des énoncés atomiques et des énoncés complexes obtenus à partir des premiers avec

les connecteurs suivants : ¬, ∨, ∧. ([La flèche] de l'implication matérielle est un symbole défini.) Les énoncés atomiques et complexes sont des *énoncés simples de L*. Les énoncés simples représentent des affirmations des langues naturelles.

Les opérateurs illocutoires sont les suivants :

⊢ le signe de l'assertion	⊣ le signe de la dénégation
⊔ le signe de supposer vrai	⊓ le signe de supposer faux

Un énoncé simple préfixé par un opérateur illocutoire est un *énoncé complet* [completed sentence] *de L* ; il n'y a pas d'autres énoncés complets. Les énoncés complets représentent des actes illocutoires.

Une affirmation peut être acceptée ou rejetée. Une personne effectue un acte lorsqu'elle accepte une affirmation. Dès qu'elle l'accepte, elle continue à accepter l'affirmation tant qu'elle ne change pas d'avis ou qu'elle n'oublie pas qu'elle accepte l'affirmation.[4] Le fait de continuer à accepter une affirmation n'est pas un acte. Une personne qui accepte une affirmation peut effectuer un acte de réaffirmation de l'affirmation ou, comme je préfère le dire, un acte *reflétant son acceptation prolongée de* l'affirmation. Une assertion est entendue ici comme un acte consistant à produire et accepter une affirmation, ou comme un acte consistant à produire et refléter son acceptation de l'affirmation (une assertion de ce genre n'a pas besoin d'un auditoire). Une dénégation est de façon similaire un acte consistant à rejeter une affirmation (comme étant fausse), ou un acte effectué pour refléter son rejet de celle-ci.[5]

Une affirmation peut être supposée vraie ou supposée fausse. Une fois faite, une supposition reste en vigueur tant qu'elle n'est pas déchargée (annulée) ou simplement abandonnée. Un argument qui commence par des assertions et des dénégations peut aboutir à une conclusion qui est une assertion ou une dénégation. Un argument qui **(5)** commence par au moins une supposition ne peut pas conclure (correctement) par une assertion ou une dénégation, tant que la supposition demeure en vigueur. La conclusion aura la force d'une supposition, et nous pouvons l'appeler une supposition.[6]

L'explication sémantique du langage L a deux étages. Le premier étage concerne les affirmations indépendamment de la force illocutoire. Cette explication sémantique donne les conditions de vérité des énoncés simples et des affirmations que ces énoncés représentent. Le premier étage de l'explication sémantique présente l'*ontologie* que les affirmations encodent ou représentent. L'explication des conditions de vérité des énoncés simples de L est entièrement standard. Une *fonction d'interprétation de L* assigne la vérité et la fausseté aux énoncés atomiques simples et elle détermine une *valuation en valeur de vérité* des énoncés simples, dans laquelle les valeurs des énoncés complexes sont déterminés par des tables de vérité courantes.

Le second étage de l'explication sémantique concerne les énoncés complets et les actes illocutoires qu'ils représentent. Dans le cas de L, l'explication sémantique concerne les assertions, les dénégations et les suppositions. Cette sémantique traite de l'*engagement rationnel*. Le terme 'engagement' est utilisé par les philosophes et les logiciens d'un grand nombre de manières différentes. Mais je suis convaincu que l'engagement que j'ai en tête est celui qui figure en fait dans notre pratique inférentielle déductive. Cela dit, puisque tout le monde semble posséder son propre concept d'engagement, je vais expliquer mon concept sans essayer de faire appel à une quelconque signification standard de 'engagement'.

L'engagement en question est l'engagement *rationnel*, par opposition à l'engagement moral ou éthique. C'est un engagement à faire ou ne pas faire quelque chose. Ce peut être également un engagement à prolonger un certain état, comme l'engagement à continuer d'accepter ou rejeter une affirmation donnée. Le fait de décider de faire X rationnellement engage une personne à faire X. Si, avant d'aller au travail, je décide d'acheter du gasoil sur le chemin du travail, je m'engage à le faire. Mais si j'oublie, ou si je change d'avis et continue mon chemin sans acheter d'essence, je n'ai pas fait quelque chose de moralement incorrect. Je peux m'en vouloir d'être oublieux, mais ce n'est pas là une faute morale. Les décisions engendrent des engagements, mais le fait d'effectuer un acte peut engager également une personne à en effectuer d'autres. Par exemple, le fait d'accepter l'affirmation que c'est aujourd'hui vendredi engagera une personne à accepter (ou continuer d'accepter) l'affirmation que c'est demain samedi. (Cet engagement rationnel est assez similaire à ce que Vanderveken appelle un *engagement faible*, dans [Vanderveken 1990]. Il oppose celui-ci à un *engagement fort*, que je considère de peu d'intérêt.)

Certains engagements sont des engagements inconditionnels, tel mon engagement d'acheter de l'essence sur le chemin du travail. D'autres sont conditionnels et n'apparaissent que dans certaines situations, tel l'engagement à fermer les fenêtres du haut s'il pleut pendant que je suis à la maison. Lorsque j'accepte l'affirmation que c'est aujourd'hui vendredi, l'engagement à accepter l'affirmation que c'est demain samedi est conditionnel. Je m'engage à le faire seulement si la question se pose et que je choisis d'y donner une issue. (Et je n'aurai plus l'engagement si je change d'avis au sujet de l'assertion initiale.)

L'engagement rationnel qui nous intéresse ici est un engagement à agir, à se garder d'agir ou à prolonger un état comme le fait d'accepter ou de rejeter une affirmation. Certains auteurs parlent du **(6)** fait de s'engager à l'égard de la vérité d'une affirmation, mais cela n'est pas le type d'engagement présent. Cela étant, une personne pourrait s'engager dans mon sens à reconnaître ou admettre la vérité d'une certaine affirmation.

Pour accepter ou rejeter réellement une affirmation, une personne doit considérer l'affirmation et "prendre position" sur l'affirmation. Les affirmations qu'une personne a réellement acceptées ou rejetées, qui n'a pas oubliées et sur lesquelles elle n'a pas changé d'avis, constituent ses croyances et incroyances *explicites*. Personne ne peut réellement accepter la totalité des affirmations qu'elle s'engage à accepter, ni rejeter la totalité de celles

qu'elle s'engage à rejeter. Personne ne le voudrait. Mais un argument déductivement correct qui commence par des assertions et des dénégations peut amener une personne à étendre la classe des affirmations qu'elle accepte réellement ou celles qu'elle rejette réellement. Un argument de ce genre commence par des croyances et incroyances explicites, et il suit la destination des engagements pour produire des croyances ou incroyances plus explicites.

Un engagement fournit le "pouvoir de motivation" qui propulse quelqu'un des prémisses vers la conclusion d'un argument déductif. Les prémisses et la conclusion sont des actes illocutoires. La personne qui fait ou qui suit un argument doit reconnaître [*recognize*] (ou penser qu'elle reconnaît) un engagement rationnel à effectuer un acte de conclusion. Si par exemple la conclusion est une assertion $\vdash$A, alors si l'argument engage une personne à accepter ou continuer d'accepter A, nous comprendrons en cela que l'argumentateur s'engage à effectuer l'acte $\vdash$A (à effectuer l'acte représenté par l'énoncé complet). Un argument peut être tel que la vérité de ses affirmations de prémisse "exige" la vérité de l'affirmation de conclusion. Mais tant qu'un argumentateur ne reconnaît pas la connexion entre prémisses et conclusion, le fait d'accepter ou de supposer les prémisses ne le conduira pas à accepter ou supposer la conclusion. La récognition du fait que ses actes de prémisse l'engagent est ce qui le pousse à effectuer l'acte de conclusion.

Un engagement à effectuer ou ne pas effectuer un acte est toujours un engagement *d'une certaine personne*. Nous présentons la sémantique d'engagement adaptée à une personne idéalisée, que nous appelons le *sujet désigné*. Ce sujet a des croyances et incroyances qui sont *cohérentes*, au sens où les croyances pourraient être toutes vraies et les incroyances toutes fausses. Le second étage de la sémantique concerne l'*épistémologie* plutôt que l'ontologie, mais l'épistémologie doit s'adapter à l'ontologie. Les engagements engendrés en effectuant certains actes illocutoires dépendent de la compréhension qu'a l'utilisateur du langage des conditions de vérité des affirmations qu'il asserte, dénie ou suppose. Nous considérons le sujet désigné à un moment particulier. Il y a certaines affirmations qu'il a considérés ou acceptés, dont il se souvient et qu'il continue d'accepter. Il y a des affirmations similaires qu'il a considérées et rejetées. Ces croyances et incroyances explicites l'engagent, à ce moment précis, à accepter d'autres affirmations et à rejeter d'autres affirmations. Nous utilisons '+' pour la valeur des assertions et dénégations qu'il s'engage à effectuer à ce moment précis.

Une *valuation d'engagement* est une fonction qui assigne + à certaines des assertions et dénégations de L. Une valuation d'engagement V est *basée sur* une fonction d'interprétation f si et seulement si (dorénavant : ssi) (i) Si V($\vdash$A) = +, alors f(A) = 1, et (ii) Si V($\dashv$A) = +, alors f(A) = 0. Une valuation d'engagement est *cohérente* ssi elle est basée sur une fonction d'interprétation.

(7)

Soit V_0 une valuation d'engagement cohérente. Celle-ci peut être comprise comme répertoriant les croyances et incroyances explicites du sujet désigné à un moment donné. La valuation d'engagement déterminée par V_0 est la fonction V telle que (i) V($\vdash$A) = + ssi A est

vraie pour toute fonction d'interprétation sur laquelle V_0 est basée, et (ii) $V(-|A) = +$ ssi A est fausse pour toute fonction d'interprétation sur laquelle V_0 est basée. La valuation V indique quelles assertions et dénégations le sujet désigné s'engage à effectuer sur la base de ses croyances et incroyances explicites.

Une valuation d'engagement est *acceptable* ssi elle est déterminée par une valuation d'engagement cohérente. Les matrices qui suivent montrent comment des valuations d'engagement acceptables "fonctionnent". Dans les matrices, la lettre 'b' représente la place vide ou blanc pour les positions dans lesquelles aucune valeur n'est assignée :

$\vdash A$	$\vdash B$	$\dashv A$	$\dashv B$	$\vdash \neg A$	$\dashv \neg A$	$\vdash (A \wedge B)$	$\dashv (A \wedge B)$	$\vdash (A \vee B)$	$\dashv (A \vee B)$
+	+	b	b	b	+	+	b	+	b
+	b	b	b	b	+	b	b	+	b
+	b	b	+	b	+	b	+	+	b
b	+	b	b	b	b	b	b	+	b
b	b	b	b	b	b	b	+,b	+,b	b
b	b	b	+	b	b	b	+	b	b
b	+	+	b	+	b	b	+	+	b
b	b	+	b	+	b	b	+	b	b
b	b	+	+	+	b	b	+	b	+

Le sujet désigné peut s'engager à asserter/accepter A et B à la fois (cette combinaison est indiquée dans la première ligne). Si tel est le cas, alors il s'engage par ailleurs à rejeter la négation de A et la négation de B, ainsi qu'à accepter à la fois la conjonction et la disjonction de A et B. Il se pourrait qu'il s'engage à n'accepter qu'une seule des affirmations ; auquel cas il peut s'engager à rejeter l'autre, ou ne s'engager ni à accepter ni à rejeter l'autre. Il peut s'engager à rejeter une affirmation ou les deux à la fois. Il peut ne s'engager dans aucune des deux directions à l'égard de chaque affirmation. Dans certains cas, les valeurs (ou non-valeurs) des assertions et dénégations d'énoncés simples ne sont pas suffisantes pour déterminer les valeurs des assertions et dénégations des énoncés complexes. Par exemple, si $\dashv A$ et $\dashv B$ n'ont pas de valeur, et si A et B sont sans rapport mutuel, alors '$\dashv (A \wedge B)$' ne devrait pas avoir de valeur. Mais si $\dashv A$ et $\dashv \neg A$ n'ont pas de valeur, l'énoncé complet '$\dashv (A \wedge \neg A)$' aura la valeur +.[7]

5. LES CONCEPTS SEMANTIQUES Le langage étendu et le second étage de l'explication sémantique de ce langage fournissent les ressources conceptuelles pour comprendre, expliquer et examiner un grand nombre de phénomènes liés à la logique. Il est courant par exemple de tenter de distinguer les arguments inductifs des arguments déductifs par des considérations ayant trait à la vérité et la probabilité. Mais ceux-ci sont des concepts du premier étage. Pour faire une distinction adéquate entre arguments déductifs et arguments inductifs, nous devons

employer des concepts du second étage. Ce qui caractérise les arguments déductifs, ou les arguments déductifs corrects, c'est qu'il sont basés sur un engagement rationnel. Le fait d'effectuer des actes de prémisse d'un **(8)** argument satisfaisant du point de vue inductif n'engagera pas l'argumentateur à effectuer l'acte de conclusion ; les actes de prémisse ne font que l'*autoriser*, à un degré plus ou moins élevé, à effectuer l'acte de conclusion.

Les conditions de vérité d'une affirmation déterminent à quoi le monde doit ressembler pour que l'affirmation soit vraie. Un grand nombre d'affirmations peuvent être faites de façons différentes. Par exemple, l'affirmation :

Un certain individu (ou un autre) est géologue

peut être rendue vraie par différents individus : pour chaque individu, le fait qu'il soit géologue rendrait l'affirmation vraie. Les conditions de vérité d'une affirmation semblent être considérées de la meilleure des façons comme une caractéristique ontologique ou *ontique* de l'affirmation, si l'ontique est opposé à l'épistémique. Mais les conditions d'engagement sont épistémiques. Ce sont des personnes individuelles qui s'engagent ou non par les affirmations qu'ils acceptent et rejettent. La personne qui fait une affirmation pourvue de sens doit reconnaître les "conséquences engagées" [*engagement consequences*] de son affirmation s'il comprend ce qu'il dit. Tout du moins, il doit reconnaître les conséquences engagées *immédiates*, car personne ne peut passer en revue la totalité des conséquences de portée plus large.[8]

Dans la logique standard, de nombreux concepts sont définis en termes de conditions de vérité. Par exemple, nous pouvons définir l'implication (vériconditionnelle) comme suit : les énoncés simples (de L) $A_1, \ldots,$ An *impliquent* (*vériconditionnellement*) B ssi il n'y a pas de fonction d'interprétation f de L telle que $f(A_1) = f(A_n) = 1$, alors que $f(B) = 0$. (S'il y a une implication reliant des énoncés simples de L, alors il y a une implication reliant les affirmations que ces énoncés représentent).

Soit X un ensemble d'énoncés simples de L et soit A un énoncé simple de L. Alors X *implique* (*vériconditionnellement*) A ssi il n'y a pas de fonction interprétante de L qui assigne 1 à chaque énoncé de X mais n'assigne pas 1 à A.

Soient $A_1, \ldots, A_n$ et B des énoncés simples de L. Alors '$A_1, \ldots, A_n$ / B' est une *suite d'arguments simple de L*. Les énoncés $A_1, \ldots, A_n$ sont les *prémisses* et B la *conclusion*. (Nous considérons également les suites d'arguments dont les membres sont des affirmations. Une suite d'arguments simple de L représentera une suite d'arguments simple dont les membres sont des affirmations du langage naturel.) Une suite d'arguments simple de L est *vériconditionnellement* (*logiquement*) *valide* ssi ses prémisses impliquent vériconditionnellement sa conclusion.

L'*implication illocutoire* relie des énoncés complets de L et les actes illocutoires que ceux-ci représentent. Quelques définitions préliminaires sont requises pour définir l'implication illocutoire.

Soit V_0 une valuation d'engagement cohérente de L, soit F la valuation d'engagement déterminée par V_0, et soit A un énoncé complet de L qui est soit une assertion soit une dénégation. Alors V_0 *satisfait* A ssi V(A) = +.

(9)

Aucune valeur n'est assignée aux suppositions par des valuations d'engagement. Mais le fait de supposer certaines affirmations engagera une personne à en supposer d'autres. Si l'on suppose qu'une affirmation est soit vraie ou fausse, nous prenons en considération des valeurs de vérité pour déterminer quelles autres affirmations nous nous engageons à supposer.

Soit *f* une fonction d'interprétation de L, et soient A et B des énoncés simples de L. Alors (i) *f satisfait* ⊢A ssi *f*(A) = 1, et (ii) *f satisfait* ¬B ssi *f*(B) = 0.

Soient *f* une fonction d'interprétation de L et V une valuation d'engagement de L basée sur *f*. Alors <*f*, V> est une *paire cohérente pour L.*

Soit <*f*, V> une paire cohérente (pour L), et soit A un énoncé complet de L. Alors <*f*, V> *satisfait* A ssi (i) soit A est une assertion ou une dénégation et V satisfait A, (ii) soit A est une supposition et *f* satisfait A.

Soient $A_1, \ldots, A_n$ et B des énoncés complets de L. Alors $A_1, \ldots, A_n$ *exigent logiquement* (*impliquent illocutoirement*) B ssi (i) B est une assertion ou une dénégation et il n'y a pas de valuation d'engagement cohérente qui satisfasse les assertions et dénégations dans $A_1, \ldots, A_n$ mais ne satisfasse pas B, ou (ii) B est une supposition et il n'y a pas de paire cohérente pour L qui satisfasse chaque membre de $A_1, \ldots, A_n$ mais ne satisfasse pas B.

Soit X un ensemble d'énoncés complets de L, et soit A un énoncé complet de L. Alors X *exige logiquement* A ssi (i) A est une assertion ou une dénégation et il n'y a pas de valuation d'engagement cohérente qui satisfasse les assertions et les dénégations de X mais ne satisfasse pas B, ou (ii) B est une supposition et il n'y a pas de paire cohérente pour L qui satisfasse chaque énoncé de X mais ne satisfasse pas A.

Il est nécessaire d'avoir deux conditions dans les définitions de l'implication illocutoire, parce que si B est une assertion ou une dénégation sa valeur est indépendante des valeurs assignées aux suppositions. Considérons par exemple les énoncés complets suivants :

$$\vdash A, \neg A, \vdash B \; ; \vdash (B \wedge A)$$

Il n'y a pas de paire cohérente qui satisfasse ⊢A, ¬A et ⊢B mais ne satisfasse pas '⊢(B ∧ A)', parce qu'il n'y a pas de paire cohérente qui satisfasse ⊢A, ¬A et ⊢B. Cela étant, les

trois premiers énoncés n'exigent pas logiquement '$\vdash(B \wedge A)$', car les suppositions ne font pas de "réclamations" sur les assertions et dénégations. Les suppositions incohérentes exigent logiquement que nous supposions vraies et suppositions fausses chaque énoncé simple, mais elles n'exigent pas que nous assertions ou déniions n'importe quoi.

Soient $A_1, \ldots, A_n$ et B des énoncés complets de L. Alors '$A_1, \ldots, A_n \rightarrow B$' est une *suite d'arguments illocutoire* ; par commodité, je dirai simplement que c'est une *suite illocutoire*. Nous pouvons définir un concept de *validité illocutoire* qui s'applique à des suites illocutoires. Une **(10)** suite illocutoire '$A_1, \ldots, A_n \rightarrow B$' est *logiquement connectée* (*logiquement valide du point de vue illocutoire*) ssi $A_1, \ldots, A_n$ implique logiquement B.

J'utiliserai les termes 'consistant' et 'cohérent' pour des idées sémantiques plutôt que syntaxiques ou preuve-théorétiques. Soit X un ensemble d'énoncés simples de L. Alors X est *consistant* ssi il y a une fonction d'interprétation *f* de L pour laquelle chaque énoncé de X a la valeur 1. (Les énoncés ont la valeur 1 pour la *valuation déterminée par f*).

Soit X un ensemble d'énoncés complets de L. Cet ensemble est *cohérent* ssi il y a une paire cohérente $<f, V>$ pour L qui satisfait chaque énoncé de X.

6. LE SYSTEME DEDUCTIF S C'est un système de déduction naturelle qui emploie trois preuve arborescentes (déductions arborescentes). Chaque étape d'une de ces preuves/déductions est un énoncé complet. Une étape initiale d'une preuve arborescente est une assertion $\vdash A$, une dénégation $\dashv A$, une supposition positive $\llcorner A$ ou une supposition négative $\neg A$. Une assertion ou dénégation initiale n'est pas une hypothèse de la preuve. Une assertion ou dénégation initiale devrait plutôt exprimer une connaissance ou (in)croyance justifiée de l'argumentateur. N'importe quel énoncé $\vdash A$ n'est pas sélectionnable à titre d'assertion initiale dans une preuve construite par une personne donnée. Par opposition, n'importe quelle supposition peut être une supposition initiale. Les suppositions initiales sont des hypothèses de la preuve.

Les règles d'inférence de la conjonction sont *élémentaires* :

Introduction de $\wedge$

$$\frac{\vdash/\llcorner A \qquad \vdash/\llcorner B}{\vdash/\llcorner(A \wedge B)}$$

Elimination de $\wedge$

$$\frac{\vdash/\llcorner(A \wedge B)}{\vdash/\llcorner A} \qquad \frac{\vdash/\llcorner(A \wedge B)}{\vdash/\llcorner B}$$

L'expression '$\vdash/\llcorner$' indique que l'illustration vaut à la fois pour les assertions et les suppositions positives. Pour chaque règle, la conclusion est une assertion seulement si toutes les prémisses sont des assertions. Si au moins une prémisse est une supposition, alors la conclusion doit être une supposition.[9]

Les arguments qui suivent sont incorrects :

$$\frac{\llcorner A \quad \llcorner B}{\vdash (A \wedge B)} \qquad\qquad \frac{\vdash A \quad \llcorner B}{\vdash (A \wedge B)}$$

bien qu'elles soient préservatrices de vérité [*truth-preserving*], car une supposition prise comme prémisse n'étaiera pas une conclusion qui est une assertion. Les arguments qui suivent sont corrects :

(11)

$$\frac{\vdash A \quad \vdash B}{\vdash (A \wedge B)} \qquad \frac{\vdash A \quad \llcorner B}{\llcorner (A \wedge B)} \qquad \frac{\llcorner A \quad \llcorner B}{\llcorner (A \wedge B)}$$

Les règles élémentaires passent directement d'assertions, dénégations ou suppositions prises comme prémisses à une assertion, dénégation ou supposition prise comme conclusion. Les règles non-élémentaires incluent au moins une prémisse qui est une sous-preuve et éliminent, ou déchargent une hypothèse de la sous-preuve. La règle *Introduction de* $\rightarrow$ est une règle non-élémentaire :

$$\frac{\begin{array}{c}\{\llcorner A\} \\ \llcorner B\end{array}}{\vdash / \llcorner (A \rightarrow B)}$$

La prémisse de cette règle est une sous-preuve dans son ensemble, avec '$\llcorner A$' comme hypothèse et '$\llcorner B$' comme conclusion. Suivant un usage de cette règle, l'hypothèse '$\llcorner A$' est éliminée. La conclusion est une assertion si la sous-preuve contient une seule hypothèse '$\llcorner A$' non éliminée. Si la sous-preuve contient des hypothèses supplémentaires, la conclusion est une supposition (positive).

Etant donnés les énoncés A et B, ce qui suit est un exemple d'argument simple du système déductif S :

$$\begin{array}{l} \qquad\quad x \\ \llcorner A \quad \llcorner B \\ \text{----------}I\wedge \\ \llcorner (A \wedge B) \\ \text{----------}E\wedge \\ \quad \llcorner A \\ \text{----------}I\rightarrow, \textit{annuler } \text{'}\llcorner B\text{'} \\ \llcorner (B \rightarrow A) \end{array}$$

Un 'x' est placé au-dessus des hypothèses éliminées. Cet argument montre que la prémisse '$\llcorner$A' exige logiquement la conclusion '$\llcorner$(B → A)'.

Puisque la logique illocutoire s'intéresse à l'épistémologie et aux arguments corrects, de même qu'elle s'intéresse à l'ontologie et aux lois logiques, il est important que les arguments du système déductif soient clairs et que la différence entre les arguments directs et indirects soit nettement indiquée. Dans la perspective de la logique illocutoire, arguments et preuves ne sont pas simplement des instruments pour démontrer des résultats divers ; ce sont également des objets à étudier. Les preuves arborescentes et les opérateurs illocutoires jouent un rôle important dans la réalisation de ce but.

7. LES ASSERTIONS CONDITIONNELLES Une théorie ou système de logique illocutoire revêt un caractère empirique. Il vise à représenter ou *capturer* notre pratique effective de l'utilisation du langage. Il est vrai **(12)** que lorsqu'il s'agit du raisonnement et des arguments, nous nous intéressons à la façon dont les gens *devraient* raisonner plutôt qu'à la façon dont les gens raisonnent *en fait*. Mais la pratique de l'utilisation d'un langage est *normative*, au sens où il y a des normes pour parler correctement et pour construire des arguments corrects. Ces normes sont implicites dans la pratique, en dépit du fait que les gens parlent et raisonnent souvent en violant des normes. Les systèmes de logique illocutoire sont destinés à éclairer et expliquer notre pratique de l'utilisation du langage, et ils doivent être jugés selon qu'ils conviennent ou non à cette pratique. Ce caractère empirique rend la logique illocutoire particulièrement appropriée pour les sciences cognitives, elle montre en quoi la logique peut apporter un éclairage sur la cognition humaine.

Il y a une vaste littérature consacrée à la logique des conditionnels du langage ordinaire. Il y a un accord considérable sur l'idée qu'une affirmation en termes de conditionnel matériel ne revêt pas le même sens qu'un conditionnel du langage ordinaire, et un grand nombre de théories logiques approfondies ont été conçues pour apporter de nouvelles idées sur le mécanisme des conditionnels du langage ordinaire. Mais il n'y a toutefois pas d'explication actuelle qui soit reconnue de façon générale comme la juste explication. Cela est particulièrement surprenant, puisque la totalité des langues naturelles semblent contenir des énoncés conditionnels, et les locuteurs ordinaires de ces langages n'ont aucune difficulté à utiliser et comprendre les formes conditionnelles.

Une partie de la difficulté vient de ce qu'il y a différents types de conditionnels, et il n'est probablement pas possible de fournir une seule explication uniforme et unifiée qui s'accorde à chacun d'entre eux. Une autre partie de la difficulté, plus étendue, vient de ce que la logique standard ne reconnaît pas ni n'inclut les actes illocutoires et la force illocutoire. Il est courant, lors de l'examen des conditionnels, de distinguer les conditionnels indicatifs des conditionnels subjonctifs ou contrefactuels. Cette distinction est effectuée de manière impropre puisqu'elle est basée sur des caractéristiques syntaxiques, alors que la distinction appropriée est sémantique.

Les énoncés conditionnels indicatifs sont typiquement utilisés pour faire des *assertions conditionnelles*, un type d'acte illocutoire particulier, tandis que les énoncés conditionnels subjonctifs sont typiquement utilisés pour faire des affirmations conditionnelles modales. Les explications de ces deux types de conditionnels seront différentes.

Avec les actes de langage qui sont effectués avec des énoncés conditionnels indicatifs, il est difficile de trouver des exemples idiomatiques où ceux-ci soient niés ou déniés. Le conditionnel qui suit est idiomatique et intelligible à la fois :

Si Frank a été payé hier, alors il est sorti dîner la nuit dernière.

mais sa négation n'est ni idiomatique ni intelligible :

Il n'est pas vrai que si Frank a été payé hier alors il est sorti dîner la nuit dernière.

Cet énoncé n'est pas idiomatique, et il serait difficile de comprendre ce que cherche à dire celui qui dirait cela. Mais si le conditionnel du langage ordinaire avait le sens du conditionnel matériel, il ne devrait y avoir aucune difficulté pour comprendre sa négation.

(13)

Pour donner un sens aux conditionnels ordinaires, il nous faut les concevoir comme des actes illocutoires.[10] Les actes illocutoires typiques comportent des énoncés individuels, lesquels sont utilisés avec une certaine force illocutoire. Pour les actes illocutoires que nous avons examinés en rapport aux systèmes de logique, les actes illocutoires typiques comportent des énoncés individuels utilisés pour faire des affirmations individuelles, lesquelles sont équipées d'une force comparable à celle de l'assertion, la dénégation ou la supposition. Mais les assertions conditionnelles combinent deux affirmations ou plus ; dans les cas que nous examinerons, une assertion conditionnelle combine une affirmation comme *antécédent* et une affirmation comme *conséquent*.

Les affirmations à valeur d'antécédent et de conséquent sont combinées pour faire une assertion conditionnelle, mais il n'y a pas d'affirmation conditionnelle. En échange, *le conséquent est asserté sous la condition de l'antécédent*. La notation que j'utiliserai pour l'assertion conditionnelle et la supposition positive conditionnelle est la suivante :

$\vdash(A)/$	pour le fait d'asserter sous la condition A
$\llcorner(A)/$	pour le fait de supposer vrai sous la condition A

Cela donne l'assertion ou supposition conditionnelle de B sous la condition A :

$\vdash(A)/B$ $\llcorner(A)/B$

Les affirmations sont les porteurs premiers de vérité et de fausseté ; l'assertion d'une affirmation vraie peut être qualifiée de vraie dans un sens dérivé. Puisqu'il n'y a pas de

déclarations conditionnelles, les assertions conditionnelles ne sont pas vraies ou fausse même dans un sens dérivé, et cela permet d'expliquer pourquoi nous ne trouvons pas de formes de conditionnel idiomatiques qui soient niées. Cela dit, nous dirons qu'une assertion conditionnelle est *objectivement incorrecte* si l'antécédent est vrai et le conséquent faux ; elle est *objectivement admissible* dans les autres cas.[11]

Pour une assertion conditionnelle, être objectivement admissible n'est pas la même chose que le fait pour une affirmation d'être vraie. L'intérêt ou but de faire une assertion conditionnelle est soit de *démontrer* un engagement conduisant du fait d'accepter ou supposer vrai l'antécédent au fait d'accepter ou supposer vrai le conséquent, soit de refléter cet engagement s'il est déjà en vigueur. La valeur d'importance d'une assertion conditionnelle est relative à l'agent ; le conditionnel est *en vigueur* pour une personne si elle s'engage à asserter ou supposer le conséquent en assertant ou supposant le conséquent. Contrairement au fait d'être en vigueur, la vérité n'est pas une propriété relative à un agent. Une affirmation est vraie pour tout le monde ou pour personne.

Les assertions conditionnelles ne représentent qu'un acte illocutoire conditionnel parmi de nombreux autres. Il y a les assertions, les dénégations et les suppositions conditionnelles. Il y a les promesses et requêtes conditionnelles, ainsi que les avertissements, menaces et excuses conditionnelles ("Si j'ai marché sur tes pieds, je m'en excuse"). La caractéristique essentielle de tous ces actes est l'engagement. L'engagement qui est signalé par une assertion conditionnelle peut être justement décrit comme un principe d'inférence : partant d'une assertion, c'est un engagement à en effectuer une autre. Mais une promesse conditionnelle n'instaure pas ce qui **(14)** serait normalement considéré comme un principe d'inférence. La personne qui promet "Si A, alors je promets de faire X" a instauré un engagement conduisant de l'assertion de A ou de la simple reconnaissance que A est vraie à la réalisation de ce qui constitue en temps normal une action non-linguistique. La promesse instaure une obligation morale ainsi qu'un engagement. Une requête conditionnelle "Si A, alors veuillez faire X" n'établit pas un engagement de la part du locuteur, mais elle constitue une tentative de la part du locuteur de faire en sorte que le destinataire établisse un engagement allant de la reconnaissance que A est vraie à l'acte de faire X.

Si nous reconnaissons que les assertions conditionnelles sont une forme d'acte illocutoire distinct dont l'intérêt consiste à instaurer ou refléter un engagement, nous pouvons comprendre dans quelle mesure l'acte de faire une assertion conditionnelle constitue un cas particulier de réalisation d'un acte illocutoire conditionnel. D'autres traitements proposés pour les assertions conditionnelles ne conviennent pas pour les autres actes conditionnels. Notre explication des assertions conditionnelles permet également de comprendre et d'expliquer pourquoi les énoncés conditionnels utilisés pour faire des assertions conditionnelles ne peuvent pas être dissociés des affirmations ou associés à elles de façon idiomatique.

Les assertions conditionnelles ne sont pas des assertions d'affirmations conditionnelles. Mais la langue anglaise et d'autres langues naturelles contiennent bel et bien des déclarations

conditionnelles. Les conditionnels modaux sont des affirmations, ils sont vrais ou faux. En vertu de leur forme conditionnelle, ils sont utilisés pour instaurer ou refléter des engagements, mais ils servent également à faire une déclaration [*claim*]. Le thème des conditionnels modaux est compliqué, et je n'ai pas grand-chose à dire à ce propos. Les énoncés conditionnels sont utilisés parfois pour faire des affirmations avec le sens d'un conditionnel matériel, souvent dans des rédactions techniques. Mais le conditionnel le plus courant, celui des conversations, c'est l'assertion conditionnelle.

8. LES MODALITES SEMANTIQUES ; LE PARADOXE DE MOORE Les deux niveaux d'un système de logique illocutoire sont les niveaux ontiques et épistémiques. La logique standard, telle qu'on la comprend communément, constitue le premier niveau ou premier étage de la logique illocutoire. La sémantique associée au niveau ontique traite des conditions de vérité, avec des caractéristiques telles que la vérité logique, l'implication et la validité logique, lesquelles peuvent être expliquées en termes de conditions de vérité. La sémantique associée au niveau épistémique traite de l'engagement et des conditions d'engagement, avec des caractéristiques qui peuvent être expliquées en ces termes. Au niveau épistémique, nous nous intéressons par exemple à ce que signifie pour certains actes illocutoires d'en exiger logiquement d'autres, ou pour une suite d'arguments illocutoire d'être logiquement connectée.

Il y a des concepts distinctifs de nécessité et de possibilité associés à chaque niveau de la théorie logique. Une affirmation est *ontiquement nécessaire* si ses conditions de vérité ne peuvent pas ne pas être satisfaites. La nécessité ontique est la *vérité analytique*. La question de savoir si une affirmation est ontiquement nécessaire, ou analytique, dépend de ce que l'on pourrait appeler la "signification totale" de l'affirmation. La *nécessité logique*, ou *vérité logique*, est déterminée par la *forme logique* de l'affirmation, telle que celle-ci est révélée par un système d'analyse logique ou un autre. La nécessité logique est un cas particulier de la nécessité ontique. Une affirmation est *ontiquement possible* si elle n'est pas contradictoire, elle est *logiquement possible* s'il y a une interprétation d'un langage logique telle que l'affirmation est vraie. Une version logique illocutoire du système modal S5 constitue le système approprié pour examiner la vérité analytique et la vérité logique.

(15)

La *nécessité épistémique* est relative à une personne, ou une communauté, et à la connaissance de cette personne ou de cette communauté à un moment donné. La chose la plus commode est de développer un système illocutoire de logique modale épistémique dans la perspective du sujet désigné. Une affirmation est épistémiquement nécessaire à un moment donné si elle résulte (au sens de l'engagement) de la connaissance du sujet désigné à ce moment précis. Et une affirmation est *épistémiquement possible* à un moment si elle n'est pas exclue par la connaissance du sujet désigné à ce moment précis.

Il y a certains concepts sémantiques des deux niveaux qui sont analogues les uns aux autres. Au niveau ontique, le niveau des conditions de vérité, nous reconnaissons une *implication* qui relie un certain nombre d'affirmations à un autre affirmation, ou une

implication d'un ensemble d'affirmations à une autre affirmation. L'implication est une relation de préservation de la vérité. Nous reconnaissons également un concept sémantique de *consistance*. Un certain nombre d'affirmations sont consistantes les unes avec les autres, ou un ensemble d'affirmations est consistant si toutes les affirmations peuvent être satisfaites ensemble. Des concepts analogues au niveau épistémique sont l'implication illocutoire, en vertu de laquelle certains actes illocutoires ou un ensemble de ces actes *exigent logiquement* un autre acte illocutoire. Ainsi que le concept de *cohérence*. Des actes illocutoires (ou des énoncés complets des langages logiques) sont cohérents si le fait d'effectuer les actes n'engage pas une personne à accepter et rejeter à la fois une affirmation particulière, ni à supposer vraie et supposer fausse à la fois une affirmation particulière.

Dans les systèmes illocutoires de la logique modale épistémique, nous trouvons des cas intéressants où ces concepts analogues divergent. Nous limiterons notre attention aux contextes où les assertions faites par le sujet désigné ont le statut de déclarations de connaissance. Et nous comprendrons une affirmation modale '□A' comme signifiant *A résulte de la connaissance actuelle du sujet désigné*, ou *A est épistémiquement nécessaire à ce moment précis*. Avec cette compréhension, il est clair que les principes d'inférence qui suivent sont corrects :

Introduction de □	*Elimination de* □
⊢A	⊢□A
---	---
⊢□A	⊢A

Pour la supposition positive, nous avons également un principe *Elimination de* □:

⌞□A

⌞A

Si nous supposons que A résulte d'une connaissance actuelle, nous supposons alors que A est vraie. Mais le principe *Introduction de* □ n'est pas correct pour la supposition. Supposer que A est vraie n'est pas supposer que A résulte d'une connaissance actuelle. En temps normal, lorsqu'une inférence de '⊢A' à '⊢B' est correcte c'est parce que A implique B. Ce n'est pas le cas avec *Introduction de* ⊢. L'inférence est manifestement correcte, mais l'affirmation A n'implique pas que A résulte d'une connaissance actuelle.

(16)

Le philosophe anglais G. E. Moore fut troublé par l'affirmation "Il pleut, mais je ne le crois pas". Il reconnaissait que l'affirmation est consistante, mais il se rendait compte également qu'il serait incorrect qu'une personne accepte ou asserte cette affirmation. Son problème était qu'il ne savait pas comment caractériser l'angle sous lequel l'assertion est

incorrecte. Son embarras, ainsi que le problème d'expliquer ce qui ne va pas au juste avec l'assertion, finiront par être appelés le *Paradoxe de Moore*.

Dans le contexte présent, il sera commode de reformuler le problème de Moore en termes de connaissance. L'"essence" du problème concerne une affirmation comme celle-ci :

$$(A \wedge \neg\Box A)$$

Un grand nombre d'affirmations de cette forme sont vraies. Ni le sujet désigné, ni aucun d'entre nous ne sait tout. Mais tandis que l'affirmation ci-dessus est consistante, l'assertion suivante :

$$\vdash(A \wedge \neg\Box A)$$

est incorrecte pour le sujet désigné. Les arguments qui suivent montrent pourquoi cela est incorrect :

$$\begin{array}{l} \vdash(A \wedge \neg\Box A) \\ \text{----------} \; E\wedge \\ \quad \vdash A \\ \quad \text{--------} \; I\Box \\ \quad \vdash \Box A \end{array} \qquad \begin{array}{l} \vdash(A \wedge \neg\Box A) \\ \text{----------} \; E\wedge \\ \quad \vdash \neg\Box A \end{array}$$

Bien que l'affirmation concernée soit consistante, son assertion est incohérente. Il y a réellement quelque chose de troublant à ce sujet. Le problème de Moore est qu'il connaissait les concepts appropriés au niveau ontique de la logique, mais il n'avait pas conscience de la logique illocutoire, du niveau épistémique examiné dans les systèmes de logique illocutoire. Il manquait simplement à Moore les concepts qu'il lui fallait pour caractériser la situation qu'il envisageait.

8. PARADOXES ET AUTRES ENIGMES Le caractère singulier de la logique illocutoire concerne la dimension épistémique qu'elle apporte à la logique, complétant la dimension ontique traitée par les systèmes de logique standards. Celle-ci donne une meilleure compréhension de la contribution que le raisonnement déductif apporte à la cognition, et fournit de nouvelles ressources permettant à des systèmes logiques de capturer la pratique linguistique effective. En supplément, la perspective élargie de la logique permet de comprendre et de résoudre certaines énigmes concernent le langage et le raisonnement déductif, ainsi que de surmonter certains inconvénients de la logique standard.

J'ai déjà examiné le Paradoxe de Moore, énigme qui est résolue dès que nous introduisons des concepts logiques épistémiques pour compléter les concepts basés exclusivement sur la vérité et les **(17)** conditions de vérité. Notre examen montre que certaines affirmations vraies ne peuvent pas être acceptées de manière cohérente par certaines personnes. Une affirmation :

Il pleut mais je ne le crois pas.

est consistante, et elle pourrait être vraie du sujet désigné. Mais celui-ci n'est pas "sélectionnable" pour accepter cette affirmation. S'il l'accepte, son acceptation et sa croyance deviendraient incohérentes. (Mais aucune personne normale n'accepterait une telle affirmation à son sujet, car cela serait un acte auto-annulant [*self-defeating*]).

J'ai utilisé à un autre endroit ([Kearns 2000]) les ressources de la logique illocutoire pour apporter une explication au paradoxe de l'exécution surprise. Le prisonnier qui est condamné à l'exécution par surprise, prévue pour un des jours de la semaine à venir, fournit un autre exemple d'une personne dont les croyances sont incohérentes. Ses croyances étayent la conclusion qu'il sera exécuté et qu'il ne le sera pas. Mais des affirmations qui sont incohérentes à accepter pour le prisonnier ne sont pas incohérentes pour nous autres. Nous pouvons croire à la fois que le prisonnier sera exécuté et que cela sera une surprise sans nous engager à accepter des affirmations contradictoires, alors que le prisonnier ne le peut pas.

J'ai un autre article dans lequel j'utilise les ressources de la logique illocutoire pour analyser le Paradoxe du Menteur, montrant que notre pratique linguistique ordinaire ne nous engage pas à accepter des affirmations consistantes. Alors que nous ne pouvons pas accepter une Affirmation du Menteur de façon cohérente, nous pouvons (et devrions) la dénier/rejeter de façon cohérente. Même le Théorème d'Incomplétude de Gödel perd une part de son importance lorsqu'il touche des systèmes de logique illocutoire. Il s'avère en effet que le Théorème ne s'applique pas à certains types de système illocutoire, pas même à des systèmes illocutoire d'ordre supérieur de ces types.

9. REFERENCE Le thème de la référence a été important en logique, tout au moins depuis les travaux de Frege et Russell et même si ni Frege ni Russell n'ont utilisé le terme 'référer' comme un terme technique pour un type d'acte de langage. Cela dit, les deux hommes se sont intéressés à notre utilisation du langage pour "atteindre" les choses du monde. Dans "Sens et Dénotation", Frege déclare que les sens des noms propres et des descriptions définies apportent des *modes d'accès* à leurs référents, tandis que Russell croyait que ce sont les expressions qu'il appelait *noms logiquement propres* qui relient directement nos affirmations aux objets de notre accointance.

Il y a diverses énigmes associées aux termes singuliers et aux affirmations faites par leur biais. L'énigme initiale est peut-être celle relevée par Frege, lequel voulait comprendre pourquoi certaines affirmations d'identité semblent triviales alors que d'autres sont informatives ; et ce bien que toutes les affirmations d'identité vraies disent simplement qu'une chose est elle-même. Une autre énigme consiste à expliquer comment un terme singulier descriptif peut être utilisé parfois pour identifier un objet qui ne satisfait pas la description (de même que l'on pourrait utiliser 'l'homme avec un martini' pour une personne qui ne boit pas un martini). Ou quelle est la différence entre les expressions fournissant l'accès direct à un

objet, qui a été caractérisé sous l'appellation de désignation rigide, et les expressions qui désignent de façon non rigide ?

(18)

Des théories diverses ont été proposées pour expliquer le mécanisme des noms, des descriptions définies, des démonstratifs, des indexicaux et d'autres termes singuliers. Certaines de ces théories ont été étendues pour recouvrir les noms communs et les adjectifs. Des théories logiques de grande complexité ont été conçues pour expliquer une pratique qui ne semble ni mystérieuse ni particulièrement compliquée pour les utilisateurs du langage ordinaire.

Aucune explication parfaitement réussie ou satisfaisante n'a été offerte pour rendre compte des usages des termes singuliers. Il n'y a sans doute pas de théorie logique qui apporte davantage de perspectives. Cela est dû en partie à la compréhension standard de la logique, laquelle ne s'adapte pas comme il faut aux dimensions à la fois ontiques et épistémiques de la logique. La plupart des termes singuliers ont au moins deux usages distincts ; les deux usages que je veux examiner sont l'usage *référentiel* et l'usage *prédicatif*. Pour certains termes singuliers, l'usage référentiel est premier, et d'autres termes singuliers ne sont utilisés que pour référer. Je n'ai pas connaissance de termes singuliers qui soient utilisés seulement de manière prédicative. Dans une théorie logique, l'usage prédicatif des termes singuliers est le mieux compris et expliqué au niveau ontique de la théorie logique. L'usage référentiel d'un terme singulier est épistémique. Chaque personne, lorsqu'elle fait référence, exploite des caractéristiques qui sont propres à elle-même. Bien que différentes personnes puissent asserter la même affirmation, elles ne peuvent pas faire la même assertion. L'assertion de Jones engage Jones mais pas Smith, tandis que l'assertion de Smith joue un rôle similaire pour Smith. Et lorsque Jones fait référence à quelqu'un, par exemple Napoléon, il exploite une connexion qui le relie à Napoléon en dirigeant son attention sur Napoléon. Smith peut faire référence également à Napoléon, mais sa connexion à Napoléon est différente de celle de Jones.

Un terme singulier peut être prédiqué d'un objet. Si je dis "George Bush est le président des Etats Unis", il se peut qu'en utilisant le nom je dirige mon attention sur la personne nommée. Lorsque je prononce la description 'le président des Etats Unis', il serait inhabituel de rediriger mon attention sur le même individu. Il est bien plus probable que je prédique la description de la personne sur laquelle je porte (déjà) mon attention. De façon similaire, si je pointe en direction de quelqu'un et que je dis "C'est Daniel Baker", j'utiliserai le démonstratif 'Ce' pour diriger mon attention (ou pour exprimer mon attention) sur une personne particulière, puis j'utiliserai le nom pour caractériser cette personne comme étant appelée ou nommée "Daniel Baker". Je peux utiliser un terme singulier de façon prédicative pour un objet particulier, mais ce peut être également un usage prédicatif lorsque je dis que Hilary Clinton n'est pas le président des Etats Unis. Je ne cherche pas à dire qu'elle est différente de George Bush, je cherche à dire qu'elle ne satisfait pas les critères en vue d'être président. Bien que cela soit peu courant, un terme singulier peut être utilisé de façon prédicative comme sujet d'un énoncé. Si je dis "La plus grande femme sur Terre doit mesurer au moins

sept pieds", je n'exploite pas ici une connexion pour diriger mon attention sur la dame en question. Je vise la plus grande femme du monde qui qu'elle soit qui mesure sept pieds.

Dans "Sur la Dénotation", Russell présenta son explication de l'usage prédicatif des descriptions définies. Il pensait que les noms propres ordinaires sont utilisés également de façon prédicative. Pour Russell seuls les noms logiquement propres peuvent être utilisés pour référer, dans notre sens de 'référer'. Dans les termes d'une théorie logique, Russell ne concevait les descriptions comme des termes singuliers ; ce sont des prédicats combinés à un opérateur de description liant des variables. Frege ne semble pas avoir reconnu la distinction entre les usages prédicatifs et référentiels des termes singuliers, car il considérait les noms et **(19)** les descriptions comme les mêmes types d'expression élémentaires et fut probablement enclin à utiliser les constantes d'individu d'un langage logique pour l'un ou l'autre type de terme singulier. Dire que le sens d'un nom fournit un mode d'accès à l'objet nommé rappelle la façon dont Frege pense que les noms ont l'habitude de référer, mais en admettant un type de sens uniforme à la fois pour les noms et pour les descriptions Frege semble se concentrer sur l'usage prédicatif des noms.

Le logicien polonais Lesniewski formula un système appelé *Ontologie*, en guise de système logique élémentaire. Le langage de l'*Ontologie* ne contient pas de catégorie de termes singuliers, ou de constantes d'individu. Il concevait les termes singuliers comme des termes généraux qui dénotent un objet unique. Cela me semble un signe clair du fait qu'il admettait l'usage prédicatif des termes singuliers mais pas l'usage référentiel. Et W. V. Quine propose d'éliminer des constantes d'individu des langages de la logique. Il conçoit les termes singuliers ordinaires comme des prédicats, capturant l'usage prédicatif mais pas l'usage référentiel de ces termes singuliers. Même John Searle, qui a largement écrit sur les actes de discours et les actes illocutoires et qui a présenté et défendu vigoureusement ce qu'il appelle une théorie de la référence, donne une explication des termes singuliers qui ne convient que pour leur usage prédicatif.

Il est probable que de nombreux facteurs jouent un rôle dans l'explication de ce qui semble être une "préférence" répandue pour l'usage prédicatif des termes singuliers par opposition à leur usage référentiel ou, dans le cas de Frege, son incapacité à distinguer les deux types d'usage. Cela étant, un facteur précis est sans doute l'incapacité des systèmes de logique standards à concilier à la fois l'ontique et l'épistémique. La dimension épistémique de la logique est prise à la légère et ignorée, et les théoriciens n'ont pas conscience des ressources qui peuvent être déployées pour traiter des caractéristiques comme la référence et la force illocutoire.

10. L'ACCES DIRECT Une personne peut diriger son attention sur un objet qui est présent, un objet dont il a actuellement conscience. Une personne peut également porter son attention sur ou songer à un objet qui n'est pas présent à lui. Ce pourrait être un objet dont il a eu l'expérience dans le passé et dont il se souvient encore, ou ce pourrait être un objet dont il n'a pas eu l'expérience, un objet dont il a seulement entendu parler ou sur lequel il a lu quelque

chose. (S'il porte son attention sur un objet dont il a eu l'expérience mais dont il ne se souvient pas, cela se fait probablement sur la base de ce qu'il a entendu dire ou de ce qu'il a lu de l'objet).

L'attention que je vise ici dote une personne d'un accès *direct* à l'objet de son attention. L'accès direct peut être soit *immédiat* soit *médiat*. L'accès immédiat est toujours direct, il implique la conscience d'un objet sans aucun intermédiaire. La conscience d'un objet dont on a présentement l'expérience constitue au moins une bonne approximation de la conscience immédiate. Si le fait de voir un objet "à l'œil nu" fournit une conscience immédiate de l'objet, alors le fait de voir l'objet en portant des lunettes fournit une conscience médiate. L'expérience de l'objet est produite par la médiation des lunettes, mais la conscience est dirigée sur l'objet, pas sur les verres.

Lorsqu'une personne aveugle utilise une canne pour explorer les objets autour d'elle, son accès et sa conscience des objets explorés sont directs mais médiats tant que son attention est dirigée sur les objets et non sur sa main, ou sa canne. L'accès *indirect* à un objet est toujours médiat. Pour accéder indirectement à B, on doit accéder directement à un objet A et utiliser consciemment les caractéristiques de A pour déterminer à quoi B ressemble. Utiliser des empreintes laissées sur le sol pour déterminer le type d'animal qui est **(20)** passé par là, pour évaluer son poids, etc., c'est accéder indirectement à cet animal. Cela étant, ce qui fournit un accès indirect à un objet pour une personne pourrait fournir un accès direct (mais médiat) pour quelqu'un d'autre. Il est possible qu'un traqueur très expérimenté puisse porter son attention sur un animal par le biais des traces que l'animal laisse.

Pour porter l'attention sur un objet directement mais de façon médiate, il doit y avoir une certaine connexion que la personne prêtant attention *exploite*. L'agent qui porte son attention n'est pas obligé d'en avoir conscience ou d'être au courant de cette connexion. On peut porter son directement attention sur un objet qui est passé et n'existe plus, de même que sur un objet qui existe présentement. On peut porter directement son attention sur un objet futur. Lors d'une attention directe mais médiate portée sur un objet, aucune expérience ou signe infaillible n'est exigée concernant l'existence d'un objet. La différence entre l'attention directe et médiate sur un objet et la tentative de porter son attention sur un objet non existant se résume à l'existence d'un objet obtenu par la connexion que la personne exploite. Une personne qui pense à tort que Sherlock Holmes a été un véritable détective peut essayer de porter son attention sur lui et penser qu'elle porte son attention sur lui, mais sa tentative ne sera pas couronnée de succès.

La connexion qui est exploitée lors d'une attention directe mais médiate portée sur un objet est un *mode d'accès* qui relie la personne portant l'attention à l'objet de son attention. Elle peut accéder à un objet dont elle a l'expérience maintenant ou dont elle a eu l'expérience (et se souvient encore). Elle peut accéder à un objet sur la base de ce qu'elle a entendu ou lu. Mais les modes d'accès de chaque personne commencent avec elle et la relient aux objets. Deux personnes n'exploiteront pas les mêmes modes d'accès lorsqu'ils portent leur attention

sur les mêmes objets. D'ordinaire, une personne capable d'exploiter une certaine connexion pour porter son attention sur un objet aura plus d'une connexion exploitable pour porter son attention sur cet objet. Si ce sont là des connexions différentes dont elle sait qu'elles aboutissent au même objet, alors nous pouvons considérer cette personne comme "rassemblant" ces connexions et les exploitant toutes lorsqu'elle porte son attention sur leur objet commun.

11. LES MODES D'ACCES Nous utilisons souvent une image ou une expression pour porter directement son attention sur un objet absent. Un *acte référentiel* est un acte consistant à utiliser une expression comparable à un nom ou une locution descriptive pour porter directement son attention sur un objet, ou pour exprimer son attention si cet objet est déjà l'objet d'attention. Une personne ne peut faire référence à un objet que s'il y a un objet, et la personne a besoin d'un mode d'accès, une connexion qu'elle soit en mesure d'exploiter et qui la relie à l'objet. (Nous sommes tous connectés à des objets par des moyens dont nous n'avons pas connaissance ; nous ne pouvons pas exploiter ces connexions pour porter notre attention sur les objets). Nous pouvons référer à des objets de l'expérience, une expérience soit présente soit passée (mais dont on se souvient). Nous pouvons référer à des objets dont nous avons entendu parler ou sur lesquels nous avons lu quelque chose. Si nos sources d'information exploitent leurs propres connexions pour porter notre attention sur les objets dont elles parlent, alors nos connexions aux objets incorporent les connexions de nos sources –dans la terminologie dédaigneuse de Searle, nos connexions sont parasites par rapport aux connexions de nos sources.

Nous pouvons utiliser également notre connaissance du monde pour faire référence à davantage d'objets que ces derniers. Je peux référer à la personne qui a commis un crime, même si je ne sais pas qui il ou elle est. Etant donnée ma connaissance limitée de la géographie et ma connaissance de la structure politique des Etats Unis, je peux faire référence au gouverneur de l'Idaho, même si je n'ai pas connaissance d'avoir lu quoi que ce soit au sujet de ce **(21)** gouverneur. Nous ne pouvons pas faire référence à un objet tant qu'il n'*existe* pas un objet auquel faire référence, même si nous pouvons tenter de porter notre attention sur un objet qui n'existe pas et n'a jamais existé dans le cas où nous penserions qu'un tel objet existe. Nous ne pouvons pas faire référence tant que nous ne *pensons* pas qu'il existe un objet auquel référer. Lorsqu'il réfère à un objet, l'utilisateur du langage doit accepter que le référent soit réel, tout au moins dans le contexte du discours –l'utilisateur du langage doit s'engager à accepter une affirmation disant que le référent existe. Bien que Russell ne pense apparemment pas dans "Sur la Dénotation" que les descriptions définies et les noms propres ordinaires conviennent pour faire référence à des individus, les deux types de termes singuliers *peuvent* être utilisés de cette manière et sont fréquemment utilisés de cette manière.

Diverses théories de la référence ont admis des modes d'accès qui relient l'utilisateur du langage à l'objet, même si ces théories ne parlent pas toutes de modes d'accès. Frege pensait que le sens d'un nom ou description fournit un tel accès, et les chaînes causales de Kripke (qui sont en réalité des chaînes intentionnelles, si Searle dit vrai) semblent fournir un accès

pour les noms. Et Russell comprenait manifestement l'accointance comme fournissant un mode d'accès aux objets des noms logiquement propres. L'idée que les utilisateurs d'un langage exploitent des modes d'accès pour faire référence explique comment nous utilisons le langage pour "accéder aux" objets du monde.

Il est clair que le seul fait de dire (écrire, penser) un nom ou description n'est pas suffisant pour diriger l'attention de l'utilisateur d'un langage sur le référent. Nous pensons que les perroquets n'utilisent pas les noms qu'ils prononcent pour faire référence à des objets. Il est clair également que nous pouvons diriger notre attention sur un objet sans utiliser une expression. Le fait de voir un objet devant vous fournit un mode d'accès à cet objet. La mémoire de l'objet qui a fait l'objet d'une expérience connecte également une personne à cet objet. Mais qu'est-ce que les noms et les descriptions fournissent pour relier l'utilisateur d'un langage à un objet ? Certaines descriptions identifient le lien. Pour une personne qui peut diriger directement son attention sur l'Arizona et qui comprend le [français], l'expression 'la capitale de l'Arizona' lui donne les ressources pour concentrer son attention sur la capitale de l'Arizona. Et pour certains noms, et certaines personnes, l'explication de Frege peut expliquer comment une personne utilise un nom pour accéder à un objet. Une personne qui associe une description à un nom, où la description identifie un lien à l'objet, peut exploiter ce lien lorsqu'il porte son attention sur l'objet. Mais il arrive également qu'une personne apprenne l'existence d'un objet par quelqu'un d'autre, lequel a lui-même appris l'existence de l'objet par une autre personne, etc., retournant à l'objet via un grand nombre de personnes. Cela constitue également un lien qui peut être exploité. Différentes expressions peuvent être associées à différents liens à un objet, lorsque l'utilisateur d'un langage n'a pas conscience qu'il n'y a pas un objet unique. Et une expression unique peut être associée à plusieurs liens à un objet (à un *faisceau* de liens). Lorsqu'il y a une erreur, et que les liens ne connectent pas tous une personne à un objet unique, il y a divers principes que nous pouvons employer pour déterminer l'objet de référence correct.

Il se pourrait qu'une personne sache d'une expression qu'elle est le nom d'un objet sans avoir aucune connaissance du lien qu'elle peut exploiter pour faire référence à l'objet. Il se pourrait également qu'un terme singulier descriptif ne parvienne pas à identifier un lien que l'utilisateur d'un langage peut exploiter. La description 'le roi de France' échoue parce qu'elle n'identifie pas un lien aboutissant à un objet. Ne serait-ce que dans mon cas, l'expression 'la plus grande femme au monde' échoue parce qu'elle ne localise pas quelqu'un qui serait cette personne. Je comprends ce que c'est qu'être la plus grande femme au monde, mais elle n'est pas située dans une structure appropriée (une structure que je connaisse) par rapport aux choses sur lesquelles je peux porter mon attention.

(22)

Les actes de référence exploitent des modes d'accès à des objets du monde de façon à ce que nous puissions dire et penser des choses concernant ces objets. Russell a trop exigé des modes d'accès, car il voulait que toutes les tentatives de référence réussissent. Il pensait que seule l'accointance, dans son sens fort, pourrait apporter la garantie qu'il recherchait. Bien qu'il ait développé une position plus ou moins fregéenne en ce qui concerne les sens des

termes singuliers, John Searle exige trop peu des modes d'accès, ce qui est surprenant. Il pense que ce n'est que dans certains cas que les actes référentiels relient les utilisateurs d'un langage à des objets du monde. Ces liens ont peu d'importance, car tout ce que nous avons à connaître pour faire référence c'est une propriété ou des propriétés qui caractérisent le référent de manière unique. Searle n'évalue pas à sa juste valeur l'importance de l'acte de référence pour relier nos affirmations aux objets que nous visons.

C'est parce que nous associons souvent plus d'un mode d'accès à un terme singulier que nous pouvons utiliser à l'occasion une description en vue de référer à un objet qui ne satisfait pas la description. Cela ne devrait pas arriver, en vertu de la position de Searle, car ce sont les noms et non les descriptions qui sont censées être associées à un faisceau d'expressions descriptives. La description est simplement une expression descriptive. La personne qui utilise une description pour faire référence à un objet qui ne satisfait pas la description *a* commis une erreur ; lorsque son erreur n'est pas simplement due à ce que sa langue a fourché, elle pense que le référent a une propriété qu'il n'a pas. Si son acte de référence réussit malgré cela, ce doit être parce que la connexion qu'elle exploite pour atteindre l'objet visé l'emporte sur son erreur.

Pour qu'une personne fasse référence à un objet, il doit y avoir un objet, elle doit savoir quelque chose de l'objet de façon à lui permettre de diriger son attention sur l'objet, et elle doit s'engager à admettre que l'objet existe. Il y a une chose telle que l'acte de référence fictionnel, ou l'acte de référer à des objets fictionnels, mais cela est un effet parasite par rapport à l'acte de référence ordinaire, lequel exige des objets réels. Les objets fictionnels ont un certain statut, ou permanent, et c'est ce qui rend possible l'acte de référence fictionnel. Or je m'intéresse seulement à l'acte de référer au monde réel. Celui qui réfère à un objet révèle son engagement préalable à admettre que l'objet existe, l'acte de référence n'engendre pas cet engagement. Lors de la référence à un objet et de la caractérisation de l'objet quelle qu'elle soit, c'est l'objet sur lequel l'attention est portée qui détermine si la caractérisation est vraie ou fausse. Quand bien même une description serait utilisée pour faire référence à un objet qui ne satisfait pas la description, la vérité ou fausseté de la caractérisation dépend de l'objet d'attention et non de la description.

12. UN TRAITEMENT LOGIQUE DE LA REFERENCE Dans les langues naturelles, les noms et les termes singuliers descriptifs peuvent être utilisés de façon soit prédicative soit référentielle. La portée prédicative d'un nom propre provient de son usage référentiel. Il se peut que la portée prédicative des termes descriptifs soit première. Dans la langue anglaise, en tout cas, il n'y a pas de procédé syntaxique utilisé pour distingue l'usage référentiel et l'usage prédicatif d'un terme singulier. Puisque cette distinction est importante sémantiquement parlant, elle devrait être consacrée dans un langage logique destiné à apporter un éclairage sur notre pratique linguistique ordinaire. Je propose d'utiliser des constantes d'individu simples pour représenter l'usage prédicatif d'un terme singulier, et des constantes d'individu soulignées pour représenter des actes référentiels.
(23)

Pour un langage quantificationnel ou un langage du premier ordre, les énoncés contenant seulement des constantes d'individu simples peuvent être interprétées par des moyens standards dans des domaines d'individus non vides. Pour inclure les énoncés contenant des constantes d'individu soulignées, il est nécessaire de considérer le niveau épistémique de la théorie sémantique et de développer une explication pour une personne particulière –pour notre *sujet désigné*. La sémantique doit inclure aussi bien des modes d'accès que des domaines d'individus, et des valuations d'engagement interpréteront des constantes d'individu soulignées de préférence à des fonctions d'interprétation.

Je n'ai ni le temps ni l'espace ici pour développer les détails formels d'un traitement logique adéquat de notre utilisation des termes singuliers. La seule chose que je veux faire remarquer à cet endroit, c'est qu'une explication adéquate des expressions référentielles et des actes référentiels appartient au niveau épistémique de la logique et non au niveau ontique. Dès que la logique est étendue pour inclure l'épistémique, il devient relativement facile de concevoir une explication simple et intuitive qui concilie les usages référentiels et non référentiels des termes singuliers.

13. RESUME La logique est un sujet académique et un domaine de recherche très ancien. Mais il y a un grand nombre de choses nouvelles qui se produisent en logique, un grand nombre de nouveaux thèmes et de nouveaux domaines pour la recherche logique. La logique illocutoire, la logique des actes de discours ou actes de langage, en est un. La logique illocutoire inclut ou incorpore la logique standard, elle apporte les ressources pour intégrer le souci traditionnel de la logique pour l'épistémologie dans la théorie logique moderne. Cette conception plus adéquate de la logique fournit une perspective qui nous permet de résoudre ou dissoudre certains problèmes de longue date et d'effectuer une recherche qui éclaire nos pratiques linguistiques et cognitives. J'espère pouvoir encourager d'autres étudiants et élèves à étudier ce domaine relativement inexploré de la logique.

RÉFÉRENCES

Philip N. Johnson-Laird, Vittorio Girotto, Paolo Legrenzi. 1999. "Modelli mentali: una guida facile per il profano", *Sistemi Intelligenti* **9**, 63-83 ("Mental models : a gentle guide for outsiders")

G. Frege. 1960. "On Sense and Reference", in *Translations from the Philosophical Writings of Gottlob Frege*, P. Geach and M. Black eds. Oxford, Basil Blackwell, 56-78.

John T. Kearns. 1997. "Propositional Logic of Supposition and Assertion", *Notre Dame Journal of Formal Logic* **38**, 325-349.

_____. 2000. "An Illocutionary Logical Explanation of the Surprise Execution", *History and Philosophy of Logic* **20**, 195-214.

_____. 2004. "An Illocutionary Analysis of Conditional Assertions", *The Logica Yearbook 2003*, L. Běhounek ed., Filosofia, Prague, 159-169.

_____. 2005. "Russell's Epistemic Understanding of Logic", *Teorema* **24**, 115-131.

_____. 2006. "Conditional Assertion, Denial, and Supposition as Illocutionary Acts", à paraître in *Linguistics and Philosophy*.

G. E. Moore. 1944. 'Moore's Paradox' in T. Baldwin, ed., *Selected Writings*. London: Routledge, 1993, 207-212.

B. Russell. 1905. "On Denoting", *Mind* **14**, 479-493.

John R. Searle. 1969. *Speech Acts: An Essay in the Philosophy of Language*. London: Cambridge: Cambridge University Press.

_____. 1985. *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge: Cambridge University Press.

Daniel Vanderveken et John Searle 1985. *Foundations of Illocutionary Logic*. Cambridge: Cambridge University Press.

Daniel Vanderveken. 1990. *Meaning and Speech Acts*. Cambridge. Cambridge University Press.

NOTES DU TRADUCTEUR

[1] La remarque de Kearns semble assimiler comme beaucoup d'autres Kant et le psychologisme. Il serait plus prudent de distinguer entre deux formes d'intuition, dont une seule renvoie au psychologisme : l'intuition *empirique*, distincte de l'intuition pure. L'analyse kantienne du jugement arithmétique était-elle proprement psychologiste, alors qu'elle évoquait une synthèse a priori dans l'opération de dénombrement ?

[2] Kearns fait allusion ici à la tradition des logiciens qui excluaient la sémantique de leur théorie logique, que [van Heijenoort 1969] caractérise comme les partisans de la logique comme langage universel. Voir la section 2.1.3.2.

[3] L'expression 'ontique' est employée dans la thèse pour évoquer le processus d'interprétation des expressions du langage, puisque celui-ci établit un lien entre le sujet interprétant et le monde. Contrairement à la distinction de Kearns ci-dessus entre épistémologique et ontologique *ou ontique*, nous traitons cette dernière expression comme distincte de l'ontologique et servant de compromis entre les aspects ontologiques et épistémologiques de

la logique : elle est l'acte de considérer (aspect épistémologique) un énoncé comme vrai (aspect ontologique).

[4] On peut établir un parallèle ici avec la clause de consistance (C.¬) d'un ensemble de formules dans [Hintikka 1962], qui fait de la logique épistémique et illocutoire des systèmes *synchroniques* ; ainsi qu'avec la définition de K dans [Lemmon 59], selon lequel savoir que p équivaut à l'avoir appris et ne pas l'avoir *oublié*.

[5] 'Comme étant fausse' sous-entend que Kearns assimile la dénégation au dissentiment A¬p, contrairement à notre propre définition de la dénégation.

[6] Ce passage confirme l'analogie entre quantificateurs ∀/∃, modalités □/◇ et forces illocutoires A/C : dans la théorie syllogistique des énoncés catégoriques, d'une existentielle (I ou E) aucune universelle ne peut être déduite ; dans la syllogistique modale, d'une possibilité aléthique (Mp) aucune conclusion nécessaire (Lp) ne peut être déduite ; d'une supposition (Cp) aucune assertion (Ap) ne peut être déduite. Dans une théorie des valuations, cette impossibilité de conclure la modalité forte d'une modalité faible se traduit en disant que les valeurs faibles "contaminent" les valeurs fortes : une formule conjonctive de type $(\phi \wedge \psi)$ prend la valeur la plus faible, soit $v(\phi \wedge \psi) = min[(v(\phi), v(\psi)]$.

[7] La matrice présentée ici est donc *non-vérifonctionnelle*, puisque *f* n'assigne pas une et une seule valeur déterminée à une formule complexe en fonction de la valeur des atomes qui la composent. L'allusion au rapport mutuel entre A et ¬A permet par ailleurs à Kearns de maintenir la validité de tautologies classiques au sein de son langage formel illocutoire : $\vdash(A \vee \neg A)$ et $\dashv(A \wedge \neg A)$ auront la valeur +. Cette solution est comparable à la méthode de supervaluation de van Fraassen, puisqu'elle assigne la valeur + aux formules de ce type quelle que soit la valeur de A.

[8] Le critère d'immédiateté de la relation de conséquence rappelle les tautologies de surface : une relation trop distante entre un acte illocutoire et sa conséquence logique ne sera pas forcément assumé par le sujet désigné.

[9] Voir la note 6 ci-dessus, sur le lien entre cette phrase et la règle syllogistique selon laquelle l'extension (ou modalité) de la conclusion n'est jamais plus grande (ou forte) que celle des prémisses. En termes algébriques, elle signifie également que la conclusion prend la valeur de la prémisse la plus faible : $v(A \wedge B) = min[v(A), v(B)]$.

[10] Le passage qui suit propose une explication sémantique du conditionnel en termes d'actes illocutoires ; cette approche permet d'éviter les paradoxes de l'implication matérielle : la relation conditionnelle entre A et B ne vaut que si le locuteur s'est engagé au préalable sur A, et le cas où il ne s'engage pas sur A annule la nature conditionnelle de la déclaration.

[11] Le cas problématique, celui des paradoxes de l'implication, est celui où l'antécédent est faux et n'est pas répertorié ici comme un acte conditionnel à part entière. C'est ce qu'explique Kearns en disant ci-dessous que le conditionnel exprime un but ou intérêt et que celui-ci suppose au moins l'antécédent vérifié.

Bibliographie

[Aho 1994]: "On the Philosophy of Attitude Logic", *Acta Philosophica Fennica* **57**
[Åqvist 1964]: "A solution to Moore's paradox", *Philosophical Studies* **15**, pp. 1-5
[Austin 1962]: *How to do Things with Words*, Oxford, Clarendon Press
[Barcan Marcus 1983]: "Rationality and believing the impossible", *The Journal of Philosophy* **80**(6), pp. 321-38
[Arner 1959]: "On Knowing", *Philosophical Review* **68**, pp. 84-92
[Barnes 1969]: "The law of contradiction", *Philosophical Quarterly* **19**, pp. 302-309
[Barwise & Perry 1983]: *Situations and Attitudes*, Cambridge, MA – London : The MIT Press
[Bäuerle & Cresswell 2003]: "Propositional attitudes", in *Handbook of Philosophical Logic* (2e édition), Vol. 10, pp. 491-512
[Beall 2000]: "Is the observable world consistent?", *Australasian Journal of Philosophy* **78**, pp. 113-118
[Bertman 1975]: "Non-extensional propositions in Wittgenstein", *International Logic Review* **5**, pp. 73-77
[Beyer 2001]: "A Neo-Husserlian theory of speaker's reference", *Erkenntnis* **54**, pp. 277-297
[Black 1952]: "Saying and disbelieving", *Analysis* **13**, pp. 25-33
[Blanché 1968]: *Introduction à la logique contemporaine*, Collection U, Armand Colin
[Blanché 1974]: "Modalité et temporalité", *International Logic Review* **9**(1), pp. 103-10
[Bochvar 1938] : "On a three-valued calculus and its application to analysis of paradoxes of classical exteded functional calculus", *Matématičéskij Sbornik* **4**, pp. 287-308
[Bovens 1995]: "'P and I will believe that not-P': diachronic constraints on rational belief", *Mind* **104**, pp. 737-60
[Brogaard & Salerno 2004]: "Fitch's paradox of knowability", in *Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/fitch-paradox/
[Bull & Segerberg 1984]: "Basic modal logic", in D. M. Gabbay & F. Guenthner (eds.), *Handbook of Philosophical Logic, 2nd Edition* (Volume 3), pp. 1-81
[Cargile 1967]: "On believing you believe", *Analysis* **27**, pp. 177-83
[Cargile 1970]: "A note on 'iterated knowings'", *Analysis* 30, pp. 151-5
[Carnap 1937]: *The Logical Syntax of Language* (version originale: *Logische Syntax der Sprache*, Springer, 1934), Routledge & P. Kegan (ed.), Londres
[Carnap 1946]: "Modalities and quantification", *The Journal of Symbolic Logic* **11**, pp. 33-64
[Carnap 1947]: *Meaning and Necessity* (*A Study in Semantics and Modal Logic*), University of Chicago Press
[Castañeda 1964]: "Review of Hintikka's 'Knowledge and Belief'", *Journal of Symbolic Logic* **29**(1964), pp. 132-4
[Castañeda 1970]: "On knowing (or believing) that one knows (or believes)", *Synthese* **21**, pp. 187-203
[Chalmers 2004]: "The Foundations of Two-Dimensional Semantics", in M. Garcia-Caprintero & J. Macia (eds.), *Two-Dimensional Semantics: Foundations and Applications*, Oxford University Press
[Chisholm 1963b]: "The Logic of Knowing", *Journal of Philosophy* **60**, pp. 773-95
[Chisholm 1967]: "Identity through possible worlds : some questions", *Noûs* **1**, pp. 1-8
[Chisholm 1982]: *Theory of Knowledge*, XXX
[Church 1946]: "A Formulation of the Logic of Sense and Denotation", in *Structure, Method and Meaning*, pp. 3-24
[Church 1951]: "The Need for Abstract Entities in Semantic Analysis", *Proceedings of the American Academy of Arts and Sciences* **80**(1) (1951) : 100-12.

[Church 1954]: "Intensional isomorphism and identity of belief", *Philosophical Studies* **5**, pp. 65-73
[Copeland 2002] : "The genesis of possible world semantics", *Journal of Philosophical Logic* **31**(2002), pp. 99-137
[Cresswell 1970]: "Classical intensional logics", *Theoria* **36**, pp. 347-372
[da Costa & Bueno 1996] : "Consistency, Paraconsistency and Truth (Logic, the Whole Logic and Nothing but Logic)", *Ideas y Valores* **100**, pp. 48-60
[da Costa & French 1989]: "On the logic of belief", *Philosophy and Phenomenological Research*, **49**, pp. 431-46
[da Costa & French 1990]: "Belief, contradiction and the logic of self-deception", *American Philosophical Quarterly* **27**, pp. 179-97
[da Costa & French 1995]: "Partial structures and the logic of the Azande", *American Philosophical Quarterly* **32**, pp. 325-39
[da Costa & Bueno & French 1998]: "Is there a Zande logic ?", *History and Philosophy of Logic* **19**, pp. 41-54
[Davidson 1973a] : "On the very idea of a conceptual scheme", *Proceedings of the American Philosophical Association* **47**, 5-20
[Davidson 1973b]: "On radical interpretation", *Dialectica* **27**, pp. 313-28
[de Sousa 1970]: "Knowledge, consistent belief, and self-consciousness", *The Journal of Philosophy* **67**, pp. 66-73
[Donnellan 1966]: "Reference and definite descriptions", *Philosophical Review* **75**, pp. 281-304
[Deutscher 1969]: "Hintikka's conception of epistemic logic", *Australasian Journal of Philosophy* **47**(2), pp. 205-8
[Dretske 1970]: "Epistemic operators", *The Journal of Philosophy* **67**, pp. 1007-23
[Dubucs 1991]: "On logical omniscience", *Logique et Analyse* **133-134**, pp. 41-55
[Dubucs 1998]: "Hintikka et la question de l'omniscience logique", in E. Rigal (éd.) : *Jaakko Hintikka, Questions de logique et de phénoménologie*, Librairie Philosophique J. Vrin, Paris, pp. 140-148
[Dubucs 2002]: "Feasibility in logic", *Synthese* **132**, pp. 213-237
[Dubucs & Marion 2003]: "Radical antirealism and substructural logics", in A. Rojszczak & J. Cachro & G. Kurczewski (eds.), *Philosophical Dimensions of Science. Selected Contributed Papers from the 11th International Congress of Logic, Methodology, and the Philosophy of Science, Krakow, 1999*, Kluwer, pp. 235-249
[Dummett 1973a]: *Frege, Philosophy of Language*, London, Duckworth
[Dummett 1973b]: "The justification of deduction", *Proceedings of the British Academy* **59**, pp. 201-32
[Eberle 1974]: "A logic of believing, knowing, and inferring", *Synthese* **26**, pp. 356-82
[Ellis 1976]: "Epistemic foundations of logic", *Journal of Philosophical Logic* **5**, pp. 187-204
[Engel 1989a]: *La norme du vrai* (*Philosophie de la logique*), NRF Essais, Gallimard
[Engel 1989b]: "Interprétation et mentalité prélogique (Quine, Davidson et la charité bien ordonnée)", *Revue Philosophique* **4**(1989), pp. 543-558
[Engel 1996] : *Philosophie et psychologie*, Folio Essais
[Engel 1998] : "Believing, holding true, and accepting", *Philosophical Explorations* **1**(2), pp. 140-51
[Englebretsen 1969]: "Knowledge, negation, and incompatibility", *Journal of Philosophy* **66**, pp. 580-585
[Englebretsen 1974]: "Epistemic logic and mere belief", *Logique et Analyse* **16**, pp. 375-8
[Fagin & Halpern 1988], "Belief, awareness and limited reasoning", *Artificial Intelligence* **34**, pp. 39-76

[Ferret 1998]: *L'identité*, Corpus, GF Flammarion, Paris
[Fisher 1964]: "Remarks on a logical theory of belief statements", *The Philosophical Quarterly* **14**, pp. 165-169
[Fitch 1963] : "A logical analysis of some value concepts", *Journal of Symbolic Logic* **28**, pp. 135-142
[Føllesdal 1986]: "Quantified modal logic and essentialism", in *Mérites et limites des méthodes logiques en philosophie*. Colloque organisé par J. Vuillemin, Fondation Singer-Polignac, Librairie Philosophie J. Vrin, Paris, pp. 169-182
[Føllesdal 1998a]: "Essentialism and reference", *The Philosophy of W. V. Quine* (2e éd.), Ed. Schilpp, pp. 97-113
[Føllesdal 2004]: *Referential Opacity and Modal Logic*, R. Nozick (éd.), Routledge
[Frege 1999] : *Idéographie*, Textes Philosophiques, trad. Corine Besson, Paris
[Frege 1892]: "Sens et dénotation", *Zeitschrift für Philosophie und philosophische Kritik* **100**, trad. in [Frege 1971].
[Frege 1919a]: "La pensée", *Beitrage zur Philosophie der deutschen Idealismus* **1**, pp. 58-77; trad. dans [Frege 1971]
[Frege 1919b]: "La négation", *Beitrage zur Philosophie der deutschen Idealismus* **1**, pp. 143-157 ; trad. dans [Frege 1971]
[Frege 1971]: *Ecrits logiques et philosophiques*, trad. Claude Imbert Editions du Deuil, Paris
[Frege 1999]: *Idéographie*, trad. Corinne Besson, postface Jonathan Barnes, Vrin, Paris
[French 1990] : "Rationality, consistency and truth", *The Journal of Non-Classical Logic* **7**, pp. 117-137
[Gardenförs 1980]: "Review of Brian Ellis' 'Rational belief systems'", *Synthese* **44**, pp. 161-169
[Gardenförs 1988]: *Knowledge in Flux: Modeling the Dynamics of Epistemic States*, MIT Press, Cambridge.
[Gardiès 1979] : *Essai sur la logique des modalités*, PUF, Paris
[Gettier 1963]: "Is justified true belief knowledge ?", *Analysis* **23**, pp. 121-123
[Girle 1974]: "Epistemic logic, language and concepts", *Logique et Analyse* **16**, pp. 359-73
[Gochet 1972]: "Sur la sémantique de la logique épistémique (Frege, Quine, Hintikka)", *International Logic Review* **5**, pp. 50-63
[Gochet 1978] : *Quine en perspective*, Flammarion, Paris.
[Gochet & Gribomont 2003]: "Epistemic Logic", in *Handbook of the History and Philosophy of Logic*, Gabbay D. M. & Woods J. (eds.), Amsterdam : Elsevier Science
[Gochet & Gribomont & Thayse 2000]: *Logique (Méthodes pour l'Intelligence Artificielle)*, Hermès Science Publications, Paris
[Gödel 1933]: "Eine Interpretation des intuitionistischen Aussagenkalküls", *Ergebnisse eines mathematischen Kolloquiums* **4**, pp. 34-40
[Granger 1979]: *Langage et Epistémologie*, Horizons du Langage, Paris.
[Greimann 2000]: "The judgement-stroke as a truth-operator : a new interpretation of the logical form of sentences in Frege's scientific language", *Erkenntnis* **52**(3), pp. 213-38
[Haack 1978]: *Philosophy of logics*, Cambridge University Press
[Hanson 1952]: "A note on statements of fact", *Analysis* **13**, p. 24
[Harnish 1980]: "Searle and the logic of Moore's paradox", *International Logic Review* **21**, pp. 72-6
[Hart 1980]: "Toward a logic of doubt", *International Logic Review* **21**, pp. 31-41
[Heal 1994]: "Moore's paradox: a Wittgensteinian approach", *Mind* **103**, pp. 5-24
[Henderson 1987]: "The principle of charity and the problem of irrationality (translation and the problem of irrationality)", *Synthese* **73**, pp. 225-52

[Hendricks 2003]: *Forcing epistemology*, Center for the Philosophy of Language and Information, Roskilde University
[Hendricks & Jørgensen & Pedersen 2003]: *Knowledge Contributors*, Kluwer Academic Publishers Vol. 322, Dordrecht
[Hilpinen 1988]: "The semantics of questions and the theory of inquiry", *Logique et Analyse* **29**, pp. 523-539
[Hintikka 1953a]: "Distributive Normal Forms in the Calculus of Predicates", *Acta Philosophica Fennica* **6**
[Hintikka 1957]: "Modality as referential multiplicity", *Ajatus* **20**, pp. 49-64
[Hintikka 1962]: *Knowledge and Belief*, Ithaca Press, New York
[Hintikka 1965]: "Are logical truths analytic?", *The Philosophical Review* **74**, pp. 178-203
[Hintikka 1967]: "Individuals, possible worlds, and epistemic logic", *Noûs* **1**, pp. 33-62
[Hintikka 1968]: "Epistemic logic and the methods of philosophical analysis", *Australasian Journal of Philosophy* **46**, pp. 37-51 ; réimprimé in *Models for Modalities*, D. Reidel, Dordrecht Holland (1969), pp. 3-19
[Hintikka 1969a]: *Models for Modalities*, D. Reidel, Dordrecht Holland
[Hintikka 1969b]: "Une sémantique adaptée aux attitudes propositionnelles", in *L'intentionnalité et les mondes possibles* (1989), pp. 87-111; extrait de "Semantics for propositional attitudes", in *Philosophical Logic*, J. W. Davis *et al.* (eds.), D. Reidel, Dordrecht Holland, pp. 21-45
[Hintikka 1970a]: "'Knowing that one knows' reviewed", *Synthese* **21**, pp. 141-62
[Hintikka 1970b]: "Knowledge, belief, and logical consequence", *Ajatus* **32**, pp. 32-47
[Hintikka 1975a]: "Impossible possible worlds vindicated", *Journal of Philosophical Logic* **4**, 475-484. Réimprimé in Saarinen E. (éds), *Game-Theoretical Semantics*, D. Reidel Publ. Co., Dordrecht, Holland (1979), pp. 367-79
[Hintikka 1976]: "The semantics of questions and the questions of semantics : Case studies in the interrelations of logic, semantics, and syntax", *Acta Philosophica Fennica* **28**, 200 pp.
[Hintikka 1979]: "Frege's hidden semantics", *Revue Internationale de Philosophie* **33**, pp. 716-722
[Hintikka 1981a]: "Phenomenology vs. possible-worlds semantics : apparent and real differences", *Revue Internationale de Philosophie* **35**, pp. 113-119
[Hintikka 1984] : "Are there nonexistent objects ? Why not ? But where are they ?", *Synthese* **60**(3), pp. 451-458
[Hintikka 1985]: "*Cogito ergo sum*: inférence ou performance?", *Philosophie* **6**, pp. 21-51; traduit par P. Le Quellec-Wolff de J. Hintikka: *Knowledge and the Known* (*Historical Perspectives in Epistemology*), pp. 98-122. Imprimé également in "*Cogito ergo sum*: inference or performance?", *Philosophical Review* **71**(1962), pp. 3-32
[Hintikka 1986]: "Quine on Who's Who'", in *The Philosophy of W. V. Quine*, Hahn, Lewis Edwin (ed.), La Salle Open Court (1986), pp. 209-228
[Hintikka 1989]: *L'intentionnalité et les mondes possibles*, trad. par N. Lavand, Presses Universitaires de Lille
[Hintikka 1997a]: *Lingua Universalis vs. Calculus Ratiocinator : An Ultimate Presupposition of Twentieth-Century Philosophy*, Kluwer Academic Publishers , Dordrecht : Holland
[Hintikka 1997b]: "Three dogmas of Quine's empiricism", *Revue Internationale de Philosophie* **4**(202), pp. 457-477
[Hintikka 1997c]: "Problems of Philosophy. Problem #3 : one-world assumption and Frege's sense-reference distinction", *Synthese* **112**(3), pp. 431-432
[Hintikka 1999]: "Quine's ultimate presuppositions", *Theoria* **65**, pp. 3-24
[Hintikka 2001]: "Intuitionistic logic as epistemic logic", *Synthese* **127**, pp. 7-19.

[Hintikka 2003]: "A second generation epistemic logic and its general significance", in *Knowledge Contributors*, Hendricks Vincent, Jørgensen Klaus Frovin and Pedersen Stig Andur (eds.), Kluwer Academic Publishers Vol. 322, Dordrecht, pp. 33-55
[Hintikka & Sandu 1995]: "The fallacies of the New Theory of Reference", *Synthese* **104**, pp. 245-283
[Hinzen 2000]: "Anti-realist semantics", *Erkenntnis* **52**, pp. 281-311
[Hocutt 1972]: "Is epistemic logic possible ?", *Notre Dame Journal of Formal Logic* **13**, pp. 433-53
[Jespersen 1933], *Essentials of English Grammar*, London
[Kanger 1957a]: "The Morning Star Paradox", *Theoria* **23**, pp. 1-11
[Kanger 1957b]: "A Note on Quantification and Modalities", *Theoria* **23**, pp. 133-4
[Kanger 1957c]: "On the Characterisation of Modalities", *Theoria* **23**, pp. 152-5
[Kearns 1997]: "A propositional logic of supposition and assertion", *Notre Dame Journal of Formal Logic* **38**, pp. 325-349
[Kearns 2000]: "An illocutionary logical explanation of the surprise execution", *History and Philosophy of Logic* **20**, pp. 195-213
[Kearns 2004]: "An enlarged conception of the subject matter of logic", *Ideas y Valores* **126**, pp. 57-74
[Kearns 2006]: "Deduction and cognition".
[Kielkopf 1971]: "Provability as a deontic notion", *Theory and Decision* **2**, pp. 1-15
[Kielkopf 1972]: "A note on Hintikka's logic of belief as an ethics of belief", *Philosophical Studies* **23**, pp. 135-7
[Klima 1964]: "The negation in English", in *The Structure of Language: Readings in the Philosophy of Language*, J. A. Fodor and J. J. Katz (eds.), Englewoods Cliffs, N. J., pp. 246-323
[Kripke 1963]: "Semantical considerations on modal logic", *Acta Philosophica Fennica* **16**, pp. 83-94
[Kripke 1979]: "A puzzle about belief", in *Meaning and Use* (ed.) A. Margalit, Dordrecht : Reidel, pp. 239-83
[Kripke 1971]: "Identity and necessity", in M. Munitz (ed.), *Identity and Individuation*, New York Univ. Press, N-Y, pp. 66-101
[Largeault 1993]: *La Logique*, in "Que sais-je?" Volume 225, PUF, Paris
[Laugier & Bonnay 2003]: "La logique sauvage de Quine à Lévi-Strauss", *Archives de Philosophie* **66**(1), pp. 49-72
[Lehrer 1968]: "Belief and knowledge", *Philosophical Review* **77**, pp. 491-99
[Lehrer & Paxson 1969]: "Knowledge : undefeated justified true belief", *The Journal of Philosophy* **66**, pp. 225-37
[Lemmon 1957]: "New foundations for Lewis modal systems", *The Journal of Symbolic Logic* **22**(2), pp. 176-186.
[Lemmon 1959]: "Is there only one correct system of modal logic ?", *Proceedings of the Aristotelian Society* (*Supplementary Volume*)**33**, pp. 23-40
[Lemmon 1967]: "If I know do I know that I know ?", in *Epistemology*, Avrum Stroll (ed.), Harper and Row, New York
[Lemmon 1977]: *An Introduction to Modal Logic*, in *American Philosophical Quarterly* (monographie), K. Segerberg (éd.), Oxford.
[Lenzen 1978]: "Recent works in epistemic logic", *Acta Philosophica Fennica* **30**, pp. 1-210
[Lenzen 1979]: "Epistemische Betrachtungen zu [S4, S5]", *Erkenntnis* **14**, pp. 33-56
[Lenzen 1980a]: *Glauben, Wissen und Wahrschenlichkeit*, Springer Verlag, Wien
[Lenzen 1980b]: "Wittgensteins Zweifel über Wissen und Gewissheit", *Grazer Philosophische Studien* **10**, pp. 43-52

[Lenzen 2003]: "Knowledge, belief, and subjective probability : outlines of a unified system of epistemic/doxastic logic", in Hendricks Vincent, Jørgensen Klaus Frovin and Pedersen Stig Andur (eds.), *Knowledge Contributors*, Kluwer Academic Publishers Vol. 322, Dordrecht, pp. 17-31

[Lenzen 2004]: "Epistemic logic", in I. Niiniluoto & M. Sintonen & J. Wolenski (eds.), *Handbook of Epistemology*, Kluwer Academic Publishers, Netherlands, pp. 963-83

[Lewis 1973]: "Possible worlds", in *The Possible and the Actual* (*Readings in the Metaphysics of Modality*), Michael J. Loux (eds.), Cornell University Press, Ithaca and London, pp. 182-9 ; extrait de *Counterfactuals* (Cambridge, Mass. : Harvard University Press, 1973), pp. 84-91

[Lewis 1981]: "What puzzling about Pierre does not believe ?", *Australasian Journal of Philosophy* **59**(3), pp. 283-89

[Linsky 1968]: "On interpreting doxastic logic", *The Journal of Philosophy* **65**, pp. 500-2

[Lokhorst 1998]: "La logique du relativisme logique", *Logique et analyse* (N.S.), **161-162-163**, pp. 57-65

[Łoś 1948] : "Logiki wielowartościowe a formalizacja funkcji intensjonalnych (Many-valued logics and the formalization of intensional functions)", *Kwartalnik Filozoficzny* **17**, pp. 59-78

Steven LUPER [2001]: "The epistemic closure principle", in *Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/closure-epistemic

[Łukasiewicz 2000]: *Du principe de contradiction chez Aristote*, trad. Dorota Sikora, préface de Roger Pouivet, L'Eclat, Paris.

[Malcolm 1952]: "Knowledge and Belief", *Mind* N.S. **61**, pp. 178-89

[Malinowski 1993]: *Many-Valued Logics*, Clarendon Press, Oxford.

[Martin-Löf 1996]: "On the meanings on the logical constants and the justifications of the logical laws", *Nordic Journal of Philosophical Logic* **1**, pp. 11-60.

[Mates 1954]: "Synonymity", in L. Linsky (éd.), *Semantics and the Philosophy of Language*, University of Illinois Press.

[Meyer 2001]: "Epistemic logic", in *Handbook of Philosophical Logic*, ???

[Mitton 1972]: "Professor Hintikka on Descartes' 'cogito'", *Mind* **81**, pp. 407-408

[Mohanty 1979]: "Aspects of Husserl's philosophy of logic a propos his Logic Lectures of 1906-7", in *Logic, Truth and the Modalities* (*from a phenomenological perspective*), Synthese Library Vol. 278, pp. 44-55

[Mohanty 1981]: "Intentionality and possible worlds", *Revue Internationale de Philosophie* **35**, pp. 91-112

[Montague 1974]: *Formal Philosophy*, R. Thomason (éd.), Yale Univ. Press, New Haven.

[Morgan 1973]: "Systems of modal logic for impossible worlds", *Inquiry* **16**, pp. 280-89

[Morick 1975]: "Intentionality without intensionality : reply to Lithown and Marras", *Philosophical Studies* **28**, pp. 143-46

[McGinn 1977]: "Charity, interpretation, and belief", *The Journal of Philosophy* **74**, pp. 521-35

[McLane 1979]: "On the possibility of epistemic logic", *Notre Dame Journal of Formal Logic* **20**, pp. 559-74

[Nef 1991]: "L'intentionnalité et les mondes possibles", in *L'Age de la Science n°4, Philosophie de la logique et philosophie du langage I*, Bouveresse J. (éd.), Odile Jacob, Paris, pp. 193-201

[Niiniluoto 1979]: "Knowing that one sees", in E. Saarinen & R. Hilpinen & I. Niiniluoto & M. P. Hintikka (eds.) : *Essays in Honour of Jaakko Hintikka*, D. Reidel Publishing Company, Dordrecht : Holland, pp. 249-84

[Nolte 1986]: "What are possible worlds ?", *Mind* **95**, pp. 432-445

[Nozick 1981]: *Philosophical Explanations*, Cambridge University Press

[Paoli 2003]: "Quine and Slater on paraconsistency and deviance", *Journal of Philosophical Logic* **32**(5), pp. 531-48
[Pap 1957]: "Belief and propositions", *Philosophy of Science* **24**, pp. 123-136
[Priest 1985]: "Contradiction, belief and rationality", *Proceedings of the Aristotelian Society* **86**, pp. 99-116
[Priest 1992]: "What is a non-normal world ?", *Logique et Analyse* **35**, pp. 291-302
[Priest 1999]: "Perceiving contradictions", *Australasian Journal of Philosophy* **77**, pp. 439-46
[Prior 2002]: *Objets de pensée*, Mathesis, Coll. Vrin, Paris
[Pritchard 2001]: "The opacity of knowledge", in *Essays in Philosophy* (*A Biannual Journal*), Vol. 2(1): 'The Internalism/Externalism Debate in Epistemology', B. Shaeffer (ed.), pp. 1-10
[Purtill 1968]: "About identity through possible worlds", *Noûs* **2**, pp. 87-89
[Putnam 1969]: "Is logic empirical ?", *Boston Studies in the Philosophy of Science* **5**, pp. 216-241
[Quine 1943]: "Notes on existence and necessity", *The Journal of Philosophy* **40**(5), pp. 113-27
[Quine 1947]: "The problem of interpreting modal logic", *The Journal of Symbolic Logic* **12**(2), pp. 43-8
[Quine 1951a]: "Ontology and ideology", *Philosophical Studies* **2**(1), pp. 11-5
[Quine 1951b]: "On Carnap's views on ontology", réimprimé in[Quine 1997], pp. 203-11
[Quine 1953]: "Three grades of modal involvement", in [Quine 1997], pp. 158-76
[Quine 1954]: "Carnap and logical truth", in [Quine 1997], pp. 107-32
[Quine 1956]: "Quantifiers and Propositional Attitudes", in [Quine 1997], pp. 185-96
[Quine 1960] : *Word and Object*, MIT Press.
[Quine 1961] : *From a Logical Point of View*, Harvard University Press (2e éd.)
[Quine 1961a]: "On what there is", in [Quine 1961], pp. 1-19
[Quine 1961b]: "Two dogmas of empiricism" in [Quine 1961], pp. 20-46
[Quine 1961c]: "Notes on the theory of reference", in [Quine 1961], pp. 130-38
[Quine 1961d]: "Reference and Modality", in [Quine 1961], pp. 139-59
[Quine 1964]: "Les frontières de la théorie logique", *Les Etudes Philosophiques* N.S. **2**, pp. 191-208
[Quine 1968]: "Objets propositionnels", *Critica* **2**(5), pp. 3-22 ; réimprimé in *Relativité de l'Ontologie et autres essais*, Paris, Aubier-Montaigne, 1977, pp. 157-177 ; édition originale : *Ontological Relativity and other essays*, 1969.
[Quine 1974]: The Roots of Reference,
[Quine 1975]: *Philosophie de la Logique*, Aubier-Montaigne ; édition originale : *Philosophy of Logic*, Englewood Cliffs, Prentice Hall (1970)
[Quine 1976]: "Worlds away", *The Journal of Philosophy* **73**, pp. 859-63
[Quine 1977]: "Intensions revisited", *Midwest Studies in Philosophy* **2**, réimprimé in *Theories and Things*, pp. 113-23
[Quine 1981]: *Theories and Things*, Harvard University Press, Cambridge (Mass.) et Londres
[Quine 1986]: *The Philosophy of W. V. Quine*, Hahn, Lewis Edwin (ed.), La Salle Open Court, pp. 209-28
[Quine 1993]: *La Poursuite de la Vérité*, Editions du Seuil, Paris; édition originale : *Pursuit of Truth*, Harvard University Press, 1990
[Quine 1994]: "Promoting extensionality", *Synthese* **98**, pp. 143-51
[Quine 1997]: *The Ways of Paradox* (*and Other Essays*) (7e éd.), Harvard University Press.
[Radford 1966]: "Knowledge –by examples", *Analysis* **27**, pp. 1-11
[Radford 1967]: "Knowing but not believing", *Analysis* **27**, pp. 139-40

[Rantala 1979]: "Possible worlds and formal semantics", in Esa Saarinen & Risto Hilpinen & Ilkka Niiniluoto & Merrill Provence Hintikka (eds.) : *Essays in Honour of Jaakko Hintikka*, D. Reidel Publishing Company, Dordrecht : Holland, pp. 177-188
[Rantala 1982]: "Impossible worlds semantics and logical omniscience", *Acta Philosophica Fennica* **35**, pp. 106-115 ; réimprimé in I. Niiniluoto et E. Saarinen (eds.), *Intensional Logic : Theory and Applications*, Helsinki
[Rantala 2003]: "Possible worlds", in *Analytic Philosophy in Finland*, in *Poznan Studies in the Philosophy of the Sciences and the Humanities* **80**, Amsterdam, New-York, pp. 179-99
[Rescher 1960]: "The problem of a theory of belief for statements", *Philosophy of Science* **27**, pp. 80-95
[Restall 2002]: "Carnap's tolerance, language change and logical pluralism", *Journal of Philosophy* **99**, pp. 426-43
[Russell 1905a]: "On denoting", *Mind* **14**, pp. 479-493
[Russell 1905b]: "Necessity and possibility", réédité in *The Collected Papers of Bertrand Russell*, Vol. 4 : "Foundations of logic : 1903-05", pp. 507-20
[Russell 1911]: "Knowledge by acquaintance and knowledge by description", réédité in *The Collected Papers of Bertrand Russell*, Vol. 6 : "Logical and Philosophical Papers : 1909-13", pp. 148-61
[Russell 1913]: "Belief, disbelief, and doubt", réédité in *The Collected Papers of Bertrand Russell*, Vol. 7 : "Theory of Knowledge : The 1913 Manuscript", pp. 136-143
[Russell 1923a]: "Truth-functions and meaning-functions", réédité in *The Collected Papers of Bertrand Russell*, Vol. 9 : "Language, Mind and Matter : 1919-26", pp. 156-8
[Russell 1923b]: "What is meant by 'A believes p' ?", réédité in *The Collected Papers of Bertrand Russell*, Vol. 9 : "Language, Mind and Matter : 1919-26", p. 159
[Russell 1969]: *Signification et Vérité*, Champs Flammarion, Paris
[Russell 1973]: "Meinong's Theory of Complexes and Assumptions, I & II & III" (1904), réédité in Essays in Analysis, D. Lackey (ed.), London, Allen & Unwin.
[Russell 2001]: "L'essence de la philosophie : la logique", in *La méthode scientifique en philosophie*, Editions Payot & Rivage, Paris
[Russell 2002]: *Théorie de la Connaissance*, orig. *Theory of Knowledge*, trad. française de J.-M. Roy, Vrin.
[Russell & Whitehead 1973]: *Principia Mathematica* (1910-1912), versions paperback de *1 à *56, Cambridge Univ. Press
[Scott 1970]: "Advice on modal logic", in *Philosophical Problems in Logic*, K. Lambert (ed.), Dordrecht: D. Reidel Publishing Company, pp. 143-173
[Searle 1969]: *Speech Acts*, Cambridge Univ. Press
[Searle & Vanderveken 1985]: *Foundations of Illocutionary Logic*, N.-Y., Cambridge Univ. Press
[Slater 1995]: "Paraconsistent logics ?", *Journal of Philosophical Logic* **24**, pp. 451-54
[Smith 1983]: "Kantifying", *Synthese* **54**, pp. 261-73
[Stahl 1979]: "Les syllogismes avec 'savoir' chez Occam", *Logique et Analyse* **21**, pp. 437-449
[Steup 2001]: "The analysis of knowledge", in *Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/knowledge-analysis
[Strawson 1952]: *Introduction to Logical Theory*, London, Methuen
[Suszko 1977]: "The Fregean Axiom and Polish mathematical logic in the 1920's", ***Studia Logica*** **36**, 377-380
[Tarski 1944]: "The semantic conception of truth and the foundations of semantics", *Philosophy and Phenomenological Research* **4**, pp. 341-376
[Tarski 1986]: "What are logical notions?", *Philosophy and History of Logic* **7**, pp. 143-157

[Thagard & Nisbett 1983]: "Rationality and charity", *Philosophy of Science* **50**, pp. 250-67
[Triplett 1988]: "Azande logic *versus* Western logic?", *The British Journal in the Philosophy of Science* **39**, pp. 361-6
[Urmson 1952]: "Parenthetical Verbs", *Mind* **61**, pp. 480-496
[Upton 1983]: "Psychological and metaphysical dimensions of non-contradiction in Aristotle", *Review of Metaphysics* **36**, pp. 591-60
[van Benthem 1991]: "Reflections on epistemic logic", *Logique et Analyse* **134**, pp. 5-14
[van Benthem 2004]: "What one may come to know", *Analysis* **64**, pp. 95-105
[van Heijenoort 1967]: "Logic as calculus and logic as language", *Synthese* **17**, pp. 324-30
[Varzi 2002]: "On logical relativity", *Philosophical Issues* **12**, pp. 197-219
[Vernant 1997]: *Du discours à l'action*, PUF, Paris
[Vernant 2003]: "Pour une logique dialogique de la dénégation", en ligne : http://web.umpf-grenoble.fr/SH/PersoPhilo/DenisVernant/Denegation.pdf, pp. 1-10; disponible également in F. Armengaud, M.-D. Popelard & D. Vernant (éds.) : *Du Dialogue au texte*, autour de Francis Jacques, Paris, Kimé
[Vernant 2005]: "Les limites d'un traitement logique de l'assertion", en ligne : http://web.umpf-grenoble.fr/SH/PersoPhilo/DenisVernant/Limitesassertion.pdf, pp. 1-20 ; disponible également in F. Clémentz (éd.) : *Russell : Logique, pensée et connaissance*, Paris, Vrin ; version anglaise in D. Vanderveken (ed.) : Logic, Thought and Action, Netherlands, Springer, ch. 13, pp. 267-88
[Vernant 2006] : "La genèse logique du concept de dénégation de Frege à Słupecki", in R. Pouivet & M. Rebuschi (éds.) : *La philosophie en Pologne (1918-1939)*, Vrin, pp. 151-178
[von Wright 1951]: *An Essay in Modal Logic*, North-Holland Publishing, Amsterdam
[von Wright 1952]: "Interpretations of modal logic", *Mind* **61**, pp. 165-177
[von Wright 1986] : "La logique modale et le *Tractatus*", in *Wittgenstein*, trad. E. Rigal, TER, pp. 195-213
[Wansing 2002]: "Diamonds are a philosopher's best friends: the Knowability Paradox and modal epistemic relevance logic", *Journal of Philosophical Logic* **31**(6), pp. 591-612
[Wertz 1973]: "How to things with Austin's words", *The New Scholasticism* **47**, pp. 478-82
[Wertz 1976]: "Poor Austin's words : a reply to Young", *The New Scholaticism* **50**, pp. 366-75
[Williams 1979]: "Moore's paradox : one or two ?", *Analysis* **39**, pp. 141-2
[Williams 1981]: "Inconsistency and contradiction", *Mind* **90**, pp. 600-2
[Williams 1982b]: "Believing the self-contradiction", *American Philosophical Quarterly* **19**, pp. 279-285
[Williams 2000]: "Moore's problem with iterated belief", *Philosophical Quarterly* **50**, pp. 28-43
[Williamson 1990]: "Two incomplete antirealistic modal epistemic logics", *The Journal of Symbolic Logic* **55** (1), pp. 297-314
[Williamson 1992]: "On intuitionist epistemic modal logic", *Journal of Philosophical Logic* **21**, pp. 63-89
[Wiredu 1975]: "Truth as a logical constant, with an application to the principle of excluded middle", *The Philosophical Quarterly* **25**, pp. 305-317
[Wittgenstein 1969]: *Tractatus Logico-Philosophicus*
[Zemach 1969]: "The pragmatic paradox of knowledge", *Logique et Analyse* **12**, pp. 283-7
[Zemach 1971]: "Epistemic opacity", *Logique et Analyse* **14**, pp. 803-810

Glossaire

A

Affirmation locutoire / illocutoire
En logique traditionnelle, l'*affirmation locutoire* désigne la *qualité* d'un énoncé* (au sens 2.), par opposition à sa *quantité*. Pour distinguer l'affirmation locutoire de sa version illocutoire, on utilise aussi l'expression 'affirmative' pour désigner la qualité ou contenu de l'énoncé.
Exemple : l'énoncé 'Tous les hommes sont mortels' est une universelle affirmative, par opposition à l'universelle négative 'Aucun homme n'est mortel'.
En logique moderne, l'*affirmation illocutoire* désigne l'acte du locuteur visant à asserter la vérité d'un énoncé (au sens 1.). Cette affirmation désigne un acte d'énonciation et ne renvoie pas à une qualité objective de l'énoncé.
Note : le rejet dans la thèse de la distinction locutoire/illocutoire conduit à une redéfinition de la différence entre les affirmations (et négations*) : tous deux sont des actes illocutoires de *force* distincte, la première exprimant une supposition de vérité $S\phi$ et la seconde une assertion de vérité $A\phi$.
Exemple : la déclaration 'Aucun homme n'est mortel' est une assertion portant sur une universelle négative, symbolisée $A\neg\phi$.

Antiréalisme / réalisme sémantique
Le *réalisme sémantique* est une théorie de la vérité selon laquelle les conditions de vérité d'un énoncé transcendent les capacités de reconnaissance de cette vérité ; par opposition, l'*antiréalisme sémantique* définit la vérité comme la capacité d'un sujet (idéal ou non) à reconnaître ou vérifier un énoncé.
Le débat entre réalistes et antiréalistes renvoie à une approche *cognitive* de la vérité sémantique (décrire l'état du monde), par opposition à l'approche *volitive* du discours de vérité (exprimer nos attitudes vis-à-vis de cet état).
Exemple : en logique illocutoire*, l'acte assertif comporte un aspect tantôt assertif tantôt volitif : décrire un état du monde, mais aussi imposer son opinion ou son autorité sur un auditoire.

Assertion
Une assertion $A\phi$ est une affirmation ou déclaration de vérité effectuée par un locuteur par le biais d'un énoncé* (au sens 2.) ϕ ; synonyme de 'tenir pour vrai' ou 'considérer comme vrai', l'assertion est l'expression d'une croyance du locuteur.
On distingue parfois l'assertion de l'affirmation* : toutes deux portent sur un même contenu propositionnel*, mais elles se distinguent par leur force illocutoire lorsque l'affirmation correspond à une supposition de vérité.

Auto-annulation (en anglais : 'self-defeatness')
Une expression est auto-annulante [*self-defeating*] si son énonciation est contraire au but recherché.

Exemple : l'énoncé 'Il pleut, mais je ne le crois pas' est auto-annulant dans la mesure où l'expression 'je ne le crois pas' est contraire au but recherché par l'assertion de 'il pleut'.
Les formules auto-annulantes sont des cas d'indéfendabilité épistémique* ou doxastique*.

Auto-confirmation (en anglais : 'self-sustenance')
Dans [Hintikka 1962], propriété métalogique en vertu de laquelle une formule modale épistémique est *auto-confirmée* [*self-sustaining*] si elle est valide pour un agent idéal, c'est-à-dire vraie dans tous les mondes possibles compatibles avec les attitudes épistémiques de cet agent.
Exemple : la clôture sous implication du système modal **K** est une formule auto-confirmante et caractérise la propriété d'omniscience logique des agents épistémiques idéaux.

C

Clause de vérité
Appelée également 'clause de succès', la *clause de vérité* désigne en logique modale épistémique l'axiome caractéristique du système modal **T** : $K\phi \rightarrow \phi$, toute connaissance est vraie.
L'interprétation illocutoire de la clause de vérité donne une clause de *véridicité* ou de *véracité* : soit une assertion de vérité implique la vérité du contenu propositionnel, soit une assertion de vérité implique l'intention du locuteur de dire la vérité

Complétude
Propriété métalogique d'un système formel qui peut être définie soit syntaxiquement, soit sémantiquement.
Un système est *syntaxiquement complet* si, pour toute formule A, soit A soit $\neg A$ est un théorème du système : soit $\vdash A$, soit $\vdash \neg A$. Si A n'est pas un théorème d'un système, ce système est syntaxiquement complet si l'introduction de A le rend contradictoire.
Un système est *sémantiquement complet* si toute formule A valide dans ce système y est également démontrable : si $\vDash A$, alors $\vdash A$.
La combinaison de la complétude sémantique et de la consistance sémantique* donne le *métathéorème de complétude* : un système est complet si et seulement si toute formule A de ce système y est démontrable, et inversement : $\vdash A$ si et seulement si $\vDash A$.

Conséquence logique
(En allemand : 'Folgerung')
Dans un langage formel, la formule ψ est une *conséquence logique* de ϕ si et seulement si la vérité de ϕ dans un modèle entraîne la vérité de ψ dans ce modèle : $\vDash(\phi \rightarrow \psi)$, ou $(\phi \Rightarrow \psi)$.
On distingue la conséquence logique, qui est une conséquence d'ordre sémantique, et la *conséquence syntaxique*, synonyme de la déduction* ($\vdash(\phi \rightarrow \psi)$).

Consistance
Propriété métalogique d'un système formel qui peut être définie soit syntaxiquement, soit sémantiquement.
Un système est *syntaxiquement consistant* si aucune formule A de ce système n'est telle qu'on peut y démontrer à la fois A et $\neg$A : si $\vdash$A, alors $\nvdash \neg$A ; si $\vdash\neg$A, alors $\nvdash$A.
Un système est *sémantiquement consistant* si toute formule A démontrable dans ce système est également valide : si $\vdash$A, alors $\vDash$A.

Contenu propositionnel
Désigne l'énoncé interprété ϕ, sur lequel porte un acte illocutoire de forme $\odot\phi$.
Notes : Le contenu propositionnel ne porte pas sur une proposition* au sens de Frege, qui est une entité non-linguistique ; une appellation plus conforme à l'approche illocutoire inciterait à parler de 'contenu énonciatif'. Il correspond à l'aspect locutoire du langage mais, étant donné le caractère synchronique des aspects locutoire et illocutoire lors d'une énonciation, [Searle 1969] parlait d'*acte propositionnel* pour qualifier l'acte de prédication (d'assignation d'une propriété à un ou plusieurs référents).

Contingence
Propriété modale en vertu de laquelle une formule ϕ est *contingente* si elle n'est ni nécessaire ni impossible.
Formellement : $\nabla\phi =_{df} \neg\Box\phi \wedge \neg\Box\neg\phi$.

D

Déclaration (en anglais : 'statement')
1. Acte de discours par lequel un locuteur affirme quelque chose par le biais d'un énoncé (au sens 2.) ; synonyme d'affirmation* (au sens de la supposition de vérité).
La logique traditionnelle d'Aristote est une logique déclarative (ou apophantique) : elle traite des déclarations, c'est-à-dire d'énoncés* au sens 2.
2. En logique illocutoire ([Searle & Vanderveken 1985]), type d'acte illocutoire qui provoque un nouvel état de choses par le biais de sa propre énonciation (exemple : la définition d'une règle, le baptême, etc.).
La différence entre les acceptions 1. et 2. correspond à la différence entre un acte de discours (affirmer quelque chose) et son but (transformer l'état du monde).
L'aspect déclaratif (au sens 2.) des modalités épistémiques apparaît dans une *déclaration de savoir* : en disant 'je sais que ϕ', je produis un nouvel état de choses en modifiant l'opinion de mon auditoire (ils me croient) ; mais on considère en général que le concept de connaissance ne se contente pas d'être déclaré : 'savoir' connote un aspect à la fois déclaratif et assertif (on suppose que l'état de choses déclaré n'a pas été créé par la déclaration, contrairement aux actes de baptêmes ou définitions), car la déclaration de savoir doit être également justifiée.
Note : Cet aspect mixte du verbe 'savoir' fait ressortir le rôle de l'autorité dans la transmission du savoir (conditions préparatoires des actes déclaratifs) ; elle met aussi en

évidence l'aspect à la fois ontologique (fait brut) et épistémologique (fait social) des notions de vérité et de connaissance.

Déduction
Une déduction est une relation de *conséquence syntaxique* entre un ensemble de formules ϕ et une conséquent ψ. Symboliquement : $\phi \vdash \psi$.
L'application d'une relation déductive par un agent donne une inférence*.

Défendabilité
Dans [Hintikka 1962], propriété métalogique en vertu de laquelle une formule modale épistémique est *défendable* si et seulement si elle est satisfiable pour un agent idéal, c'est-à-dire vraie dans au moins un monde possible compatible avec les attitudes épistémiques de cet agent.

Déflationisme
En théorie de la vérité, une interprétation est *déflationniste* (ou déflationnaire) lorsqu'elle interprète l'expression 'est vraie' comme redondante et supprime le prédicat de vérité.
Exemple : 'Socrate est chauve' et 'L'énoncé "Socrate est chauve" est vrai' ont le même sens dans une interprétation déflationniste.
Le concept de vérité désigne la valeur assignée à un énoncé (au sens 1.) lors d'une interprétation ; il désigne alors l'opinion d'un locuteur sur l'état du monde et devient synonyme de l'énonciation comme acte assertif : asserter quelque chose, c'est avoir l'intention de décrire un énoncé vrai.

Dénégation (en anglais : 'denial')
En logique illocutoire, action de dénier que l'on distingue à la fois de la négation comme action de nier (affirmation de fausseté) et comme qualité d'un contenu propositionnel (attribution de fausseté à une formule).
Exemple : 'je ne nie pas qu'il soit riche'. Synonyme de 'rejet' ou de 'refus', la dénégation correspond à une négation illocutoire. Formellement : $\dashv\phi$, $D\phi$.
Pris dans ce sens, dénier consiste à suspendre son jugement vis-à-vis d'un contenu propositionnel ϕ et à ne le déclarer ni vrai ($A\phi$), ni faux ($A\neg\phi$). Formellement : $D\phi =_{df} A(\neg A\phi \wedge \neg A\neg\phi)$.
La dénégation (négation illocutoire) est distinguée de la négation* (négation locutoire), qui correspond soit à une affirmation de fausseté ($A\neg\phi$), soit à une supposition de fausseté ($S\neg\phi$).
Exemple : la règle du *modus tollens* signifie que, de l'affirmation de '$\phi \rightarrow \psi$' et de la négation de 'ψ', on peut inférer la négation de 'ϕ'. Formellement : $A(\phi \rightarrow \psi), A\neg\psi \vdash A\neg\phi$.

Désignation rigide
Théorie de la nomination due à Kripke, en vertu de laquelle un nom propre a la même extension dans tous les mondes possibles.
La théorie de la désignation rigide, connue également sous l'appellation de *Nouvelle Théorie de la Référence* (dont les partisans sont entre autres Barcan Marcus, Donnellan, Føllesdal),

implique une distinction entre des mondes possibles de nature distincte (métaphysique et épistémique).

Direction d'ajustement
Désigne chez [Searle 1969] la direction entre les mots du langage (le dire) et les choses (le monde) ; le type de direction détermine la signification [*meaning*] de l'acte propositionnel effectué par le locuteur.
Il y a quatre directions d'ajustement principales : langage-monde (assertifs), monde-langage (promissifs, directifs),la double direction (déclaratifs), et la direction nulle (expressifs).
Les modalités épistémiques ont un but illocutoire qui varie selon l'intention de leur usage : assertifs (savoir, avoir la preuve de), déclaratifs (asserter, déclarer, annoncer, décréter), expressifs (être certain, être convaincu, croire, penser).

Duperie de soi (en anglais : ('self-deception')
Désigne la tendance chez un agent à refuser la vérité de ses propres croyances ; la duperie de soi est considérée parfois comme l'expression de tendances contradictoires ou irrationnelles chez un agent.
Des auteurs paraconsistants ont mentionné la duperie de soi comme un contre-exemple aux théorèmes d'entraînement ($K\phi \rightarrow B\phi$) et de consistance ($B\phi \rightarrow \neg B\neg\phi$).

E

Enoncé (en anglais : 'sentence')
1. Entendu comme la proposition grammaticale, l'énoncé (le *Satz* de Frege) est une phrase susceptible d'être interprétée.
Synonymes : phrase, signe propositionnel, forme de mots, lekton ('dicible', 'énonçable').
2. Le résultat de l'interprétation d'une proposition grammaticale (synonyme de 'statement')
[Frege 1889, 1892] parle de contenu jugeable et de sens ou pensée [*Gedanke*].
Lorsque l'énoncé est interprété, on parle d'un déclaratif, interrogatif, exclamatif, etc., spécifié par un mode illocutoire (une direction d'ajustement*).
Exemple : 'Socrate est chauve' est un énoncé *déclaratif*, c'est-à-dire un énoncé pourvu de sens qui exprime une pensée et dont le but est de décrire un état de choses
'Socrate est-il chauve ?' est un énoncé interrogatif qui n'est pas déclaratif et ne vise pas à décrire un état de choses ; en revanche, il porte sur un énoncé susceptible d'être vrai ou faux (une pensée).
Note : tous les énoncés (déclaratifs, interrogatifs, exclamatifs, etc.) contiennent ou portent sur une pensée, mais ils se distinguent par le but illocutoire* de leur énonciation ; un jugement ou assertion* a pour but de dire que cette pensée est vraie ; une question a pour but de demander si cette pensée est vraie ; une exclamation a pour but d'exprimer le souhait que cette pensée soit vraie, etc.

Explosion
Propriété des systèmes logiques *consistants* en vertu de laquelle toute formule contradictoire entraîne la trivialité du système, c'est-à-dire la déductibilité de n'importe quelle formule. Formellement : $(\phi \wedge \neg\phi) \vdash \psi$, pour tout ϕ et ψ.
La paraconsistance est l'ensemble des systèmes logiques dans lesquels un ensemble de formules peut être inconsistant sans être trivial.

Exportation/importation (lois d') :
Formules de la logique modale exprimant la relation logique entre les modalités quantifiées. Elle se lisent soit dans le sens *de re-de dicto* de l'importation (du monde extérieur vers le monde intérieur) :
$\Box(\exists x)Fx \rightarrow (\exists x)\Box Fx$
soit dans le sens *de dicto-de re* de l'exportation (du monde intérieur vers le monde extérieur) :
$(\exists x)\Box Fx \rightarrow \Box(\exists x)Fx$
La validité de ces formules dépend de la nature du domaine d'interprétation (constant, non constant), leur adéquation formelle dépend du contexte de discours.

Extensionalisme
Principe sémantique en vertu duquel la signification d'une expression est son extension, c'est-à-dire l'objet ou ensemble d'objets qui lui correspond.
L'application de ce principe varie selon la nature de l'expression interprétée (constante individuelle, prédicative, propositionnelle).
L'interprétation extensionnelle du langage a été remise en cause par le discours des modalités, synonymes d'intensions*.

H

Hyperintensionalité
Propriété en vertu de laquelle certaines expressions d'un langage n'ont pas le même référent dans tous les mondes possibles.
Le contexte épistémique est un contexte hyperintensionnel : un nom propre peut ne pas avoir le même référent dans tous les mondes possibles des agents.

I

Illocutoire
Aspect des actes de discours visant à exprimer l'action accomplie lors d'une énonciation.
Exemple : 'je baptise cet enfant' est un acte de discours dont l'aspect illocutoire est l'acte de baptême (cf. la notion de performatif chez [Austin 1962]).

La typologie des différents types de force illocutoire a été dressée dans [Searle 1969].

Indéfendabilité
Dans [Hintikka 1962], propriété métalogique en vertu de laquelle une formule modale épistémique est *indéfendable* si et seulement si elle est impossible à satisfaire pour un agent idéal, c'est-à-dire vraie dans aucun monde possible compatible avec les attitudes épistémiques de cet agent.

Inférence
(En allemand : 'Schluß')
Application d'une relation de déduction entre un ensemble de *prémisses* ϕ, c'est-à-dire un ensemble de formules dont on admet la vérité, et une conclusion ψ.
Symboliquement : $J\phi \Rightarrow J\psi$ (où 'J' représente l'acte de jugement, c'est-à-dire de considérer comme vrai).
On distingue l'inférence de la relation de conséquence (syntaxique ou sémantique) : dans une inférence, les hypothèses ϕ sont considérées comme vraies, c'est-à-dire que l'agent s'engage sur la vérité des formules.

Intensions
Par opposition aux extensions, les intensions correspondent au sens* des expressions.
Si l'on définit l'intension comme la règle qui détermine l'extension d'une expression dans chaque monde possible, deux expressions auront la même intension (sont co-intensionnelles) si elles ont la même extension dans tous les mondes possibles. L'intension d'un énoncé correspond à sa proposition* (au sens 2. de la Proposition).
Synonyme de 'compréhension'.

L

Logique appliquée/pure
Une logique appliquée est un langage formel (pur ou non interprété) dans lequel les expressions reçoivent une interprétation dans la langue naturelle. Exemple : la logique épistémique est une logique modale interprétée dont le but est d'expliquer le comportement logique des concepts de connaissance et de croyance.
Synonyme : logique philosophique*.

Logique déviante/enrichie
Une logique déviante est une logique non-standard dans laquelle certains théorèmes de la logique standard ne sont pas des théorèmes. Exemple : les logiques intuitionniste, quantique, floue, paraconsistante sont des logiques déviantes.
Une logique enrichie est une logique qui contient tous les théorèmes de la logique standard et en ajoute d'autres. Les logiques modales classiques sont des logiques enrichies.

Logique philosophique
Ensemble des logiques appliquées* destinées à analyser certaines expressions des langues naturelles, parmi lesquelles des concepts traités par la tradition philosophique (le temps, la connaissance, le devoir, la vérité, l'existence).
Dans la mesure où les logiques philosophiques formalisent les énoncés des langues naturelles et l'intention signifiante des locuteurs, elles correspondent toutes à des logiques intentionnelles exprimant un certain rapport du sujet-locuteur au monde (au temps : logique temporelle ; à l'éthique : logique déontique ; à la vérité : logique épistémique, logique déclarative). Les énoncés déclaratifs et épistémiques sont formalisables au sein d'une logique commune des modalités assertoriques.

M

Modalité
Ensemble des modes exprimant un jugement sur un énoncé au sens 2. (est vrai de manière nécessaire, possible, impossible, contingente, probable, certaine, indubitable, etc.)
La modalité correspond à l'aspect *neustique* des propositions (cf. Richard Hare), par contraste avec les aspects *phrastique* (le contenu propositionnel) et *tropique* (les modes verbaux). Par analogie, les notions de vérité/fausseté et d'existence sont des formes de modalités appliquées à des énoncés au sens 1. (non interprété).

Mode
En grammaire, manière dont le mode exprime l'état ou l'action (les six modes de la langue française) et qui peut être formalisé en logique modale.
Le subjonctif, l'impératif ou le conditionnel, mais aussi les modes de certaines langues étrangères (optatif, vocatif) correspondent à des modalités illocutoires* de type directif.
Issu du terme médiéval 'modus', désigne le marqueur de force illocutoire attaché à un énoncé (au sens 2.) ou contenu propositionnel (le *dictum*).
Exemple : dans une assertion, le mode de l'énoncé est le mode assertif dont la direction d'ajustement va des mots aux choses.
Note : on considère généralement le mode comme s'appliquant à un énoncé au sens 2., c'est-à-dire un énoncé déjà pourvu d'un sens (une valeur de vérité) ; mais il s'applique plutôt à un énoncé au sens 1., si l'on considère le fait d'associer une pensée à un signe propositionnel comme un acte de modalisation (d'application d'un mode) à part entière.
La différence entre le *modus* des médiévaux et le mode de vérité ou de fausseté (cf. [von Wright 1951] réside alors non plus dans l'argument sur lequel ils portent (tous deux portent sur un énoncé au sens 1.) mais dans leur force illocutoire : un jugement est plus fort qu'une supposition de vérité, laquelle nous avons assimilée à un contenu propositionnel ou pensée.
La différence entre les deux types de modes ci-dessus correspond à la différence entre l'interprétation d'une formule dans un modèle et la modalisation d'une formule interprétée, ou à la différence entre une fonction d'interprétation (application de $\mathcal{L}$ dans $\mathcal{M}$) et une fonction de valuation (application de $\mathcal{M}$ dans $\mathcal{M}$).

Modèle
Un modèle $\mathcal{M}$ pour une formule ϕ est une interprétation qui rend ϕ vraie ; par extension, le modèle d'un système est une interprétation qui rend vraie l'ensemble des axiomes et théorèmes de ce système.
Par opposition, on appelle *contre-modèle* de ϕ une interprétation qui rend ϕ fausse.
La distinction entre logique modale et logique non-modale porte sur la structure du modèle : il se compose d'un ensemble de formules unique (la classe d'unité $\{W\}$) en logique non-modale, d'un ensemble d'ensemble de formules (une classe de mondes $w \in W$ spécifiés par leur relation R) en logique modale.

Mondes possibles/impossibles
En sémantique des logiques modales, les mondes possibles désignent un sous-ensemble de formules w dans un ensemble d'ensembles de formules W. Synonyme d'états de choses (Wittgenstein) et de descriptions d'état (Carnap), les mondes possibles sont interprétés dans différentes structures au sein des modèles de Kripke.
Les mondes impossibles désignent un ensemble de formules inconsistants ; ils sont utilisés en logique épistémique pour fournir un contre-modèle aux théorèmes normaux système **K** (cf. Rantala).

Moore (paradoxe de)
Ensembles d'énoncés absurdes mais non contradictoires de type 'Il pleut, mais je ne le crois pas', ou 'Je crois qu'il pleut, mais il ne pleut pas'.
Formellement : $\phi \wedge \neg B\phi$, $B\phi \wedge \neg\phi$.
L'interprétation illocutoire* de ces deux formulations y révèle une contradiction entre la *condition de sincérité* de leur énonciation et leur contenu propositionnel : dire que $\phi/\neg\phi$, c'est penser que ϕ est vrai/n'est pas vrai ; dire que l'on ne croit pas/croit que ϕ, c'est penser que ϕ n'est pas vrai/est vrai. On obtient dans les deux une conjonction contradictoire : ϕ est vrai $\wedge$ ϕ n'est pas vrai.

N

Négation
Parallèlement à l'affirmation*, désigne tantôt la négation locutoire tantôt la négation illocutoire.
La *négation locutoire* renvoie à la qualité négative de l'énoncé : $\neg\phi$; par extension, désigne l'acte de supposer la fausseté du contenu propositionnel* (formellement : $S\phi$), par opposition à la *négation illocutoire* qui désigne l'acte d'asserter la fausseté du contenu propositionnel (formellement : $A\neg\phi$). On distingue la négation de la dénégation*.
Note : le signe de la négation $\neg$ symbolisant l'acte de supposition de fausseté, on confond la fausseté avec la supposition de fausseté.

Normalité
En logique modale propositionnelle, un système axiomatique est normal s'il possède les trois propriétés minimales du système modal **K** : tous les théorèmes standards sont des théorèmes de ce système ; l'opérateur modal □ se distribue sur l'implication (et la conjonction) ; la règle de nécessitation.
Un système normal satisfait la propriété de monotonie et peut être interprétée dans un modèle de Kripke minimal (où R n'a aucune propriété spécifique).

O

Omniscience logique
Propriété caractéristique du système modal **K** en logique épistémique, l'omniscience logique provient de l'application de la règle de distribution et signifie que toutes les conséquences logiques d'une connaissance doivent être également connues.
La recherche d'une solution à ce défaut structurel a conduit entre autres à la construction de modèles de mondes possibles impossibles* (modèles de Rantala).

Opacité référentielle
Par opposition à la transparence référentielle, se dit d'un contexte de discours dans lequel une expression référentielle n'est pas purement désignative, c'est-à-dire ne désigne pas un objet du domaine d'interprétation (pas de référence : Quine) ou ne désigne pas un seul objet du modèle d'interprétation du langage (référence multiple : Hintikka).
La seconde explication de l'opacité (référence multiple) signifie que l'interprétation peut assigner plus d'un objet à une seule expression : la fonction d'interprétation n'est pas bijective, car à une expression ne correspond pas un et un seul objet du modèle.

P

Pensée
1. Le processus de pensée, défini comme une opération appliquée à un ensemble d'éléments et dont le résultat est une pensée (au sens 2.) ;
2. Le résultat de cette opération : la pensée en tant qu'objet de pensée.
La distinction ci-dessus est consacrée dans l'histoire de la philosophie du langage par plusieurs paires de notions : noèse/noème (Husserl) ; modus, dictum (médiévaux), signe propositionnel/proposition* (Wittgenstein), signe/sens* (Frege), lekton/noêton (stoïciens), antécédent/image, contenu propositionnel/énoncé*.

Proposition
1. Proposition grammaticale synonyme de l'énoncé au sens 1. (le *Satz* de Frege), entendue comme une expression susceptible d'être interprétée dans un modèle. Synonyme : phrase.

Exemple : 'La neige est blanche', 'Schnee ist weiβ' et 'snow is white' sont trois propositions grammaticales distinctes issues de trois langues naturelles différentes.
2. Désigne l'intension ou signification commune à plusieurs énoncés* (au sens 2) ou pensées (les *Gedanken* de Frege), distinguée parfois par la formule majuscule 'Proposition'. Exemple : 'La neige est blanche', 'Schnee is weiβ' et 'snow is white' sont trois énoncés (au sens 2.) distincts exprimant le même sens, c'est-à-dire la même proposition au sens 2.
Le rejet des intensions* ou Propositions (cf. la thèse d'indétermination de la traduction de Quine) conduit au rejet de la 'Proposition' comme entité objective saisie par des locuteurs.
Si l'on rejette l'existence autonome des entités intensionnelles de Frege, le sens désigne l'acte particulier de référence réalisé par un locuteur (dans un ou plusieurs de ses mondes possibles). Cf. la définition par [Hintikka 1969b] du sens* comme référence multiple.
Note : plusieurs locuteurs expriment la même proposition lorsque leurs énonciations ont le même but illocutoire et lorsque leur contenu propositionnel ont la même extension (leurs constantes d'individu et de prédicat ont la même interprétation).
Exemple : 'Ferme la porte !', 'Je t'ordonne de ne pas laisser cette salle ouverte' et 'Close the door !' sont trois énoncés exprimant la même proposition si leurs locuteurs effectuent le même acte illocutoire directif (donner un ordre) sur le même contenu propositionnel (l'individu *a* ferme la porte P à un temps *t*).

S

Sémantique formelle
Interprétation des expressions d'un langage logique dans un modèle*.
La sémantique formelle assigne des valeurs sémantiques à chacune de ces expressions : elle interprète les expressions prédicatives en termes d'objets et propriétés dans un domaine de référence, les expressions propositionnelles en termes de valeurs de vérité.

Sens (et référence)
En logique modale : le sens est le référent d'une expression dans tous les mondes possibles*.
Si $I(\phi)$ est l'interprétation de ϕ, symbolisée $\vDash \phi$, alors le sens de ϕ est son interprétation dans tous les mondes possibles : $w \vDash \phi$, alors $\forall w^* R w, w^* \vDash \phi$.
Le problème de l'identité du sens des énoncés (au-delà des locuteurs et des langues naturelles) se pose dans les contextes modaux. Comment garantir l'identité de la référence dans *tous* les mondes possibles, et qu'est-ce qu'un *monde possible* ?
Exemple : dans l'énoncé déclaratif 'la neige est blanche', le sens de l'énoncé (au sens 2.) dépend de l'intention signifiante du locuteur (cf. opacité référentielle*) ; mais Frege suppose à l'inverse que le sens d'un énoncé transcende les langues et les locuteurs. Le rôle du sens fregéen, entendu comme une entité idéale intermédiaire entre l'énoncé et la référence, exprime l'autonomie de la pensée vis-à-vis des actes de pensée des locuteurs.
Chez [Frege 1892], l'identité du sens [*Sinn*] des phrases [*Sätze*] est incarnée par l'entité objective appelée pensée [*Gedanke*] : une pensée est la proposition* (au sens 2.) qui correspond au sens d'un énoncé. Chez Quine, le rejet de ces entités (les significations

objectives) conduit à la thèse d'indétermination de la traduction : le sens dépend des locuteurs et se construit par un processus intersubjectif d'acquisition du langage référentiel.
Par extension : le sens est le processus ou la méthode d'interprétation d'une expression d'un langage, par opposition à l'extension visée par l'interprétation. On distingue la méthode d'interprétation d'une expression et son interprétation par les paires sens/référence, intension/extension, ou compréhension/extension.
Par extension : si l'on reprend l'idée de sens comme projection (application) ou direction d'un point vers un autre, le sens correspond en logique illocutoire* à la direction d'ajustement* des actes de discours (cf. [Searle 1969], [Searle & Vanderveken 1985]) et détermine les règles d'interprétation des expressions en termes de relation langage-monde.
Dans un métalangage formel, le sens est décrit par les règles d'application du langage dans un modèle (direction langage-monde), ce qui correspond au mode de donation [*Gegebenheitsweise*] d'un antécédent vers son image ; on retrouve dans ce processus d'application le sens défini par Frege, mais sans associer d'identité idéale intermédiaire entre l'antécédent (l'expression) et l'image (l'objet de référence).

Signification
(En anglais : 'meaning' ; en allemand : 'Bedeutung').
Souvent confondue avec le sens*, la signification d'une expression est définie tantôt par sa seule référence (son extension), tantôt par son sens et sa référence.
D'après les théoriciens de la désignation rigide* (et Russell), un nom propre (logique) n'a pas de sens* et sa signification tient uniquement dans sa référence ; par opposition, la tradition frégéenne considère la signification de toute expression comme un composé de son sens et de sa référence.
Une théorie de l'intentionnalité présente la signification comme le processus global d'assignation d'une valeur à une expression par le biais d'un interprétant (le sujet). La logique illocutoire* adhère à ce schéma de type fregéen et le complète, expliquant le sens* comme une *direction d'ajustement** entre les mots et les choses. Le processus d'interprétation (des mots d'un langage vers les choses du modèle) est une condition nécessaire mais non suffisante de la détermination du sens d'un énoncé (au sens 2.).
Note : 'to mean' est la traduction de 'signifier', mais aussi de '*vouloir* dire' ; le préfixe 'be-' de 'bedeuten' semble suggérer une *action* effectuée sur le signe ; le suffixe '-ifier' de 'signifier' indique de la même manière l'action de *mettre* quelque chose *en* signes, de le représenter par une marque qui le symbolise.
Les trois expressions ci-dessus sous-entendent un lien intentionnel établi entre le signe et son objet, ce qui semble plaider en faveur de l'approche fregéenne de la signification (la référence dépend d'un sens, c'est-à-dire d'une méthode d'assignation d'une extension à un signe).

Symbolisme

Symboles formels :

p, q, r	atomes propositionnels, formules atomiques
x, y, z	variables d'individus
a, b, c	constantes d'individus
P, Q, R	constantes de prédicats
ϕ, ψ, φ	complexes propositionnels, formules complexes
$\ulcorner\phi\urcorner$	déclaration, énonciation de ϕ
$\llcorner$	affirmation, supposition de vérité
$\neg$	négation classique (non, ne … pas), supposition de fausseté
$\sim$	négation intuitionniste ($\sim p =_{df} (p \rightarrow \bot)$)
$\wedge$	conjonction (et)
$\vee$	disjonction (ou)
$\veebar$	disjonction exclusive (soit … soit)
$\rightarrow$	conditionnel, implication matérielle (si … alors)
$\equiv$	biconditionnel (si et seulement si)
$\Vdash$	conséquence logique (syntaxique ou sémantique)
$\vdash$	conséquence syntaxique, barre de jugement (assertion)
$\vDash$	conséquence sémantique, satisfiabilité
$\Rightarrow$	implication (conditionnel valide)
$\Leftrightarrow$	équivalence (bi-conditionnel valide)
1	vrai
0	faux
$\Rightarrow$	présupposition, implicature
$\in$	appartenance
$\notin$	non-appartenance
$\subseteq$	inclusion large
$\cap$	intersection
$\cup$	réunion
$\varnothing$	ensemble vide
$\bot$	antilogie, contradiction, absurdité
$\top$	tautologie
$\therefore$	conclusion syllogistique (donc)
$\forall$	quantificateur universel
$\exists$	quantificateur existentiel
$\prec\!\!-$	implication stricte
$\succ\!\!-\!\!\prec$	équivalence stricte
$\Box$	modalité forte (nécessité)
$\Diamond$	modalité faible (possibilité unilatérale)
∇	possibilité bilatérale, contingence
@	actualité (c'est le cas que)
L	nécessité logique
M	possibilité logique

K_a	nécessité épistémique (*a* sait que)
P_a	possibilité épistémique (il est possible, pour autant que *a* sache, que)
B_a	nécessité doxastique (*a* croit que)
C_a	possibilité doxastique (il est compatible avec ce que *a* croit que)
$-$	barre horizontale (du contenu jugeable), considération, supposition
$\dashv$	rejet
$\odot$	marqueur de force illocutoire (fort ou faible)
A	nécessité illocutoire, assertion (je tiens pour vrai que)
S	possibilité illocutoire, supposition (je suppose vrai que)
D	nécessaire contingence illocutoire, doute, dénégation, rejet (je dénie que)
I	contingence illocutoire, ignorance (j'ignore si)
W	ensemble de mondes possibles (ensemble d'ensemble de formules)
w, w*	mondes possibles (ensemble de formules)
R	relation d'accessibilité binaire dans W (wRw*)
$\mathcal{L}$	langage
$\mathcal{D}$	domaine d'interprétation
f	fonction d'interprétation des expressions prédicatives
v	fonction de valuation des expressions propositionnelles
$\mathcal{F} = \langle W,R \rangle$	structure, cadre
$\mathcal{M} = \langle v,W,R \rangle$	modèle

Propriétés des opérateurs doxastiques et épistémiques :

(C.K*) = **K**	Si $w \vDash K\phi$, alors pour tout w^*R_kw on a : $w^* \vDash K\phi$
(C.B*)	Si $w \vDash B\phi$, alors pour tout w^*R_bw on a : $w^* \vDash B\phi$
(C.K) = **T**	Si $w \vDash K\phi$, alors $w \vDash \phi$
(C.KK*) = **4**	Si $w \vDash K\phi$, alors pour tout w^*R_kw on a : $w^* \vDash K\phi$
(C.BB*)	Si $w \vDash B\phi$, alors pour tout w^*R_bw on a : $w^* \vDash B\phi$
(C.PK) = **5**	Si $w \vDash P\phi$, alors pour tout w^*R_kw on a $\vDash P\phi$
(C.k*) = **D**	Si $w \vDash K\phi$, alors pour au moins un w^*R_kw on a $w^* \vDash \phi$
(C.b*)	Si $w \vDash B\phi$, alors pour au moins un w^*R_bw on a $w^* \vDash \phi$

Schémas logiques :

(1) $K_aK_b\phi \rightarrow K_a\phi$

(2)	$K_b\phi \rightarrow K_aK_b\phi$
(3)	$K_a\phi \rightarrow K_aK_b\phi$
(4)	$\phi \wedge \neg B\phi$
(4B)	$B(\phi \wedge \neg B\phi)$
(4K)	$K(\phi \wedge \neg K\phi)$
(4BK)	$B(\phi \wedge \neg K\phi)$
(4KB)	$K(\phi \wedge \neg B\phi)$
(5)	$B\phi \wedge \neg\phi$
(5B)	$B(B\phi \wedge \neg\phi)$
(5K)	$K(K\phi \wedge \neg\phi)$
(5BK)	$B(K\phi \wedge \neg\phi)$
(5KB)	$K(B\phi \wedge \neg\phi)$
(6)	$\Diamond(K\phi \wedge \neg\phi)$
(7)	$(K\phi \wedge \Diamond\neg\phi)$
(I1)	$(\forall x)(\forall y)(\forall F)\ [(Fx \equiv Fy) \rightarrow (x = y)]$
(I2)	$(\forall x)(\forall y)(\forall F)\ [(x = y) \rightarrow (Fx \equiv Fy)]$
(LL)	$(\forall x)(\forall y)[(\forall F)(Fx \equiv Fy) \equiv (x = y)]$
(LL$\Box$)	$[\Box Fa \wedge (a = b)] \rightarrow \Box Fb$
(GE$\Box$)	$\Box Fa \rightarrow (\exists x)(\Box Fx)$
(BF_e)	$\Box(\exists x)Fx \rightarrow (\exists x)\Box Fx$
(BF_i)	$(\exists x)Fx \rightarrow \Box(\exists x)Fx$
(8)	$K_a(\exists x)(x = b)$
(9)	$(\exists x)K_a(x = b)$
(10)	$K_a(x = c)$
(11)	$(\exists x)K_a(x = c)$
($\Box$I)	$(\forall x)(\forall y)[(x = y) \rightarrow \Box(x = y)]$
(12)	$B(\phi \wedge \neg\phi)$
(13)	$B\phi \wedge B\neg\phi$
(14)	$B\phi \wedge \neg B\phi$
(15)	$A(\phi \rightarrow \psi)$
(16)	$[A(\phi \rightarrow \psi) \wedge A\phi] \rightarrow A\psi$

Index des Auteurs

Aho : 239f.
Åqvist : 188-9f., 195f.
Aristote : i-xii, xv-xvii, 4, 5, 7, 31, 130, 168f., 179f., 255, 326f., 338, 376
Arner : 208, 209f., 210, 211
Austin : xii, xvi, 158, 185f., 280f., 323f., 343
Barcan Marcus : 6, 41, 221f., 228, 233f., 234, 245f., 246, 292f., 294, 374f.
Barwise & Perry : 210
Bäuerle & Cresswell : 294f.
Bazhanov : 7f., 275f., 300
Beall : 285f., 298f.
Beyer : 111f.
Béziau : 16f., 17f-18f., 33f., 272f., 284f.
Black : 198, 324f.
Blanché: 5f., 32, 33f., 130, 195, 319, 367f.
Bochvar : 6, 7f., 189f., 300, 331, 336
Bovens : 190f., 331
Brogaard & Salerno : 261f., 263
Bull & Segeberg : 3
Cargile : 197
Carnap : 6, 19-24, 29-30, 31f., 39, 43f., 48, 91, 102f., 123, 125, 128, 131, 132, 135, 137-8, 155f., 172, 187, 228, 230, 237f., 250f., 268, 272, 298
Castañeda : 27f., 148f., 191f., 221f.
Chalmers : 160f.
Chisholm : 4f., 181, 187, 208, 221f., 234, 240f., 242f., 243-5, 247f., 258, 354f., 373f.
Church : 29, 30, 250f., 261f.
Copeland : 18f.
Costa-Leite : 303f.
Cresswell : 9f., 154f., 165
da Costa & Bueno : 141f.
da Costa & Bueno & French : 296f., 300f.
da Costa & French : 284, 285, 294f., 300f.
Davidson : 270-3, 275-6, 279f., 296f.
De Sousa : 179
Donnellan : 111f., 232, 233f., 234, 274f.
Dretske : 209
Dubucs : 61f., 163, 177f., 311f.
Dubucs & Marion : 142f., 177f.
Dugundji : 17-8
Dummett : 35-6, 110f., 133f., 256, 257, 261, 267, 276f.
Eberle: 150
Ellis : 58, 107, 114-22, 207f., 268-9, 286f.
Engel : 59, 108, 113, 114f., 152f., 200, 285, 275f., 277f.
Englebretsen : 321, 322f.
Fagin & Halpern : 170, 178
Favrholdt : 93, 95, 96f., 97f., 98
Ferret : 245

Fisher : 146f., 157f.
Fitch : xviii, 190f., 196, 261f., 262-7, 303f., 342
Føllesdal : 225-6, 234
Frege : 30, 36, 38, 50, 54, 56, 58-60, 62, 63f., 64-74, 76, 77f., 79-82, 84, 92, 94, 95, 97f., 98, 101f., 106-10, 113-7, 121, 123f., 124-5, 129f., 131-4, 135f., 136-7, 154, 158f., 172f., 218-9, 228, 231, 234, 236f., 239, 240, 247, 248f., 250f., 288f., 302-3, 304f., 305-11, 306-8f., 314-8, 326-7, 331, 333, 334f., 341-2, 376-7, 394, 396, 398-9
French : 269f., 285, 297
Gardenförs : 114, 116f., 118f., 121-2, 291f.
Gardiès : 31, 33f., 162, 141f., 160f., 162f., 303f., 326f., 331, 332f.
Gettier : 209, 254, 257-60, 302, 342
Girle : 169-70, 202, 221f.
Gochet : 31, 39, 40f., 74f., 237f.
Gochet & Gribomont & Thayse : 144f.
Gödel : 11, 17, 264, 280, 329, 330, 334, 337, 394
Granger : 46, 105, 142, 144
Greimann : 331f.
Haack : 140
Hanson : 80, 187, 312f., 331
Harnish : 185, 186f.
Hart : 32, 324f., 327, 334, 353-74
Heal : 185f.
Henderson : 295f., 296f.
Hendricks : 8f., 29f., 207
Hilpinen : 307f.
Hintikka : 1-3, 6, 8-9, 18, 20, 21-9, 33-8, 42-4, 46-9, 51-2, 57, 60, 68, 71, 82-4, 86-7, 89-90, 92-9, 102, 105, 110-4, 117, 121-5, 128-41, 145, 146-62, 164-70, 172-92, 195-6, 198-204, 206-8, 210-4, 216, 219-22, 224-30, 234, 236-52, 254, 258, 260-2, 264, 269, 271, 275, 286-7, 291, 295, 309-10, 313, 316, 319, 321, 325, 333, 335, 340, 343-4, 354-5, 358-60, 366-7, 369-74, 404
Hinzen : 112, 266-7
Hocutt : 53, 63, 144, 150f., 152f., 154-5, 156f., 169, 180, 187-8, 196, 212, 309, 373
Jaskowski : 192f.
Kanger : 2, 18, 23
Kearns : 16f., 303f., 312-3, 317, 318-9f., 322, 324, 333, 376-404
Kielkopf : 63, 145-6, 151-5, 188
Klima : 324f.
Kripke : 2-3, 10-1, 17-8, 19f., 20-5, 28, 31f., 33, 41-2, 43f., 52, 63, 87, 136, 138, 141, 144-5, 148, 162, 186, 225-6, 227f., 228, 233f., 236f.
Laugier & Bonnay : 298f.
Lehrer : 195
Lehrer & Paxson : 258
Lemmon : 13, 14f., 67f., 145, 146f., 155-6, 201, 203-5, 225f., 242f., 369, 403
Lenzen : 13, 151f., 157f., 167f., 207f., 221-2f., 233f., 247f., 259f., 265-6, 324,
Lewis C. I. : 5, 8, 9-11, 13, 17, 33-4, 54, 100, 105, 143f., 145, 163, 236, 309, 367, 369
Lewis D. : 42-3, 48-50, 210f., 240, 293, 294f.
Linsky : 149
Lokhorst : 282f.
Łoś : 13
Łukasiewicz : 6-7, 14, 16f., 49, 130f., 208f., 279, 280f., 307f., 317f., 326

Malcolm : 208-9
Malinowski : 331
Martin-Löf : 112, 113f., 267f.
Mates : 250f.
Meyer : 3, 140f., 171
Mitton : 200
Mohanty : 110-1, 244f.
Montague : 2, 136, 343
Moore : xviii, 98, 184-200, 209, 214, 216, 264, 284-5, 294, 301-2, 319-20, 324, 331, 342, 357, 366, 369-70, 392-4
Moretti : 33f.
Morick : 243f.
McGinn : 290f.
McLane : 156f.
Nef : 42, 244f.
Niiniluoto : 243f.
Nolte : 49
Nozick : 163, 210, 236
Paoli : 147f., 271, 288f.
Pap : 3
Peirce : xii, 6f., 76, 132, 140, 158f., 199f., 307f.
Priest : 7f., 168f., 272f., 282f., 285, 299-300
Prior : 374
Pritchard : 181, 211
Purtill : 16f., 244-5
Putnam : 301f.
Quine : iii, xvi, 5f., 6, 14, 17f., 30, 31f., 33, 37, 38f., 39-43, 47, 51, 53, 60, 63, 67, 73, 87, 99-107, 117, 118-9f., 123, 127-8, 129f., 132-8, 143-4, 146, 147f., 151, 156, 158-9, 206, 215, 217-8, 224, 227-41, 243, 245-7, 248f., 250, 257f., 260, 267-73, 275-85, 287f., 288-91, 294f., 296, 298-9, 301, 336, 341, 396
Radford : 180-1, 195, 203
Rantala : 174, 177-8, 296,
Rescher : 3, 75f., 146f.
Restall : 1, 137f., 298f.
Russell : ii, ix, xvi, 6f., 37, 39, 49f., 54, 63, 67, 75-88, 92-3, 95, 97f., 98, 100, 103, 106, 110, 120, 123f., 124, 132-3, 136, 137f., 160f., 200, 232, 234, 236f., 238f., 239, 247f., 302f., 303-5, 308, 309f., 310, 313, 322-3, 325-9, 332f., 333-5, 337-8, 341, 345-52, 394, 396, 398-9
Ryle : xii, 158
Salerno : 261f.
Scott : 207, 210
Searle : 160f., 185, 187, 189f., 200, 239, 325, 396, 398, 400
Searle & Vanderveken : xii, 248, 303, 311, 313, 324-5, 329, 333-4, 338, 343, 378
Smith : 226f., 244
Strawson : ix-xii, 1, 39, 158, 215, 236f.
Suszko : 36f., 319f., 337f.
Tarski : 6, 22, 33f., 58, 85f., 102, 107, 119f., 132-3, 135, 146-7, 160f., 230, 268, 276, 377
Thagard & Nisbett : 287f., 295f., 297
Triplett : 300f.
Urmson : 211
Upton : 4f.

van Benthem : 24f., 112, 137, 197f., 263f., 265f., 343
van Heijenoort : 124f., 129-31, 134f., 136-7f., 138, 402
Varzi : 102f.
Vernant : xvi, 33f., 178, 252-3, 303, 304-7f., 308-9, 313, 317, 318f., 323-5, 326f., 327-33, 335, 336f., 352, 368, 371
Von Wright : xii, 3, 5f., 9f., 11f., 18, 91f., 298, 340, 343
Wansing : 262-3
Wertz : 323-4f.
Williams : 15, 191f., 197, 256, 263, 284-5, 288, 289-90f.
Williamson : 112f., 256, 263
Wiredu : 256f., 281f., 316f.
Wittgenstein : 18f., 19-21, 37-8, 44, 47, 49f., 54, 63, 66, 86, 88-99, 103, 106, 113f., 129f., 131-4, 136, 157-8, 161, 172, 185f., 208-11, 239, 300f., 341, 343
Zemach : 203, 205, 212-3, 214f., 215, 218f., 247, 248f., 250, 251f., 254, 256, 260, 264, 302, 309f.,

SCD NANCY 2
46, av. de la Libération
B.U. LETTRES

Philosophie des Modalités Epistémiques

La logique assertorique revisitée

Résumé :

La valeur d'une analyse logique réside dans sa capacité à résoudre des *paradoxes* et à comprendre l'origine de nos problèmes conceptuels ; à ce titre, le rôle d'une *logique épistémique* est de traiter des paradoxes liés au concept de connaissance.

On présente généralement la logique épistémique comme une analyse logique de concepts centraux en épistémologie : connaissance, croyance, vérité, justification. Une autre approche sera proposée ici en vue de clarifier ce genre de discours, centrée sur la notion d'*assertion* et décrite en termes d'*énonciation*.

Parce que les modalités épistémiques expriment des attitudes, c'est l'intentionnalité du discours qui sera mise en valeur dans le cadre d'une logique *modale illocutoire*. Deux thèses transversales parcourront l'ensemble du travail : les énoncés déclaratifs et épistémiques partagent la même logique (logique assertorique) ; la pluralité des jeux de langages n'implique pas un pluralisme logique.

Philosophy of Epistemic Modalities

Assertion logic revisited

Summary :

The relevance of any logical analysis lies in its ability to solve *paradoxes* and trace conceptual troubles back; with this respect, the task of *epistemic logic* is to handle paradoxes in connection with the concept of knowledge.

Epistemic logic is currently introduced as the logical analysis of crucial concepts within epistemology, namely: knowledge, belief, truth, and justification. An alternative approach will be advanced here in order to enlighten such a discourse, as centred upon the word *assertion* and displayed in terms of *utterance*.

Insofar as epistemic modalities express some attitudes, the intentionality of discourse will be emphasized within an *illocutionary modal* logic. Two large views will range over the whole study: declarative and epistemic sentences have one and the same logic (assertion logic); the plurality of languages games doesn't entail any logical pluralism.