Département de formation doctorale en informatique
UFR STMIA

École doctorale IAE + M

Thèse

présentée pour l'obtention du titre de

Docteur de l'Université Henri Poincaré, Nancy 1

en Informatique

par

Susanne Salmon-Alt

Référence et dialogue finalisé : de la linguistique à un modèle opérationnel

Soutenue publiquement le 16 mai 2001

Membres du jury

Président :	Jean-Marie Pierrel	Professeur, Université H. Poincaré, Nancy 1
Rapporteurs :	Jean Caelen	Directeur de Recherche CNRS, IMAG Grenoble
	Francis Corblin	Professeur, Université Paris 4 – Sorbonne
Examinateurs :	François Lamarche	Directeur de Recherche INRIA, Loria, Nancy
	Norbert Reithinger	Principal Researcher, DFKI Saarbrücken, Allemagne
Invité :	Jacques Moeschler	Professeur, Université de Genève, Suisse
Directeur de thèse :	Laurent Romary	Chargé de Recherche CNRS (HDR), Loria, Nancy

Laboratoire Lorrain de Recherche en Informatique et ses Applications - UMR 7503

MERCI...

Je me fais un plaisir de remercier Laurent pour m'avoir attendu un jour de février pluvieux à la gare de Nancy. Et depuis, pour son enthousiasme et ses encouragements qui m'ont accompagnés tout au long des trois dernières années. Parmi toutes les choses que j'ai apprises en travaillant avec lui, il y en a une que je n'oublierai pas : c'est qu'« il faut savoir être patient avec soi-même »...

Merci aussi à Jean-Marie pour la confiance qu'il m'a accordée en me proposant l'intégration dans l'équipe Langue et Dialogue. En tant que linguiste-informaticienne, je m'y suis toujours sentie à ma place.

Je tiens à remercier Jean Caelen et Francis Corblin qui m'ont fait le plaisir d'être rapporteurs pour les retours aussi bien informatiques que linguistiques que j'ai reçus de leur part. Merci bien sûr aussi à François Lamarche pour avoir accepté d'examiner cette thèse en tant que rapporteur interne et à Jaques Moeschler pour l'intérêt avec lequel il a suivi l'évolution de ce travail. Et un « Dankeschön » particulier à Norbert Reithinger pour avoir examiné le manuscrit en français !

Mais au fait, avant ce jour pluvieux du mois de mai où tout s'est terminé, il y avait des gens sans qui ce manuscrit n'aurait pas vu le jour sous sa forme actuelle : je pense en particulier à Laurence qui a eu la gentillesse de relire les bouts successifs de ma thèse entre des aller-retour Paris-Nancy. Je pense à Nadia, spécialiste de la détection des fautes d'orthographe dans les notes de bas de page, à Hélène, impitoyable avec les conjonctions en début de phrase, à Fred, inconditionnel des espaces insécables (et des étiquettes recto-verso), à Patrice, traqueur des fautes de frappe dans les algorithmes et à Olivier pour son aide considérable sur la dernière partie. Plus largement, j'aimerais remercier celles et ceux qui m'ont accueilli, aidé et supporté dans tous les sens du terme : Jean-Luc, Armelle, André, Yannick, Sandrine, Fred (Wolff), Florence, les deux Patrices, Bertrand, Evelyne, Anne, Azim et tous les autres... sans oublier Isabelle, Chantal et Noëlle.

Enfin, quelque fût le temps, je remercie de leur intérêt et de leur soutien à distance ma famille et mes amis berlinois ainsi que Uta au Québec. Et si je ne suis pas encore recluse au fond des Quatre Montagnes, c'est grâce à Stéphane qui s'est fait embarquer dans cette histoire et qui s'est tout simplement révélé un merveilleux mari !

SOMMAIRE

INTRODUCTION

PREMIÈRE PARTIE

DEUXIÈME PARTIE

TROISIÈME PARTIE

CONCLUSION

1 Thème et variations

1.1 Linguistique et dialogue finalisé : vers la conception d'interfaces plus naturelles

La conception d'interfaces homme-machine conviviales est devenue une des préoccupations essentielles de l'informatique. Si l'exécution de tâches simples et bien définies s'accommodait d'interfaces limitées aux menus hiérarchiques, formulaires ou lignes de commande, la complexité croissante des tâches effectuées par les machines – assistance dans des activités de raisonnement, recherche d'informations, résolution de sous-problèmes, proposition d'outils pour la réalisation d'un but concret – s'accompagne d'un changement profond des modes d'interaction homme-machine. La langue naturelle apparaît alors comme un outil de communication idéal : non seulement elle s'adapte facilement à de nouvelles situations de communication, mais elle est aussi le seul outil capable de modifier dynamiquement la situation de communication et de se prendre elle-même comme thème du discours (Sabah, 1997).

Depuis une vingtaine d'années, l'intégration de la langue naturelle dans les interfaces fait partie des thèmes de recherche centraux en informatique. Grâce à cette dynamique, nous disposons aujourd'hui d'un certain nombre d'applications effectives du dialogue homme-machine, en particulier dans des domaines bien définis comme l'accès à des bases de données (systèmes de réservation de vols – GUS, Bobrow et al., 1977) ou la commande de systèmes simples (gestion de messagerie électronique – PARTNER, Morin et Pierrel, 1987), édition graphique – ICP-Draw, Caelen, 1991). Le principe sur lequel reposent ces applications est le suivant : à partir d'un énoncé en langue naturelle, on construit, en passant par des analyses syntaxiques plus ou moins sophistiquées, une représentation interne du sens, le plus souvent proche de la logique du premier ordre. Cette représentation sert ensuite de base, par exemple, pour l'instanciation d'une requête SQL[1] (dialogues de renseignement) ou d'un schéma d'action (dialogues de commande). Si la conception de telles applications a été encourageante dans un premier temps, les limites des systèmes développés sont essentiellement dues au fait qu'ils reposent sur l'hypothèse simplificatrice d'un parallélisme entre les structures d'une application informatique et les structures langagières, voire cognitives (Duermael, 1994 ; Sabah, 1997). Cela a pour conséquence immédiate la limitation de la partie « naturelle » de la langue, dans la mesure ou celle-ci est réduite à un sous-ensemble de vocabulaire et de constructions acceptées. En fin de compte, l'utilisateur se trouve dans une situation peu confortable : il doit apprendre, une fois de plus, à respecter les limites de la machine et ceci pour interagir dans des domaines dont il doit tellement bien connaître les structures, qu'il est permis de s'interroger sur la réelle utilité de la langue naturelle dans ces cas-là.

Or, la flexibilité qu'offre la langue naturelle peut certainement être mise à profit dans l'interaction homme-machine. Encore faut-il que l'on ne réduise pas *a priori* les possibilités d'expression de la langue naturelle par sa réduction à des structures compatibles avec celles du système. Le qualificatif « finalisé » ne s'applique donc pas à la composante langagière du dialogue, mais bien aux situations de communication : en tout état de cause, il n'est pas pour l'instant réaliste d'envisager des conversations « libres » avec une machine. Nous nous intéressons donc ici à des situations qui s'appuient sur des applications clairement définies : celles-ci comportent un gestionnaire informatique qui inclut une représentation visuelle de l'état de l'application et une interface de commande que l'on souhaite encapsuler dans un système de dialogue naturel assurant la communication avec l'usager. Mais à la différence des systèmes mentionnés ci-dessus, les structures langagières autorisées n'ont pas à être isomorphes aux structures des commandes élémentaires du gestionnaire. Pour aller plus loin, il faut au contraire envisager un intermédiaire dont le fonctionnement soit le plus analogue et compatible avec

[1] *Standard Query Language*, langage normalisé d'interrogation de bases de données

celui de la cognition humaine. C'est la seule voie qui permettra à l'utilisateur de se libérer de toute contrainte sur le « comment faire » pour se concentrer exclusivement sur le « quoi faire » (Pierrel et Romary, 1997).

Une telle redéfinition des objectifs de la conception d'interfaces naturelles déplace considérablement l'enjeu des recherches subséquentes. Contrairement à des travaux relevant plutôt du génie logiciel, nous visons l'étude des possibilités offertes par une communication spontanée de la part des utilisateurs avant de rendre compte des capacités de calcul des machines. Il apparaît alors que cette approche se situe à l'intersection de plusieurs disciplines, ayant pour objet d'étude les facultés cognitives spécifiquement humaines : psychologie, logique, intelligence artificielle, neurosciences et linguistique. A l'intérieur d'un tel cadre, le rôle de la linguistique consiste à identifier les phénomènes caractéristiques du langage humain, à travers l'étude des langues dans toute leur richesse et à en donner des descriptions plus ou moins formalisées. Elle met ainsi en évidence les règles et inventaires qui fondent la compétence linguistique d'un sujet. Or, ce sont précisément ces règles qui devront remplacer certaines heuristiques intégrées dans des systèmes de dialogue actuels, rendant difficile leur extension et leur réutilisabilité.

Enfin, l'étude du dialogue homme-machine dans le cadre plus vaste du traitement automatique des langues permet de percevoir des phénomènes qui sont parfois négligés ou oubliés par les modèles classiques. Parmi ces phénomènes, nous avons choisi de nous concentrer sur les problèmes de la référence aux objets, posés avec acuité dans le dialogue homme-machine qui ne peut traiter du langage hors situation. Notre travail est alors motivé par la nécessité constante de déterminer les objets dont parle l'utilisateur pour ensuite être capable d'exécuter une action ou une requête. Le fait que ces objets peuvent être accessibles dans l'environnement visuel ou même désignés par des gestes rend nécessaire la prise en compte du caractère multimodal d'une grande partie des interactions homme-machine. Enfin, le problème de la référence dans le dialogue homme-machine présente l'intérêt de remettre en cause, à travers la nécessité d'une représentation profonde commune, la distinction traditionnelle entre processus de compréhension et de génération.

1.2 Référence et linguistique : le « paradoxe de la reprise immédiate »[2]

La modélisation du calcul de la référence aux objets peut être considérée comme la modélisation d'une partie du fonctionnement cognitif de l'être humain : celle qui consiste à mettre « *en relation une expression de la langue (dite en générale expression référentielle) en emploi dans un énoncé et l'objet dans le monde que cette expression désigne* » (Reboul et Moeschler, 1994 : 535). La notion d'objet doit être considérée au sens large. Il peut s'agir d'un objet concret ou abstrait, d'un objet fictif ou encore d'un groupe d'objets. Les expressions référentielles, par lesquelles ces objets sont désignés, se trouvent matérialisées par des syntagmes d'une grande variété : groupes nominaux à différentes déterminations, groupes nominaux sans ou avec tête nominale et groupes pronominaux. L'objectif du linguiste consiste alors à étudier le système de ces catégories linguistiques associées à l'interprétation référentielle.

Il nous semble que les études linguistiques consacrées aux expressions impliquées dans l'interprétation référentielle constituent un point de passage obligatoire pour qui veut modéliser le calcul référentiel en vue d'un traitement automatique. Tout d'abord, elles présentent l'avantage de décrire les données linguistiques, observées ou imaginées par le linguiste, dans toute leur richesse et sans poser de simplifications *a priori*. Mais au-delà d'une simple accumulation de données, l'intérêt consiste en la proposition de régularités associées à l'usage d'une forme particulière. Le souci

[2] expression forgée par F. Corblin (1983)

essentiel des linguistes s'intéressant aux expressions référentielles est de dégager un lien entre le type d'une expression et les conditions de son usage. Le point commun de beaucoup de travaux linguistiques est de partir de l'hypothèse que chaque expression représente « *un marqueur référentiel original [...], un outil référentiel* per se *qui a ses propriétés identificatoires propres, non réductibles à celles des autres marqueurs proches* » (Kleiber, 1994 : 41). Sur la base de cette hypothèse, l'objectif de la description linguistique est de pouvoir prédire l'apparition de telle ou telle expression, étant donné telles ou telles circonstances. Nous illustrerons par la suite cette approche par un débat – celui de la reprise immédiate – que nous considérons comme exemplaire : il permet à la fois d'appréhender les approches linguistiques de façon concrète et d'en faire apparaître les limites.

1.2.1 Le paradoxe de la reprise immédiate

Le débat concernant le « *paradoxe de la reprise immédiate* » pose le problème du rôle joué par différents types de marqueurs référentiels dans différents contextes discursifs. A la suite des travaux de G. Guillaume (1919), C. Blanche-Benveniste et A. Chervel (1966), E. Padučeva (1970) et J.-C. Milner (1982), est considérée comme « paradoxe » la répartition des déterminants définis et démonstratifs dans les anaphores fidèles[3] reprenant un syntagme indéfini. Les configurations typiques sont celles de (1) et (2) :

(1) J'ai vu une voiture. ? La voiture / Cette voiture roulait vite.

(2) J'ai vu une voiture et un camion. La voiture / ? Cette voiture roulait vite.

Ce qui a été jugé comme un paradoxe, c'est que le défini semble exclu du site (1) et obligatoire en (2), alors que la situation se présente à l'inverse en ce qui concerne le démonstratif.

- *L'explication en termes de contrastes*

Il ressort d'une série de travaux de F. Corblin (1983, 1987, 1995) que l'on n'assiste pas à une distribution complémentaire stricte des deux types de déterminants : aussi bien le défini que le démonstratif sont souvent possibles dans les mêmes configurations. Le problème n'est donc pas un problème d'ambiguïté référentielle (il n'y a aucun problème pour identifier le référent), mais plutôt un problème de présentation optimale des référents à l'intérieur d'un cadre donné. La question qui se pose alors est de savoir quelles sont les constantes linguistiques qui assurent une telle optimalité.

L'hypothèse avancée par F. Corblin repose sur une différence dans la saisie référentielle du défini et du démonstratif :

(a) Le défini *le N* saisit son référent grâce à son contenu descriptif, par signalement singularisant. Comme *N* est le seul principe permettant de singulariser ou d'identifier le référent dans un ensemble d'entités de référents potentiels, le défini instaure alors de fait un contraste entre l'individu vérifiant la propriété *N* et les autres, qui ne la vérifient pas.

(b) Le démonstratif *ce N* saisit son référent en vertu de critères externes (mention récente, proximité, saillance). Cette entité est alors reclassifiée comme étant un *N*, ce qui suppose une extraction d'une valeur de la classe des *N*. De ce fait s'instaure un contraste entre un *N* particulier (le référent) et d'autres *N*.

[3] On parle d'anaphore lorsque la saturation interprétative d'une expression est dépendante d'une autre expression qui la précède (*J'ai vu une voiture. Elle / l'engin / ce bolide roulait vite.*). Par anaphore fidèle, on entend la reprise d'une mention par une expression possédant la même tête lexicale (*J'ai vu une voiture. Cette voiture roulait vite.*).

Il convient de souligner que cette notion de contraste n'est jamais posée comme fournissant le principe de fonctionnement des définis et démonstratifs : il s'agit d'une conséquence du mode d'identification propre à chacun des marqueurs référentiels (Corblin, 1995 : 74). Toujours est-il que dans certaines configurations caractéristiques, considérées hors contexte d'usage, cette notion permet d'expliquer les intuitions que l'on peut avoir sur les deux volets du « paradoxe » :

(3) J'ai vu une voiture. ? La voiture / Cette voiture roulait vite.

Ainsi, les intuitions face au premier volet du paradoxe (exclusion du défini) se justifient par le fait que l'emploi du défini apparaît non motivé, en l'absence d'un ensemble contextuel hétérogène à l'intérieur duquel pourrait s'installer une opposition de type *voiture* vs. *non voiture*. L'emploi d'un démonstratif, en revanche, passe mieux, car en l'absence d'un domaine empirique, il semblerait que l'opposition puisse toujours s'instaurer à l'intérieur de la classe des *N*[4].

(4) J'ai vu une voiture et un camion. La voiture / ? Cette voiture roulait vite.

Pour ce qui est du deuxième volet du paradoxe (exclusion du démonstratif), il s'explique par la mise à disposition d'un ensemble contextuel explicite, à l'intérieur duquel le défini peut en effet opérer une opposition de type *voiture* vs. *camion*. L'usage du démonstratif semble ici peu motivé, car s'il permet d'identifier sans problème le référent, il l'extrait de son domaine contextuel pour l'insérer dans un nouvel ensemble (virtuel, qui plus est).

- *L'explication par le prolongement des circonstances d'évaluation*

A la suite des travaux de F. Corblin (1983), G. a publié une série d'articles (1986, 1988, 1989) dans laquelle il défend une position concurrente sur le « paradoxe » : la motivation de cette nouvelle position repose essentiellement sur l'observation qu'une application systématique[5] (et quelque fois caricaturale) de la notion de contraste n'arrive pas à rendre compte d'un certain nombre d'exemples, sur lesquels l'auteur émet par ailleurs des jugements d'acceptabilité relativement tranchés. Le Tableau 1 en donne quelques exemples :

Exemples	***Prédiction par contraste***[6]	***Jugement de G. Kleiber (1988)***	***LE***	***CE***
Un prince aimait une princesse. ? prince aimait aussi les fleurs.	LE	CE (LE « moins naturel », p. 77)	37 %	63 %
Dans mon jardin, il y a un cerisier. ? cerisier a été planté par mon père.	LE	CE (LE « interdit », p. 78)	21%	79%
Un avion s'est écrasé hier. ? avion venait de Miami.	CE	LE (CE « nettement moins satisfaisant », p. 78)	47%	53%

Tableau 1 : Comparaison de quelques prédictions sur le « paradoxe de la reprise immédiate »

A propos de ces exemples, deux remarques s'imposent : premièrement, les prédictions en termes de contraste sont abusives, dans la mesure où la notion de contraste est ici considérée comme fournissant le principe même de l'interprétation. Or, F. Corblin précise bien que seuls des contextes « *en*

[4] La reprise par le démonstratif paraît toutefois moins appropriée par rapport à une reprise pronominale : nous l'expliquerons dans la suite par le fait que le démonstratif n'opère ici ni une reclassification, ni une opposition à d'autres *N* d'un domaine explicite. De ce fait, il a effectivement le même effet interprétatif que le pronom.

[5] L'application systématique (et caricaturale) du principe de contraste amène à prédire l'apparition d'un démonstratif lorsqu'un seul référent est introduit dans la phrase précédente et celle d'un défini lors de l'introduction de plusieurs référents.

[6] Attention : il s'agit de l'interprétation que donne G. Kleiber (1986) des travaux de F. Corblin (1983).

opposition manifeste avec des contrastes inhérents aux catégories peuvent donner matière à des jugements nets directement motivés par des phénomènes de contrastes » (1995 : 75). Les exemples (1) et (2) peuvent être considérés comme de tels contextes, mais il est évident que l'interprétation de la grande majorité des occurrences naturelles demande la prise en compte d'autres facteurs que la seule notion de contraste. Une deuxième remarque concerne les jugements d'acceptabilité de G. Kleiber : il nous a semblé que ceux-ci doivent être manipulés avec prudence. Nous les avons donc vérifiés sur un ensemble de 68 locuteurs, auxquels nous avons demandé d'indiquer la meilleure des deux variantes de reprise. Les résultats de cette enquête[7], reproduits dans les deux dernières colonnes du tableau, laissent en effet penser que la frontière entre ce qui est « naturel » et « moins naturel » n'est pas aussi nette.

Toujours est-il qu'à partir de ses jugements, G. Kleiber (1986) propose une nouvelle explication du « paradoxe », fondée sur la différence entre saisie indirecte (défini) et saisie directe (démonstratif) :

(a) L'article défini réfère indirectement au référent à travers une circonstance d'évaluation constituée par la phrase introductrice (p1). Il figure en reprise immédiate lorsque le locuteur, dans la seconde phrase (p2), entend parler du référent introduit dans la première phrase comme étant celui qui vérifie la condition d'être un *N* dans la circonstance tracée par p1.

(b) L'adjectif démonstratif renvoie directement au référent, par l'intermédiaire obligé du contexte d'énonciation de son occurrence. Il appréhende le référent comme un objet non nommé reclassifié. *Ce* est utilisé donc en reprise immédiate lorsque le locuteur entend parler directement du référent, indépendamment de la circonstance d'évaluation tracée en p1.

Cette hypothèse ne met donc pas en jeu un contraste éventuel entre entités contextuelles, mais le rôle prépondérant des cadres situationnels posés par les deux phrases p1 et p2 : elle postule en particulier qu'une anaphore définie n'est possible que s'il y a prolongement en p2 de la circonstance évaluative tracée par p1. Une anaphore démonstrative est toujours possible, mais elle instaure une rupture avec p1. L' « algorithme » que propose G. Kleiber pour prédire le fonctionnement du défini et du démonstratif en reprise immédiate serait donc le suivant :

```
si p1 fournit une circonstance d'évaluation pertinente alors
      si p2 prolonge cette circonstance alors
            "LE"
      fin si
sinon
      "CE"
fin si
```

Figure 1 : « Algorithme » pour la décision entre CE et LE en reprise immédiate (Kleiber, 1986)

Il apparaît tout de suite que cet algorithme est inopérant tant que les notions de « circonstance d'évaluation » et de « prolongement de circonstance » ne sont pas précisées plus avant. G. Kleiber aborde cet aspect essentiellement à travers des exemples dont certains sont reproduits dans le Tableau 2 : il apparaît que la « circonstance d'évaluation » se définit à travers l'éventualité introduite en p1. En fonction de différentes caractéristiques, celle-ci est dite « circonstance d'évaluation pertinente » ou non. Lorsqu'elle n'est pas « circonstance d'évaluation pertinente », elle est considérée comme non prolongeable et une reprise par CE s'impose. Autrement, les caractéristiques de l'éventualité p2 déterminent s'il s'agit d'un prolongement ou non.

[7] Nous y reviendrons de façon plus détaillée dans la section suivante.

Circonstance d'évaluation pertinente	*Caractéristiques de p1*	*Exemple de p1*	*p2 prolonge p1 => LE*	*p2 ne prolonge pas p1 => CE*
non	impersonnelle existentielle	Dans mon jardin, il y a un cerisier.	[impossible]	CE cerisier a été planté par mon père.
	événement « standard »	Un chien aboie.	[impossible]	CE chien a faim.
oui	compte rendu de perception	Le tableau représente une jeune fille.	LA jeune fille sourit.	CETTE jeune fille s'habille habituellement en vert.
	s'inscrit dans une scène ou une narration plus vaste	Un homme entra.	L' homme dit bonjour.	J'avais déjà vu CET homme chez un ami.
	événement jugé « exceptionnel »	Un avion s'est écrasé hier.	L'avion venait de Miami.	CET avion reliait habituellement New York à Miami.
	introduction de plusieurs référents	Un prince aimait une princesse.	LE prince était grand et fort.	CE prince aimait aussi les fleurs.

Tableau 2 : Explication de la distribution CE/LE selon Kleiber (1986, 1990)

Si tant est que l'on puisse juger réellement de la pertinence d'une circonstance d'évaluation ainsi que d'une éventualité prolongeante, cette hypothèse permet en effet de prédire la distribution des déterminants du Tableau 1 selon G. Kleiber. Essayons à présent de la mettre à l'épreuve sur l'objet initial du débat, à savoir le paradoxe de la reprise immédiate dans des exemples tels que (5) et (6) :

(5) J'ai vu une voiture. ?La voiture / Cette voiture roulait vite.

La première question est donc de savoir si p1 constitue une condition d'évaluation pertinente. Si l'on le considère comme introduisant un « événement standard », la réponse serait alors « non ». Mais il serait aussi possible de penser à un « compte rendu de perception » et la réponse serait alors « oui ». La deuxième question est de savoir si p2 peut être considéré comme un prolongement du cadre instauré en p1 ? Là, nous manquons réellement d'indices qui permettraient de l'affirmer. Les éléments de réponse ne sont donc pas vraiment évidents : tout au plus, raisonne-t-on par analogie aux exemples.

(6) J'ai vu une voiture et un camion. La voiture / *Cette voiture roulait vite.

Concernant le deuxième volet du paradoxe, il est traité indépendamment des conditions d'évaluation ou du contexte d'énonciation : ici, il s'agit de facteurs syntactico–sémantiques qui empêchent la reprise par *CE*.

1.2.2 Enquête sur les exemples

Étant donné la durée et la ferveur des discussions autour d'un nombre restreint d'exemples[8], majoritairement construits par les auteurs eux-mêmes, il nous a semblé intéressant de vérifier les jugements d'acceptabilité avant de nous prononcer sur l'apport de chacune des deux hypothèses. Nous avons donc repris, des articles de G. Kleiber (1986, 1988)[9], 18 exemples autour desquels s'organise l'argumentation. Ces exemples ont été présentés dans les deux versions (reprise immédiate au défini vs. démonstratif) à 68 locuteurs, étudiants de première année à l'université de Nancy 2. Nous leur avons demandé de choisir, parmi ces deux variantes, celle qui leur paraissait la plus naturelle.

Cette enquête n'était en aucun cas conçue en tant qu'expérience psycholinguistique : notre seul objectif a été de recueillir les intuitions d'un plus grand nombre de personnes sur les exemples. Le

[8] Le champ d'étude s'est tout de même ouvert aux occurrences littéraires, en particulier avec les travaux de F. Corblin (1995).
[9] G.Kleiber reprend lui-même certains de ses exemple de F. Corblin (1983).

détail des résultats est donné dans le Tableau 4, page 14. Il regroupe les exemples en quatre blocs : un premier, pour lequel aussi bien la théorie de G. Kleiber que l'application (caricaturale, toujours) de la notion de contraste prédisent, pour un effet optimal, l'apparition d'un démonstratif ; un deuxième, pour lequel G. Kleiber prédit l'apparition du défini et la structure contrastuelle des énoncés du démonstratif ; un troisième, pour lequel les deux hypothèses prédisent l'apparition du défini ; enfin, un quatrième, pour lequel G. Kleiber prédit l'apparition du démonstratif et la structure contrastuelle des énoncés du défini. Enfin, G. Kleiber ne se prononce pas sur un dernier exemple, pour lequel la structure contrastuelle permet de prédire l'apparition du démonstratif. Les deux dernières colonnes du tableau indiquent le pourcentage des réponses en faveur du défini et du démonstratif, respectivement.

A partir de ce tableau, il est possible de dégager quelques grandes tendances :

- Tout d'abord, toutes questions confondues, les décisions en faveur du démonstratif sont majoritaires, avec une proportion de 65 %, alors que les deux hypothèses prédisent l'apparition d'un démonstratif pour seulement la moitié des cas (9 selon le contraste, 7 selon les circonstances d'évaluation). En faisant abstraction des applications caricaturales des deux principes concurrents, cet avantage général pour le démonstratif s'accorde bien avec les hypothèses sur la saisie référentielle des deux marqueurs : à l'origine des principes, F. Corblin (1983) caractérise bien le démonstratif en tant que « *anaphorique positionnel local* », c'est-à-dire comme reprise d'une mention donnée nécessairement par le contexte. G. Kleiber (1986) le considère comme « *renvoyant directement, c'est-à-dire abstraction faite de toute circonstance d'évaluation, au référent introduit par le SN*[10] *indéfini* ». Or, ces conditions sont effectivement remplies dans tous les exemples : l'introduction textuelle d'un référent met à disposition un candidat à reprise immédiate par un démonstratif. Dans aucun cas, il n'y a de difficulté d'interprétation référentielle et au pire, le démonstratif indiquerait, pour reprendre les termes de G. Kleiber (1986 : 56), « *une sorte d'excès référentiel dans l'identification* », ressenti comme « *un zèle référentiel trop poussé* ».

- Dans un deuxième temps, un examen plus attentif des regroupements permet de constater que là où les deux principes s'accordent sur les prédictions, les résultats confirment ces prédictions (Tableau 3). Ceci est particulièrement flagrant pour le bloc LE/LE : il est le seul à avantager systématiquement la reprise par le défini. En revanche, en cas de désaccord, c'est le démonstratif qui reprend le dessus.

Enquête auprès de 68 locuteurs		***Prédiction par prolongement des circonstances d'évaluation***	
		CE	***LE***
Prédiction par contraste	***CE***	CE : 78 % LE : 22 %	CE : 58 % LE : 42 %
	LE	CE : 77 % LE : 23 %	CE : 34 % LE : 66 %

Tableau 3 : Résultat d'une enquête sur la distribution des déterminants en reprise immédiate

[10] SN = syntagme nominal

N°	*Exemple*	*Contraste*	*Cond. d'éval°*	*LE (%)*	*CE (%)*
1	Il était une fois un prince. ? prince ne pouvait avoir de fils.	CE	CE	10	90
2	J'ai acheté une voiture hier. ? voiture est bleue.	CE	CE	35	65
3	Un chien aboie. ? chien a faim.	CE	CE	29	71
4	J'ai vu une voiture. ? voiture roulait vite.	CE	CE	13	87
				22	78
5	Un homme entra. ? homme dit bonjour.	CE	LE	53	47
6	Un avion s'est écrasé hier. ? avion venait de Miami.	CE	LE	47	53
7	J'ai heurté un camion. ? camion venait de droite.	CE	LE	47	53
8	Il était une fois un prince très malheureux. ? prince ne pouvait avoir de fils.	CE	LE	22	78
				42	58
9	Le tableau représente une jeune fille. ? jeune fille sourit.	LE	LE	53	48
10	Paul a tué un renard. ? renard voulait s'introduire dans son poulailler.	LE	LE	53	48
11	J'ai vu une voiture et un camion. ? voiture roulait vite.	LE	LE	91	9
				66	34
12	Une femme entra dans la pièce. J'avais déjà vu ? femme chez un ami.	LE	CE	3	97
13	Paul a tué un renard. Il poursuivait ? renard depuis trois mois.	LE	CE	18	82
14	Dans mon jardin, il y a un cerisier. ? cerisier a été planté par mon père.	LE	CE	21	79
15	Le tableau représente une jeune fille. ? jeune fille s'habille habituellement en vert.	LE	CE	25	75
16	Il était une fois un prince qui vivait dans un beau château. ? prince ne pouvait avoir de fils.	LE	CE	35	65
17	Un prince aimait une princesse. ? prince aimait aussi les fleurs.	LE	CE	37	63
				23	77
18	J'ai vu un camion et une voiture. ? camion et pas un autre roulait vite.	CE	?	25	75
Total				35	65

Tableau 4 : Résultats de l'enquête sur la préférence entre définis et démonstratifs des exemples 1 - 18

- En ce qui concerne les extrêmes, il faut retenir le deuxième volet du paradoxe (l'exemple 11 du Tableau 4), avec une très grande majorité pour une reprise au défini. Mais de façon plus surprenante, il ne s'oppose pas au premier volet (l'exemple 4 du Tableau 4), mais à l'exemple 12. Ici, on ne peut qu'admettre, avec G. Kleiber (1986 : 71) que « *l'article défini se trouve totalement exclu des p2 qui présentent explicitement le référent comme saisi en dehors de la circonstance d'évaluation de p1* ». Mais cette affirmation n'est pas tout à fait satisfaisante, parce qu'elle ne dit

pas pourquoi cet effet n'est pas aussi net dans tous les exemples présentés comme tels. Force est de constater qu'il y a probablement des nuances qui échappent à l'explication en termes de prolongement de circonstance. Cette impression est particulièrement forte pour l'exemple 17 du Tableau 4 : ici, il y a visiblement d'autres facteurs qui entrent en jeu. L'introduction, en p1, d'un domaine sémantiquement homogène, regroupant des référents jouant des rôles thématiques principaux (agent et patient) et l'instanciation d'une relation de parallélisme discursif semblent augmenter sensiblement les chances de saisie par un défini.

On pourrait tenter de conclure que notre enquête réconcilie les deux principes qui, au final, semblent complémentaires : les résultats ne s'expliquent pas exclusivement par l'une ou l'autre des propositions. On constate plutôt qu'en cas de rupture des circonstances d'évaluation, le démonstratif l'emporte à plus de 75%. Mais dans les autres cas, la distribution optimale semble s'accorder avec la structure contrastive fournie par le contexte. En tout état de cause, les résultats de cette étude sont à manipuler avec beaucoup de prudence : ils concernent seulement 18 exemples, testés dans des conditions qui ne permettaient pas un contrôle de tous les paramètres susceptibles d'influer sur la décision des sujets. L'intérêt de cette enquête a surtout été de confirmer les prédictions en cas d'accord des deux principes et de nuancer les jugements en cas de désaccord.

1.2.3 Apports et limites des approches linguistiques pour une modélisation de la référence

Notre objectif n'est pas tant d'entrer dans les détails de la polémique autour du « paradoxe de la reprise immédiate »[11] que de tenter d'évaluer les apports et limites des études linguistiques pour la conception d'un modèle opérationnel de la référence. Dans cette perspective, nous verrons que ce débat contradictoire peut être considéré comme exemplaire à plusieurs égards.

- *Apports*

Un premier apport fondamental des approches linguistiques consiste en l'étude systématique des liens entre l'apparition de telle ou telle expression référentielle et les structures du contexte. Ce type d'étude permet de dégager deux points essentiels :

D'abord, un classement des expressions référentielles en « introducteurs de nouvelles entités » (indéfini) et « expressions anaphoriques » (défini, démonstratif, pronom) ne s'avère pas assez fin pour prédire la distribution effective des marqueurs référentiels. Outre le fait que tout le « paradoxe de la reprise immédiate » n'aurait même plus lieu d'être, l'étendue du débat autour de ce seul phénomène montre bien qu'une vision dichotomique des expressions référentielles laisse de côté bon nombre de facteurs entrant en jeu pour le choix des expressions disponibles.

Ensuite, il apparaît clairement que le rôle à accorder aux structures contextuelles dépasse de loin les seuls facteurs de récence ou de saillance : F. Corblin explique en effet les différences d'usage par des recouvrements plus ou moins grands entre structures contextuelles données et valeurs contrastives inhérentes aux déterminants. G. Kleiber y ajoute l'idée que l'enchaînement discursif des éventualités a également un rôle à jouer dans la distribution des déterminants en reprise immédiate. Au final, les deux propositions semblent effectivement entrer en jeu, d'autant plus qu'elles font les mêmes prédictions (bien que pour des raisons différentes) dans les cas les plus marqués, comme la difficulté d'extraction, par un démonstratif, d'un référent d'un groupe coordonné. Or, la mise à jour de facteurs aussi complexes que les structures contrastives des ensembles contextuels ou l'enchaînement des

[11] Pour se faire une idée plus précise des arguments entrant en jeu, on pourra se reporter aux ouvrages de G. Kleiber (1989) et de F. Corblin (1995).

éventualités passe nécessairement par des études fines et focalisées sur des phénomènes linguistiques précis et les linguistes sont évidemment les mieux placés pour de telles entreprises. Il appartient alors aux chercheurs des autres domaines soucieux de modéliser la langue naturelle – informaticiens, chercheurs en intelligence artificielle, psychologues – de s'y intéresser de plus près afin d'en tirer un bénéfice pour leur modélisation.

Mais au-delà de l'étude de phénomènes précis, le souci des linguistes est de proposer des explications qui s'insèrent dans des visions plus générales sur les systèmes linguistiques étudiés et c'est là un deuxième apport fondamental de ces approches. Là aussi, le débat du « paradoxe » est paradigmatique : F. Corblin (1987) propose en effet une explication qui repose non seulement sur une théorie globale des constructions linguistiques de la référence en français, mais qui permet encore d'expliquer les deux volets du paradoxe. Cela n'est pas exactement le cas de G. Kleiber qui, lui, traite le deuxième volet indépendamment du premier. En revanche, dans ses travaux ultérieurs (1990, 1994), il élargit le problème au « *troisième larron de la reprise immédiate qui est le pronom personnel* » (1986 : 78), ce qui va effectivement vers une couverture plus large des formes de la reprise immédiate. Or, nous pensons que seules de telles approches unifiées du système des marqueurs référentiels permettent d'éviter l'écueil d'une modélisation trop simplifiée des processus d'interprétation référentielle.

- *Limites*

Si les études linguistiques du système des expressions référentielles constituent pour nous une étape incontournable de la modélisation, elles ne sont pas pour autant sans soulever quelques problèmes de taille.

Une première limite de ces approches consiste en l'absence d'une opérationnalité immédiate des critères proposés. Là aussi, le « paradoxe de la reprise immédiate » peut servir d'exemple : pour mémoire, G. Kleiber explique en effet la distribution optimale du défini et du démonstratif par un « algorithme » dont les deux critères-clé sont la « circonstance d'évaluation » et le « prolongement du cadre d'évaluation ». Or, ceux-ci s'avèrent très difficilement opérationnels, comme l'a montré l'essai de leur application au premier volet du paradoxe (5). En l'absence d'une définition plus claire et/ou d'indices linguistiques, on ne voit effectivement pas en fonction de quoi il serait possible de décider si un événement est « standard », relate un « compte rendu de perception » ou peut être jugé « exceptionnel ». La définition de ce que peut être le prolongement d'un cadre évaluatif n'est pas non plus donnée. Toutefois, en examinant les exemples, il serait possible de retenir des critères comme certaines relations de discours (RST, Mann et Thompson, 1988 ; S-DRT, Asher, 1993), la continuité dans un script au sens de R. Schank (1975) ou encore l'adjacence temporelle des éventualités relatées (Romary, 1989). Des facteurs de rupture seraient surtout des cas de rupture temporelle – la « sortie » d'un cadre événementiel – linguistiquement marqués par un changement de temps verbal ou certains adverbes (*habituellement*, *aussi*). Mais ce sont là tout au plus des hypothèses qui demandent à être validées par des expériences plus poussées et plus systématiques que l'enquête que nous avons pu réaliser.

Ce problème de la fixation des notions-clé proposées par G. Kleiber en rejoint un autre, relevé par F. Corblin (1995 : 78) : G. Kleiber (1989 : 26) justifie l'apparition du défini dans certains exemples en avançant que « l'article défini présente dans ce cas le contexte d'énonciation comme une circonstance d'évaluation ». Or, si c'est le déterminant qui détermine la nature de p1 (circonstance d'évaluation pertinente ou non), on est en droit de s'interroger sur la valeur prédictive de l'hypothèse, dont le but est justement de prédire l'apparition de tel ou tel déterminant en fonction de la nature de p1. Il est

évident que ce problème de circularité persistera tant qu'on ne dispose pas d'une spécification plus avant des ingrédients de l'algorithme.

Un deuxième problème d'opérationnalité est illustré par l'utilisation du principe des contrastes par G. Kleiber : nous avons déjà mentionné à plusieurs reprise que l'application « radicale » de la notion de contraste mène à des prédictions erronées, d'autant plus que F. Corblin (1983, 1995) souligne explicitement qu'il ne s'agit pas de prédire une distribution complémentaire, mais des préférences d'usage, en particulier dans des contextes à forte valeur contrastive. Cependant, cette utilisation « radicale » des contrastes soulève un réel problème qui prolonge le problème d'opérationnalité de certaines notions linguistiques. Si l'on veut éviter une implémentation caricaturale des idées de F. Corblin, il faudra très probablement gérer de façon beaucoup plus fine différentes échelles contrastives : bien que tous les exemples (7) à (10) introduisent deux entités contextuelles, la mise en contraste de celles-ci n'est certainement pas la même :

(7) J'ai vu *une voiture* et *un camion.*

(8) *Paul* a tué *un renard.*

(9) *Une femme* entra dans *la pièce.*

(10) Dans *mon jardin*, il y a *un cerisier.*

Il apparaît clairement que le mode de coordination ainsi que la fonction syntaxique (ou le rôle thématique) des constituants ont un rôle à jouer. Un syntagme nominal coordonné tel qu'en (7) introduit plusieurs entités jouant un même rôle thématique. Il semble alors qu'elles forment un groupe plus homogène dans la mesure où aucune entité n'est mise en valeur par rapport à une autre. Regrouper des entités participant à un même événement, mais jouant des rôles différents, se révèle déjà plus délicat : l'on peut faire l'hypothèse – et la théorie du Centrage (Grosz et al., 1995) la fait – que des entités en position sujet sont plus saillantes que celles en position objet direct, elles-mêmes mises en avant par rapport aux objets indirects et aux circonstanciels, mais il ne s'agit là que d'une hypothèse parmi d'autres (cf. Schnedecker et Bianco, 1995, pour une synthèse).

Enfin, un dernier point problématique est directement lié aux remarques précédentes : tant que les concepts mis en jeu par les approches linguistiques ne sont pas minimalement formalisables, les hypothèses avancées reposent le plus souvent sur les seuls exemples et extraits traités et il est très difficile de les évaluer sur d'autres données. Or, un module de la référence qui puisse être intégré dans des systèmes de dialogue sera confronté à des données dont la nature est très différente des exemples construits par les linguistes et même des exemples réels tirés de la littérature française. La question qui se pose alors est celle de l'adaptation des hypothèses proposées à des corpus de dialogues oraux et dépassant le cadre strictement linguistique, entre autres par le rôle que joue le contexte de la situation de communication. Cela nous amène à la deuxième variation de notre thème qui est le lien entre référence et dialogue finalisé.

1.3 Référence et dialogue finalisé ou « Ça va être dur avec ces formes-là... »[12]

1.3.1 La référence aux objets dans le dialogue homme-machine

Si les études linguistiques appréhendent la référence essentiellement à travers l'étude des constructions langagières que sont les expressions référentielles, résoudre la référence dans le cadre d'un système de communication homme-machine revient à identifier ou ré-identifier des entités

[12] Extrait d'un corpus de dialogues finalisés, portant sur la manipulation de formes géométriques (Ozkan 1994). Ce corpus est présenté plus en détail ci-dessous.

désignées par ces expressions référentielles. Aux nombreux facteurs linguistiques qui contribuent à ce processus d'identification s'ajoutent alors des facteurs spécifiques aux circonstances d'élocution. Parmi ces facteurs, on peut citer en premier lieu l'intentionnalité d'un dialogue homme-machine : étant donné que nous excluons pour l'instant la possibilité de conversations libres sur n'importe quel sujet, les dialogues qui nous intéressent ici seront nécessairement fondés sur un but. Or, qu'il s'agisse de buts dirigés par les états, les objets, la tâche ou les objectifs (cf. la classification proposée par M. L. Bourguet, 1992), il est évident que l'interprétation des expressions référentielles est soumise à ces buts.

En deuxième lieu, nous aimerions attirer l'attention sur le rôle prépondérant que joue l'environnement visuel dans un dialogue de ce type : la conception linguistique classique de la référence a surtout été centrée sur l'opposition entre deux phénomènes – anaphore et déixis – l'un nécessitant, pour être interprété, le recours au cotexte linguistique, l'autre à la situation d'énonciation. A l'instar de certains travaux linguistiques plus récents (Kleiber, 1990, 1994 ; Reboul et Moeschler, 1994), le dialogue homme-machine remet en question l'adéquation de cette distinction (Frey, 1989). La résolution de la référence va en effet au-delà de la mise en correspondance de segments linguistiques, dans la mesure où l'objectif ultime est d'établir un lien entre des segments linguistiques – les expressions référentielles – et des objets appartenant à l'application informatique à commander. Or, ce processus demande la prise en compte constante de la situation d'énonciation et en particulier du retour visuel qu'un utilisateur peut avoir sur l'état de l'application.

Une autre particularité de l'étude de la référence dans la perspective du dialogue homme-machine est liée à la conciliation de contraintes spécifiques à l'application d'une part et à l'utilisateur d'autre part. Les contraintes spécifiques à l'application suivent une logique informatique : la commande d'une application passe par l'exécution de procédures ou de fonctions, respectant une syntaxe rigide, portant sur des objets typés et requérant des paramètres précis. Dans un dialogue homme-machine, ces éléments doivent être extraits des données linguistiques, produites par un humain à partir de représentations cognitives que celui-ci se fait de l'application. Or, puisqu'un dialogue homme-machine n'est une forme appropriée d'interaction que lorsque l'utilisateur n'est pas parfaitement au courant du langage de commande sous-jacent, il n'y a plus de raisons de supposer un parallélisme entre les représentations internes à la machine et celles de l'utilisateur. On peut en effet raisonnablement supposer que les représentations de l'utilisateur sont façonnées par au moins trois facteurs.

- D'abord, le retour visuel ne reproduit évidemment pas la totalité des informations relatives à la tâche. Il s'agit d'une représentation abstraite d'un sous-ensemble d'objets repartis sur un espace visuel. On peut dès lors faire l'hypothèse que cette représentation focalise l'attention de l'utilisateur à ce sous-ensemble d'objets. Par ailleurs, la structuration spatiale, par exemple une proximité visuelle entre objets, peut faire croire à une proximité fonctionnelle, même si l'application fonctionne selon d'autres principes.

- Ensuite, les connaissances que l'utilisateur a de la tâche vont influer sur la catégorisation des objets. Cette tendance est particulièrement bien visible lors d'une tâche de construction : nous présenterons ci-dessous un corpus, collecté lors de la conception de dessins à partir de figures géométriques, qui illustre par de nombreux exemples la variation dans la désignation des objets, selon qu'il s'agit de l'espace source, fourni par les figures géométriques (*triangle, cercle,...*), ou de l'espace cible, fourni par le dessin final (*toit, soleil, ...*).

- Enfin, si l'utilisateur d'une application pilotée par le langage naturel est supposé avoir une certaine connaissance de la tâche à effectuer, il n'est pas pour autant supposé connaître parfaitement bien

l'outil informatique. Cela peut entraîner en particulier un décalage dans la représentation de la complexité d'une tâche : ce qui semble une seule action à l'utilisateur (*Mets un triangle à gauche !*) doit en effet être décomposé en actions plus élémentaires (prise d'un triangle, placement à un endroit sur l'écran) exécutables par la machine.

Pour résumer, l'enjeu du calcul référentiel dans le dialogue homme-machine consiste à maîtriser constamment le décalage entre des représentations identifiantes pour les objets de l'application informatique à commander et des représentations cognitives que l'utilisateur s'en fait. Parmi les deux solutions possibles – demander à l'utilisateur de s'adapter à la machine ou donner à la machine la capacité de s'adapter à l'utilisateur – nous choisirons la seconde.

Or, donner à la machine la capacité de s'adapter à l'utilisateur demande que celle-ci soit pourvue d'un fonctionnement compatible avec celui de l'utilisateur : « *Pour garantir l'ergonomie des interprétations construites par la machine, c'est-à-dire leur conformité aux attentes des utilisateurs, le fonctionnement du système mis en œuvre doit être analogue à celui de la cognition humaine.* » (Sabah, 1997 : 24). Il s'ensuit qu'il est d'abord nécessaire d'étudier les capacités cognitives des utilisateurs : pour la partie du fonctionnement cognitif qui nous intéresse ici – l'interprétation référentielle – cela signifie qu'il faut s'intéresser prioritairement aux processus de référenciation dans des situations naturelles. C'est pour cette raison que nous avons choisi d'accorder une place importante à l'observation de données réelles, issues d'un corpus de dialogues homme-homme, collecté par N. Ozkan à l'IMAG (Ozkan, 1994).

1.3.2 Un corpus de dialogues finalisés naturels : le corpus Ozkan

Le choix de ce corpus a été déterminé par plusieurs critères. L'absence de réels corpus homme-machine ne nous permettait pas de travailler sur des données issues de ce type d'interaction. Une autre solution aurait pu être d'étudier un corpus enregistré en « Magicien d'Oz », c'est-à-dire confrontant des sujets à un compère simulant les capacités de la machine. Mais ce dernier point nous a semblé problématique à plusieurs égards : d'abord, attribuer des capacités à la machine sous-entend de faire d'emblée l'hypothèse d'une machine à capacités limitées (« *limiter la production des anaphores en répétant, là où il est naturel de le faire, les unités référençables* » ou « *refuser la compréhension des ellipses* » ; Morel, 1989) et donc de demander à l'utilisateur de s'adapter. Pousser cette logique à l'extrême conduit en effet à « *construire une machine à faire formuler et reformuler des consignes par paraphrases jusqu'à obtenir la précision que requiert son exécution* » (Vivier et Nicolle, 1997 : 249). Or, c'est exactement le contraire qui motive notre travail. Ensuite, l'expérience acquise au cours de telles simulations (Morel, 1989 ; Vivier et Nicolle, 1997) montre que les compères ont souvent du mal à respecter des contraintes supposées être celles d'une machine. Par ailleurs, il a été remarqué à plusieurs reprises (Morel, 1989 ; Luzzati, 1995 ; Vivier et al., 1998) que la communication avec une machine, fut-elle simulée, induit des comportements spécifiques de la part des utilisateurs. Si ces spécificités méritent sans aucun doute de l'attention, il ne nous semble pas pour autant judicieux de fonder un modèle de la référence sur ces phénomènes particuliers. Au contraire, ce n'est que par rapport à des échanges linguistiques naturels et non contraints que l'on pourra évaluer la spécificité de ce type de comportement. Enfin, parmi les corpus homme-homme disponibles, nous avons écarté ceux qui ont été enregistrés via une liaison téléphonique (*CIO*, *SNCF* ; Morel, 1989), pour ne pas compromettre d'avance la possibilité d'étudier la relation entre un environnement visuel commun et les processus référentiels.

Nous avons alors retenu le corpus collecté par N. Ozkan (1994) lors de sa thèse consacrée à un modèle dynamique du dialogue. Ce corpus, que nous appellerons désormais « corpus Ozkan », a l'avantage de combiner le « finalisé » de la tâche avec le « naturel » du langage. Il s'agit d'un corpus

enregistré lors de la collaboration de deux sujets autour d'une tâche de conception de dessins simples à partir de figures géométriques. Le scénario à la base des dialogues est le suivant : un manipulateur (M) suit les instructions d'un instructeur (I) qui est le seul à connaître le dessin final à réaliser. I et M sont placés dans des salles séparées. Ils partagent un même écran, comportant la scène en construction et la palette des figures géométriques disponibles. Par ailleurs, ils peuvent intervenir tous les deux dans la construction du dessin, mais en aucun cas, I exclut M de la participation à la tâche.

D'un point de vue matériel, la partie du corpus sur laquelle nous avons pu travailler[13], comprend 33 dialogues, répartis entre sept couples de sujets et portant sur la reconstruction de sept scènes différentes. Il s'agit de la partie du corpus pour laquelle une transcription orthographique avait déjà été réalisée à l'IMAG. Cette transcription, assortie de différentes annotations supplémentaires (actions et gestes des sujets, nature des actes dialogiques, stratégies des interlocuteurs, implicites) nous a permis tout au long de ce travail d'appuyer nos réflexions sur des données concrètes et réelles. Pour la suite de notre travail, le corpus Ozkan nous servira donc de matière de référence, qu'il s'agisse d'illustrer des phénomènes linguistiques, de tester certaines modélisations existantes ou de traiter les données selon notre propre modélisation.

L'exemple (11) correspond à la transcription[14] du dialogue *C5Égypte*[15] en entier. La Figure 2 montre l'aspect de la palette des objets disponibles et le dessin final, dont seul I a connaissance au début du dialogue.

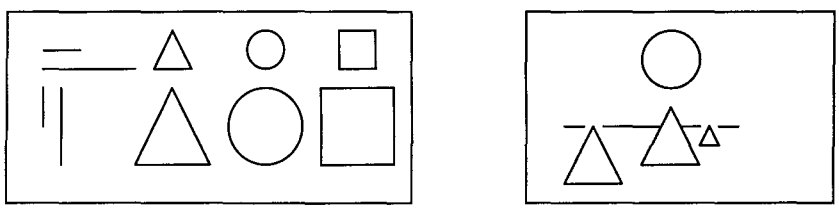
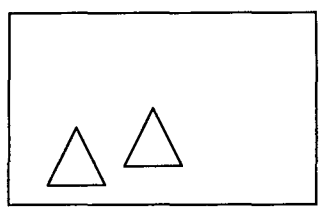


Figure 2: Palette des objets disponibles (corpus Ozkan)

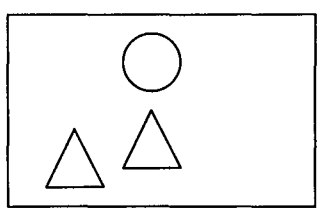
(11)

I_1 bon, alors il faut faire des pyramides
donc je vais essayer de la faire quand même
[I pose une pyramide]
M_1 ouais
I_2 voilà la première
[I pose une deuxième pyramide]
I_3 trop haute...tu peux pas essayer de la redescendre un petit peu
M_2 oui
I_5 un peu plus bas voilà comme ça
I_6 non un peu plus haut
I_7 voilà comme ça



Scène 1 (I_7)

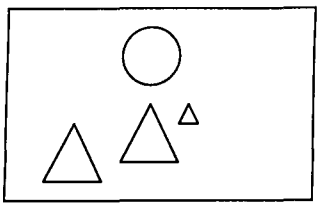
I_8 et puis ensuite est-ce que tu peux me prendre le grand rond
M_3 oui bien sûr
I_9 et le placer au dessus de la pyramide
M_4 laquelle ?
I_{10} celle de droite
M_5 collé ?
I_{11} non non, au-dessus, de sorte que ça fasse un soleil en fait
[M pose un rond]
M_6 d'accord... comme ceci ?
I_{12} voilà comme ça



Scène 2 (I_{12})

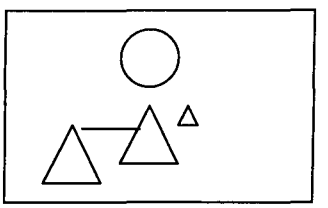
[13] Nous tenons à remercier Jean Caelen et Jean-François Sérignat pour nous avoir permis l'accès à ces données.
[14] En ce qui concerne les gestes de manipulation, nous en reproduisons que ceux qui sont nécessaires à la compréhension du dialogue.
[15] Le code des exemples est composé du n° du couple (C5) et du titre du dessin à réaliser (Égypte).

ah j'ai oublié une petite pyramide
est-ce que tu pourrais me prendre le petit triangle
M_7 oui
I_{13} et le placer tout à droite de la grande pyramide
en fait du grand triangle
M_8 de celui de droite ?
I_{14} de celui de droite oui
[M tente un emplcament]
M_9 à ce niveau-là ?
I_{15} un peu plus rapproché
[M pose le triangle]
I_{16} voilà comme ça



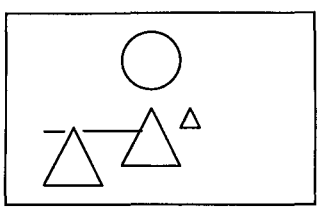
Scène 3 (I_{16})

et puis maintenant faudrait faire la ligne d'horizon
il faut prendre une grande euh une grande horizontale
I_{17} et la placer euh à la pointe des triangles
des deux grand triangles
M_{10} ah mais elle va pas être horizontale alors
parce qu'il y en a un qui est plus haut que l'autre
I_{18} ouais pff ... on s'est mal débrouillées
[M manipule la barre]
un peu plus vers le bas
I_{19} un peu plus vers le bas...encore, encore
[M manipule la barre]



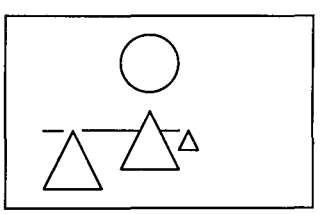
Scène 4 (I_{19})

I_{21} voilà comme ça et tu en prends une deuxième, une petite
M_{11} une petite
I_{23} tu la places à gauche de la euh pyramide de gauche
M_{12} oui [M manipule la barre]
I_{24} voilà
M_{13} comme ceci ?
I_{25} voilà comme ça,



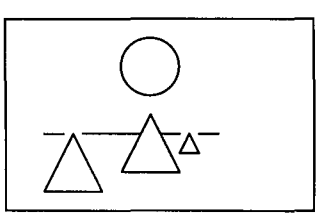
Scène 5 (I_{25})

et t'en prends... ah oui ce sera pas exactement le même dessin
et t'en prends une autre petite
M_{14} oui
I_{26} et tu la places à droite de la pyramide de droite
I_{27} euh comme pour l'autre avant
M_{15} qui la touche ?
[M manipule la petite barre]
I_{28} oui, ce qu'il faudrait c'est que tu descendes un peu le petit triangle
M_{16} oui... jusqu'où ?
I_{29} bah jusqu'à la ligne d'horizon que ça passe vers la pointe
[M manipule le triangle]
M_{17} que ça passe par la pointe ?
I_{30} un peu plus la pyramide en haut
M_{18} un peu plus en haut
[M manipule le triangle]
I_{31} voilà comme ça..



Scène 6 (I_{31})

et tu prends une autre petite verticale
M_{19} une autre petite verticale
I_{32} euh horizontale pardon
M_{20} ah
I_{33} et puis tu la places dans la même lignée
à droite de la petite pyramide
M_{21} oui de sorte à continuer l'horizontale-là
I_{34} voilà, bah c'est bon tu peux vider le dessin



Scène 7 (I_{34})

1.3.3 Variété des phénomènes référentiels du dialogue naturel finalisé

Comme le montre cet exemple, ce corpus est extrêmement riche du point de vue référentiel et ceci malgré l'univers restreint de la tâche. Il illustre en effet toute la variété des phénomènes référentiels dans un dialogue finalisé et multimodal, combinant langage, perception visuelle et gestes.

D'abord, il contient tout type d'expressions référentielles (indéfini : *des pyramides* ; défini : *la pyramide* ; pronom personnel : *la* ; démonstratif : *celle de droite*) qui entretiennent par ailleurs des liens bien plus variés que simplement coréférentiels (*des pyramides – la première, celle de droite* ; *une grande horizontale – une deuxième, une autre petite* ; *le petit triangle – la pointe*).

Ensuite, en partant d'une tâche simple et d'un univers référentiel initialement constitué par les dix figures géométriques de la palette, les locuteurs font évoluer cet univers de façons multiples et dynamiques : en particulier, ils passent de la palette (*le grand rond*) à la scène cible (*un soleil*), sans toujours marquer (et remarquer...) ce phénomène explicitement comme en I_{12}/ I_{13} (« *est-ce que tu pourrais me prendre le petit triangle et le placer tout à droite de la grande pyramide, en fait du grand triangle* »). Nous avons ici une illustration parfaite de la catégorisation des objets en fonction de la représentation que les locuteurs se font de l'avancement de la tâche.

Ce dialogue confirme également l'hypothèse selon laquelle la structure visuelle des entités de l'application influe sur les possibilités référentielles : en I_{26}, *la pyramide de droite* ne réfère pas à la pyramide la plus à droite de la scène (la petite pyramide), mais à celle qui est la plus à droite parmi le sous-ensemble perceptivement homogène que forment les deux grandes pyramides. Par ailleurs, nous avons souligné le rôle important que peut jouer la connaissance de la tâche dans l'interprétation. Ce rôle est mis en évidence par les interventions I_9 (« *le placer au dessus de la pyramide* ») et M_5 (« *collé ?* ») : tant que I n'annonce l'objectif de la pose du rond – *de sorte que ça fasse un soleil en fait* – M est en effet dans le doute sur la position verticale exacte de celui-ci. Comme l'ont relevé par ailleurs J.-M. Pierrel et L. Romary (1997), un utilisateur peut parfaitement être au courant de la tâche tout en faisant abstraction de la structure élémentaire des procédures et paramètres à transmettre à l'application. C'est ce que l'on peut observer en I_{21} et I_{23} (« *tu en prends une deuxième, tu la places à gauche de la pyramide de gauche* »), où la formulation linguistique de la requête de placement ne comporte ni le type de l'objet (barre), ni son orientation (horizontale), ni la hauteur du placement (même hauteur que le premier élément de la ligne d'horizon).

Par ailleurs, la question de la pertinence de la distinction entre anaphore et déixis est soulevée à travers le fait que les locuteurs ancrent leurs expressions référentielles aussi bien sur le contexte perceptif que discursif : un défini considéré comme « anaphorique » tel que *la première* (I_2) nécessite en effet la prise en compte du contexte d'énonciation pour s'interpréter, tout comme le pronom personnel *la* « sans antécédent » (Kleiber, 1990) en I_1.

Enfin, ce dialogue permet de constater un certain nombre de phénomènes typiques de l'oral, susceptibles d'interférer avec les processus de résolution référentielle. Mentionnons en premier lieu le nombre très élevé d'ellipses sur la tête nominale des groupes nominaux (*la première, celle de gauche, une deuxième*). Les reprises et corrections (*la pyramide – laquelle* ; *une autre petite verticale – euh horizontale, pardon*) méritent également qu'on leur porte un intérêt particulier : non seulement elles doivent être traitées correctement, mais un modèle qui tient compte des représentations respectives des deux interlocuteurs devrait même être capable de prédire leur apparition. Le dialogue donné en exemple montre aussi que les structures syntaxiques sont loin de la norme du français écrit : cela signifie que des modèles faisant intervenir la syntaxe dans les processus de résolution (co)référentielle ne sont que partiellement transposables aux dialogues naturels. Enfin, la composante gestuelle ne peut

pas être négligée : des expressions telles que *à ce niveau-là* ou *comme ceci* (même si leur interprétation ne fait pas partie des objectifs de cette thèse) demandent obligatoirement la prise en compte des gestes de désignation et plus largement des manipulations des protagonistes.

1.4 Référence et modélisation opérationnelle : Quo vadis ?

Le constat de la richesse des phénomènes référentiels d'un dialogue naturel, fût-il finalisé, nous renforce en premier lieu dans notre conviction quant à l'importance à accorder aux approches linguistiques, préconisant une prise en compte globale du système de la détermination : au vu de la fréquence et de la diversité des expressions référentielles, un partenaire dialogique qui refuserait de comprendre ou produire des pronoms, ellipses et anaphores (comme l'envisage par exemple M. A. Morel, 1989), ne pourra plus, selon nous, prétendre au qualificatif « naturel ». Une première contrainte que nous nous imposons dès lors pour une modélisation de la référence est un modèle unifié pour tout type d'expression référentielle.

Mais la présentation rapide du type de données auxquelles nous nous intéressons ici a également montré qu'il faudra aller au-delà des phénomènes purement linguistiques : notre objectif final étant l'identification des objets dont parle un locuteur, nous devons obligatoirement tenir compte d'informations de natures diverses : linguistique, évidemment, mais aussi perceptive (visuelle en particulier), gestuelle (manipulation ou désignation d'objets), conceptuelle (par exemple, sur la composition des objets), ou encore spécifique à la tâche. De cette nécessité découle une deuxième contrainte que nous nous imposerons pour la suite de ce travail : nous envisagerons une modélisation permettant d'intégrer dans un cadre représentationnel unifié des informations hétérogènes, c'est-à-dire linguistiques, perceptives et conceptuelles.

En plus des deux premiers objectifs, notre travail est motivé par le souhait de proposer un modèle du calcul référentiel qui soit plausible d'un point de vue cognitif. Cette plausibilité pourrait être évaluée d'abord à travers la comparaison des prédictions faites par notre modèle avec les phénomènes effectivement observables dans le corpus Ozkan. Mais étant donné qu'un corpus de dialogues naturels ne comporte pas nécessairement tous les phénomènes prédictibles, il serait également envisageable de faire confirmer certaines prédictions par des expériences psycholinguistiques ciblées. Enfin, comme nous faisons l'hypothèse que les processus cognitifs impliqués dans l'analyse et la génération d'expressions référentielles ne sont pas fondamentalement différents l'un de l'autre, la réversibilité de notre modèle pour la génération automatique d'expressions référentielles serait non seulement un bon indicateur pour la plausibilité cognitive de notre proposition, mais encore intéressante d'un point de vue informatique : comme les rôles d'un système de dialogue changent entre celui de locuteur et celui d'auditeur, un modèle réversible permettrait en effet d'implémenter un module unique pour l'analyse et la génération d'expressions référentielles.

Enfin, un dernier impératif de notre modélisation est le critère d'opérationnalité. Si nous nous intéressons d'abord à la modélisation des processus référentiels, nous ne perdons pas de vue que nous le faisons dans l'objectif d'améliorer la communication homme-machine par l'introduction du langage naturel. Nous envisageons donc non seulement un modèle unifié, prédictif et cognitivement adéquat, mais aussi un modèle qui puisse effectivement être implémenté dans des systèmes de dialogue. Cela demande alors de spécifier de façon précise l'ensemble des structures de données utilisées ainsi que les opérations définies sur ces structures. L'intérêt de la contrainte d'opérationnalité n'est pas seulement de garantir la possibilité d'une implémentation, mais aussi de poser les bases pour une évaluation objective du modèle sur des données réelles. Enfin, une telle approche a l'avantage de mettre clairement en évidence les limites actuelles de l'automatisation du traitement de la langue naturelle.

Afin de répondre à ces objectifs, nous développerons une modélisation des processus d'interprétation référentielle fondée sur la notion de **domaine de référence**. L'idée fondamentale sur laquelle repose notre travail est issue d'un certain nombre de travaux informatiques (Gaiffe, 1992; Romary, 1993; Dale et Reiter, 1995), linguistiques (Corblin, 1987) et cognitifs (Olson, 1971 ; Langacker 1991, Reboul et al., 1997) qui ont mis en évidence l'importance de l'ancrage de tout acte de référence dans un contexte d'interprétation assis sur la notion de contraste ou de différenciation. Il s'en dégage l'hypothèse que l'interprétation d'une expression référentielle demande l'isolement de l'objet désigné dans un contexte local – ou domaine de référence – sur la base de critères distinctifs le long d'une axiologie.

La conséquence de ce point de vue est que l'interprétation d'une expression est plus que l'identification d'un référent : c'est aussi l'identification de son domaine de référence et donc d'un ensemble de référents alternatifs exclus. A partir de là, nous pensons que l'identification systématique des « alternatives exclues » peut être mise à profit de processus des compréhension à la fois plus efficaces d'un point de vue informatique et plus proches du fonctionnement cognitif. Ces alternatives sont en effet souvent considérées comme fournissant le domaine d'interprétation pour les expressions elliptiques, ambiguës ou ayant trait à l'altérité (Vivier et al., 1997, Kievit et Piwek, 2000). Cette observation traduit, selon nous, un principe d'interprétation plus fondamental : les domaines de référence forment l'espace d'ancrage contextuel préférentiel pour toutes les expressions à interpréter. Nous proposons, par conséquent, une modélisation qui structure le contexte en domaines de référence. Nous faisons ensuite l'hypothèse que différents types d'expressions référentielles imposent des contraintes spécifiques sur l'extraction de leur référent de ce contexte et qu'elles le restructurent différemment. A partir de là, nous sommes capables de formuler des prédictions portant à la fois sur l'acceptabilité de différents types d'expressions dans un contexte donné et sur des interprétations préférentielles en cas d'ambiguïté.

Les grandes lignes autour desquelles s'organise notre thèse sont les suivantes : dans une première partie (chapitres 2 à 5), nous dresserons un panorama des travaux en linguistique, psycholinguistique et linguistique computationnelle ayant trait au problème de l'interprétation référentielle. Le deuxième chapitre, consacré aux travaux linguistiques, présente notamment un certain nombre de problèmes récurrents – liés à l'usage des indéfinis, définis, démonstratifs et pronoms personnels – qui justifient que ces différentes formes ne soient pas traitées globalement en fonction de leur seule capacité à reprendre ou non un antécédent, mais bien plus sur les contraintes que chacune d'entre elles expriment sur son contexte d'interprétation. Le troisième chapitre étudie plus en détail une de ces contraintes – l'accessibilité ou la focalisation – ayant donné lieu à un nombre important de travaux et expérimentations psycholinguistiques. A partir du constat que le facteur d'accessibilité à lui seul est insuffisant pour rendre compte des contraintes contextuelles sur l'interprétation référentielle, nous nous intéresserons, dans un quatrième chapitre, à d'autres modélisations contextuelles, proposées dans des cadres théoriques plus formels ou dédiées directement à la compréhension automatique de la langue. Il s'en dégage en effet d'autres contraintes, mais l'étude des implémentations effectives, au cinquième chapitre, montre que les systèmes actuels ne couvrent en général qu'une faible portion des phénomènes concernés. A travers cette première partie, il sera malgré tout possible de dégager des traits constants et utiles pour une démarche plus globale, même si une remise en perspective apparaît nécessaire.

La deuxième partie de cette thèse (chapitres 6 à 9) est consacrée à la présentation de la modélisation que nous proposons. Nous dessinerons, au sixième chapitre, les grandes lignes de la notion de domaine de référence. Celle-ci permet d'appréhender le traitement des expressions référentielles non comme une simple assignation de référents, mais plutôt comme un mécanisme dynamique de structuration et de focalisation de sous-ensembles contextuels. Ce chapitre montre en

particulier comment la notion de domaine de référence peut être retrouvée dans un certain nombre de travaux précédents sans qu'elle n'ait été en général posée comme principe fondateur. Nous développerons alors une série d'arguments – dépassant le simple champ du traitement automatique des langues, pour impliquer également des travaux issus des domaines cognitifs, psycholinguistiques et discursifs – en faveur de cette notion. Les trois chapitres suivants présentent en détail les caractéristiques du modèle proposé, en abordant successivement la description des structures de base du modèle contextuel, leur utilisation pour le calcul référentiel par identification et restructuration de domaines de référence et enfin, l'illustration de la prédictivité du modèle par sa mise en œuvre effective.

La troisième partie de la thèse (chapitres 10 et 11) porte sur la validation du modèle proposé : d'abord, nous proposons une validation, sur le corpus Ozkan, du traitement d'un phénomène référentiel spécifique qui est l'interprétation des expressions d'altérité (*autre*). Au passage, nous élaborerons une méthodologie d'annotation de corpus qui dépasse les limitations imposées jusqu'alors par le cantonnement à la notion de coréférence, notoirement insuffisante pour tenir compte du statut référentiel de *autre*. Dans un deuxième temps, nous présenterons l'implémentation qui a été faite de notre modèle dans le cadre d'une collaboration avec la société *Thalès* (ex-*Thomson CSF*). Nous montrerons en particulier comment cette implémentation a pu s'intégrer dans une application de dialogue homme-machine réel portant sur un domaine extérieur à celui du corpus Ozkan, qui nous a servi de référence tout au long de ce travail.

Enfin, nous terminerons en ouvrant différentes perspectives qui nous sont apparues au cours de cette thèse : nous reviendrons en particulier sur la validation cognitive du modèle proposé, en détaillant des protocoles expérimentaux possibles et en faisant des hypothèses sur l'adaptation possible de notre modèle à la génération d'expressions référentielles. D'autres prolongements possibles seraient une reconsidération, en termes de décalage domanial, de ce que l'on peut entendre par « malentendu dialogique » et une étude multilingue comparative du fonctionnement des marqueurs référentiels.

2 La linguistique et l'irréductibilité des marqueurs référentiels

2.1 Introduction

Ce chapitre est consacré au système des catégories linguistiques associées à l'interprétation référentielle. Il s'agira d'une présentation dans une perspective linguistique, c'est-à-dire relevant de l'hypothèse selon laquelle chaque marqueur référentiel est irréductible. Cette irréductibilité se traduit par un principe d'interprétation propre à chacune des catégories, principe qui doit permettre de les délimiter les unes par rapport aux autres, tout en assurant une couverture maximale des différents usages observés pour une même catégorie.

L'étude des catégories référentielles du français nous amènera successivement à nous intéresser aux expressions indéfinies, définies, démonstratives et pronominales[16] : pour chacune d'entre elles, nous nous efforcerons de dégager, en nous appuyant sur des travaux linguistiques, ce qui la caractérise et ce qui permet de la distinguer des autres catégories. En même temps, nous nous attacherons à mettre en évidence comment ces caractéristiques permettent d'expliquer, dans un cadre cohérent, différents usages d'un même marqueur : ce problème se posera avec acuité pour les expressions définies, trop souvent considérées comme relevant de principes d'interprétation spécifiques pour des usages anaphorique, associatif ou encore générique.

Enfin, nous conclurons cette présentation par une synthèse qui propose d'aller au-delà des différences irréductibles, en constatant qu'il y malgré tout des points communs à toutes les catégories étudiées : d'une part, l'usage d'une expression référentielle est toujours conditionné par l'organisation contextuelle et d'autre part, l'isolement d'un référent par l'emploi d'une expression référentielle a toujours comme effet une réorganisation du contexte d'interprétation. A partir de là, il nous sera possible de classer les différents facteurs de conditionnement et de ré-organisation, apparus tout au long de ce chapitre, dans un schéma synthétique (Figure 4, page 45). Ce schéma nous servira non seulement de guide de lecture pour l'état de l'art de la modélisation cognitive et computationnelle des processus référentiels (chapitres 0, 4 et 5), mais aussi de référence pour la modélisation que nous proposerons par la suite (chapitres 6, 7, 8 et 9).

2.2 Les indéfinis

2.2.1 Principe d'interprétation des indéfinis

Selon F. Corblin (1987), le principe fondamental de l'interprétation des indéfinis de la forme *un, deux,..., n N* (*un triangle, deux triangles, quelques triangles*) consiste en une **opération de dénombrement**[17] sur la classe nominale de type *N* : ce qu'il est possible de dénombrer sont des sous-espèces (12), des individus (13) ou des quantités standardisés (14). Le dénombrement est effectué par des nombres définis (*un, deux,...*) ou des nombre indéfinis (*quelques, certains,...*) :

(12) Je cultive trois *haricots verts.* (Corblin, 1987)

(13) Il reste *deux haricots verts* sur l'assiette. (Corblin, 1987)

(14) Et *deux haricots verts* pour le 22 ! (Corblin, 1987)

[16] Nous excluons les noms propres de notre étude.
[17] ou de *prélèvement*, dans les termes de M. Galmiche (1986)

La caractéristique la plus importante de cette opération – permettant de séparer la classe des indéfinis des définis ou des démonstratifs – est son **indépendance vis-à-vis du contexte** : « *La règle interprétative centrale qui gouverne l'indéfini est que chacune des extractions possibles est strictement indépendante de toute autre, ne pouvant être définie que pour une classe N, un nombre n et relativement à une énonciation. Ce fonctionnement exclut toute connexion au contexte d'usage, ou entre deux extractions. Même la disjonction entre le produit d'extractions distinctes qui constituerait une forme de connexion interprétative n'a rien d'obligatoire, elle est seulement typique.* » (Corblin 1987 : 78)

L'introduction d'une nouvelle variable pour les indéfinis dans le cadre des sémantiques formelles (cf. par exemple Kamp et Reyle, 1993) essaie de refléter cette propriété d'indépendance contextuelle. Néanmoins, comme le souligne F. Corblin ci-dessus et dans son article « *La condition de nouveauté comme défaut* » (1994), cette opération demande à être aménagée, dans la mesure où le principe d'indépendance contextuelle ne peut même pas garantir la disjonction des extractions successives. Si la lecture disjonctive est plus probable tant que l'identité n'est pas marquée explicitement, elle peut se en effet se substituer à une lecture coréférentielle dans certaines configurations textuelles particulières, comme l'ont noté récemment L. Danlos (1999), B. Gaiffe (Danlos et Gaiffe, 2000) et nous-mêmes (Salmon-Alt et al., 2000).

Une deuxième caractéristique des indéfinis est de pouvoir s'interpréter dans le champ d'un opérateur de réitération ou de multiplication, lorsque le contenu propositionnel est vérifié plus d'une fois. Cela rejoint le problème posé par les analyses logiques en termes d'ambiguïté de portée. *Un homme* en (15) peut effectivement être analysé comme ayant une portée maximale (a) ou comme étant sous le champ d'un événement vérifié plusieurs fois (b).

(15) *Un homme* est venu à plusieurs reprises. (Corblin, 1987)

a. $\exists x$ (homme(x) $\wedge$ est_venu_à_plusieurs_reprises(x))
b. $\forall e$ (venue(e) $\Rightarrow \exists x$ (homme(x) $\wedge$ venir(e,x)))

Cette deuxième caractéristique permettra de départager plusieurs interprétations, traditionnellement attribuées aux indéfinis : l'interprétation spécifique, l'interprétation non spécifique et l'interprétation générique.

2.2.2 Interprétations spécifique, non spécifique et générique

L'interprétation spécifique suppose la portée maximale de l'expression indéfinie : il s'agit de la lecture hors champ d'un opérateur de réitération. On parle donc d'interprétation spécifique lorsqu'un énoncé comportant un groupe nominal de type *nombre + N* correspond à la mention de ce nombre d'individu de type *N*, en tout. Comme le mentionne F. Corblin (1987), cette lecture est l'interprétation fondamentale, dans la mesure où c'est la seule à être toujours disponible. Les tests linguistiques permettant de conclure à une lecture spécifique sont l'identification possible par un désignateur détaché (16) et la possibilité de réponse à la question « *De combien de N est-il affirmé quelque chose ?* » (17) :

(16) *Un homme* est venu à plusieurs reprises : Jean. (Corblin, 1987)

(17) *Trois hommes* sont venus à plusieurs reprises : trois.

Une deuxième possibilité est l'interprétation non spécifique de (16) et (17) : elle provient d'une lecture sous le champ d'un opérateur de réitération. Les tests linguistiques présentés ci-dessus ne s'appliquent plus, car il est impossible de connaître l'identité et le nombre précis des individus impliqués dans la répétition de l'événement en question. Il existe d'autres interprétations

habituellement considérées comme non spécifiques. Cela concerne les indéfinis impliqués dans un procès dont les arguments ne sont pas nécessairement fixés, comme c'est le cas pour les procès sous le champ d'une modalité :

(18) Il faudrait qu'*un homme* vienne m'aider. (Corblin, 1987)

Comme le note F. Corblin (1987), il est impossible de départager ce deuxième type d'emploi du premier par des critères pragmatiques de type « *Le locuteur n'a pas d'individu particulier en tête* ». Ce critère peut être faux pour des exemples comme (18). Considérer avec A. Reboul qu'un indéfini traduit simplement l'idée selon laquelle l'interlocuteur n'a pas besoin d'identifier l'individu en question (Reboul et al., 1998) semble être une meilleure solution, d'autant plus que cela se ramène au principe fondamental d'extraction qui ne stipule en rien l'extraction d'un individu particulier.

(19) *Un garçon* ne pleure pas.

(20) *Un carré* a quatre côté. (Corblin, 1987)

Le troisième type d'interprétation est l'interprétation générique, telle qu'en (19) et (20) : cet emploi s'insère facilement dans le paradigme proposé par F. Corblin, car il est simplement « *produit par un opérateur qui multiplie la validité d'un contenu propositionnel vérifié à chaque fois que l'on considère une valeur sur la classe* » (Corblin, 1987 : 50). Il s'agit donc d'un cas particulier de l'interprétation non spécifique, faisant intervenir un opérateur qui donne un procès vérifié pour un grand nombre de fois, voire toujours. De ce fait, cette lecture est incompatible avec des contextes propositionnels relatant un fait ponctuel (21) :

(21) *Un garçon* est venu acheter un gâteau.

En revanche, c'est la seule lecture pour des contextes faisant intervenir des propriétés définitoires, car ceux-ci ne permettent évidemment pas l'extraction d'un individu particulier (20). Enfin, l'usage générique d'un indéfini atteint l'espèce à partir de l'individu :

(22) *Un homme* a inventé la Théorie de la Relativité. (Corblin, 1987)

Il est en effet impossible d'utiliser un indéfini générique dans des propositions non vérifiées individuellement pour chaque élément extrait de la classe des *N* (22).

2.3 Les définis

2.3.1 Problèmes liés aux conditions d'unicité et de (pré-)existence

Depuis l'analyse logique des descriptions définies, proposée par B. Russell (1905), on considère que leur emploi est lié à deux conditions fondamentales : l'unicité et l'existence. Cette analyse a son origine dans l'intérêt que portait B. Russell à l'énigme référentielle « ontologique » (Linsky, 1967). B. Russell, considérant que la signification d'un nom propre est le porteur de ce nom, faisait l'observation suivante :

(23) *Le cercle carré* n'existe pas.

Pour une proposition comme (23), si la signification de la description est l'objet lui-même, il faudra admettre qu'elle est fausse. Il préfère donc considérer qu'une description définie n'a pas de signification en elle-même. Elle ne prendrait sa signification qu'en contexte. Pour un exemple tel que (24), la signification du défini le cercle carré serait une paraphrase (25), se traduisant par la formule logique de (26) :

(24) *Le cercle carré* est bleu.

(25) a. Au moins, un objet est cercle et carré.
b. Au plus, un objet est cercle et carré.
c. Quelque soit cet objet, il est bleu.

(26) $\exists x$ (cercle_carré(x) $\wedge$ $\forall y$ (cercle_carré(y) $\Rightarrow$ x = y)) $\wedge$ bleu(x)

L'analyse de B. Russell revient donc à considérer qu'une phrase ayant pour sujet une description définie est vraie si et seulement si le référent remplit les conditions d'unicité et d'existence et si ce que l'on en prédique est vrai. Cette analyse a été modifiée par P. Strawson (1977), pour qui existence et unicité font partie des présupposés, c'est-à-dire doivent être vérifiées, sous peine d'avoir une proposition sans valeur de vérité. Ce que les analyses centrées sur la détermination du référent (et non sur les conditions de vérité) en retiennent généralement, c'est une contrainte de pré-existence et d'unicité restreinte à un contexte donné (discours, situation). Indépendamment des orientations particulières des analyses, ces conditions peuvent être paraphrasées de la façon suivante : « *Il faut qu'il existe un et un seul x qui soit un N dans le contexte d'interprétation.* » Or, cette formulation est à l'origine de plusieurs problèmes, concernant à la fois la condition de pré-existence et celle d'unicité.

En ce qui concerne la contrainte de pré-existence, B. Gaiffe, A. Reboul et L. Romary (1997) posent le problème du statut des anaphores associatives (*un bar – le patron*). Nous traiterons ce point plus en détail par la suite (2.3.5), mais notons d'ores et déjà qu'une description définie peut introduire un nouveau référent de discours (*le patron*), ce qui est *a priori* contradictoire avec une condition de pré-existence.

Un deuxième souci est le fait que les contraintes d'unicité et de pré-existence ne permettent en rien de faire des prédictions sur l'emploi du défini dans des configurations discursives telles qu'en (27) : les deux conditions – unicité et pré-existence – sont parfaitement remplies en (a) et (c), mais le défini semble beaucoup plus approprié en (a) qu'en (c), en particulier si on le compare à un pronom dans le même site.

(27) a. Un homme et une femme entrèrent dans un bar. *L'homme/ ? il* portait un chapeau mou.
b. Un homme entra dans un bar. *? L'homme/ il* portait un chapeau mou.

En ce qui concerne la condition d'unicité, F. Corblin (1987) relève d'autres problèmes : parallèlement aux observations de B. Gaiffe et al. (1997) sur la délimitation des définis par rapport aux pronoms, F. Corblin constate que le principe d'unicité ne permet pas de distinguer les définis par rapport aux démonstratifs et indéfinis : « *Après tout* », constate-t-il, « *le N, ce N, un N sont tous au singulier et impliquent tous d'une manière ou d'une autre singularité* » (Corblin, 1987 : 102). Ensuite, l'unicité ne s'applique pas au pluriel et aux termes de masse, pour lesquels le concept de « totalité » a été introduit par Hawkins (1978). Cela rapproche le défini du quantificateur universel, mais n'explique plus la différence avec une expression linguistique concurrente qui est *tout* et surtout pas pourquoi le défini générique ne peut pas être paraphrasé par *tous les N* :

(28) L'homme a marché sur la lune.
*Tous les hommes ont marché sur la lune. (Corblin, 1987)

Par ailleurs, le défini générique semble réfractaire à une analyse standard en termes d'unicité, comme le montre (29) qui ne peut pas s'analyser ainsi :

(29) L'homme a deux bras.
$\exists x$ homme(x) $\wedge$ avoir_deux_bras(x)

Afin de résoudre ce problème posé par l'analyse du générique, il faudra, comme le propose F. Corblin (1987 : 104), modifier le principe d'analyse en : « *Il existe un x et un seul que N a le pouvoir de séparer des autres x* ». Formulée de telle façon, l'unicité ne s'applique plus

obligatoirement à un et un seul individu particulier de type N, mais elle peut aussi s'appliquer à l'espèce et restituer ainsi sa lecture générique à l'exemple (29).

2.3.2 *Principe d'interprétation des définis*

Afin de pallier les problèmes rapportés dans la section précédente, aussi bien B. Gaiffe et al. (1997) que F. Corblin (1987) aboutissent à un nouveau principe d'interprétation pour les définis. F. Corblin considère que « *Dans* le N, le *s'interprète comme instruction d'avoir à conférer capacité désignative au contenu lexical préfixé* N. *Cela signifie que c'est par l'intermédiaire de ce contenu et par ce seul intermédiaire qu'un groupe nominal défini doit être en mesure de séparer un x des autres* » (1987 : 106). La piste proposée par B. Gaiffe et al. est la suivante : « *Nous imposerons [...] pour le traitement d'un groupe nominal défini non pas une présupposition d'existence du référent, mais une présupposition d'existence d'un contexte dans lequel le référent puisse être créé* » (1997 : 88).

Même si ces deux propositions peuvent sembler, au premier abord, loin l'une de l'autre, elles présentent des points communs importants : F. Corblin considère que *le N* doit séparer « *un x des autres* ». Cela **suppose donc l'existence d'un domaine d'interprétation** comprenant *x* et d'autres individus, comme l'imposent B. Gaiffe et al. Ensuite, rien ne restreint la nature des *x* à des individus particuliers. Il peut donc s'agir d'un domaine d'individus particuliers, mais aussi d'un domaine d'espèces, ce qui permet respectivement la référence générique et la référence spécifique. Enfin, la seule condition sur ce domaine n'est pas une condition de pré-existence du référent, mais une structuration permettant d'en **extraire un et un seul N** (espèce ou individu). Cette extraction peut évidemment s'opérer sur un référent pré-existant dans le domaine, mais elle peut aussi amener un référent n'étant pas jusqu'alors présent dans le contexte immédiat. Afin que cette extraction puisse se faire, le contenu lexical préfixé *N* doit fonctionner en tant que différenciateur. Cela est le rôle de la référence virtuelle, notion que nous détaillerons dans la section suivante, avant de montrer que le principe d'interprétation, formulé comme ici, couvre bien à la fois les emplois génériques et les emplois spécifiques du défini.

2.3.3 *Référence virtuelle*

La notion de référence virtuelle rejoint le *sens* frégéen : G. Frege , dans son article « Sinn und Bedeutung » (1892) distingue entre *sens* (Sinn) et *référence* (Bedeutung). Cette distinction lui sert, entre autres, pour expliquer l'énigme référentielle « de l'identité » (Linsky, 1967). Il explique ainsi pourquoi (30) n'est ni tautologique, alors que cette proposition affirme l'identité de deux choses, ni fausse, bien que deux choses ne soient jamais égales.

(30) Venus est l'étoile du matin.

La raison en est la distinction entre sens et référence. La référence d'une expression est l'objet nommé par elle. Le sens d'une expression est, selon G. Frege, saisie par toute personne suffisamment familière avec la langue et inclut le *mode de présentation*. Dans l'exemple (30), les deux expressions *Vénus* et *étoile du matin* ont effectivement le même référent, mais celui-ci n'est pas « donné » de la même façon : elles ont donc des sens différents et c'est cela qui apporte une connaissance effective garantissant l'informativité de la proposition.

C. Milner (1978 : 26) introduit les notions de *référence virtuelle* et *référence actuelle* pour respectivement *sens* et *référence*. Cette terminologie souligne davantage la relation entre les deux notions : une unité lexicale nominale hors emploi n'entretient pas de relation directe avec un segment de la réalité qui serait son référent, mais elle n'est pas non plus sans rapport avec la référence, dans la

mesure où elle détermine en effet le type des segments de réalité qui peuvent être désignés par elle. Elle impose donc un ensemble de conditions que le référent doit satisfaire pour qu'il y ait référence et c'est cet ensemble qui est appelé *référence virtuelle.*

F. Corblin (1987) attire l'attention sur un fait problématique : la référence virtuelle, telle que présentée par C. Milner, se définit en fonction des propriétés qu'un référent doit posséder pour pouvoir être désigné. Or, cela n'est pas compatible avec le fonctionnement des démonstratifs dont on verra par la suite qu'ils ne réfèrent pas en fonction des propriétés du référent (cf. la section 2.4) et pour les indéfinis qui ne désignent pas. La proposition est alors de considérer la référence virtuelle simplement comme l'invariant contribuant à l'opération référentielle propre à chaque type de marqueur.

2.3.4 Du défini générique aux définis spécifiques

Le principe d'interprétation des groupes nominaux définis tel que nous l'avons présenté ci-dessus (2.3.2) permet facilement de dériver les conditions d'interprétation pour les définis génériques comme dans les exemples (28) ou (29). F. Corblin les formule ainsi : « Le *impose d'avoir à conférer capacité désignative au contenu nominal préfixé et l'interprétation référentielle est saturée par la désignation de l'espèce N* » (1987 : 116). Ce principe est donc basé sur le fait que *le N* peut, dans toutes les circonstances et indépendamment d'un contexte spécifique, séparer une et une seule espèce des autres. Cette propriété distingue très nettement l'usage générique de l'indéfini, examiné dans la section 2.2.2, de l'usage générique des définis : ce dernier est toujours possible, dans la mesure où *le N* a toujours la capacité d'isoler une espèce, alors que l'indéfini générique est tributaire d'un contexte « multiplicateur ». Il s'ensuit naturellement que les possibilités d'interprétation d'un indéfini générique sont plus restreintes que celles d'un défini générique : plus précisément, elles sont limitées à des propositions vérifiées distributivement sur les individus de l'espèce. Le défini générique n'est pas incompatible avec cette propriété, mais il couvre d'autres cas qui n'admettent pas un indéfini : cela concerne les propriétés vérifiées collectivement par l'espèce et des propositions mettant en cause l'espèce elle-même. Le Tableau 5 propose une comparaison des définis et indéfinis génériques en fonction des propriétés de la proposition hôte :

Propriétés de la proposition	***Défini générique***	***Indéfini générique***
proposition vérifiée distributivement sur l'espèce	L'homme a deux bras.	Un homme a deux bras.
proposition vérifiée collectivement sur l'espèce	Et Dieu créa la femme.	*Et Dieu créa une femme.
mise en cause de l'espèce	La Licorne n'existe pas.	*Une Licorne n'existe pas.

Tableau 5 : Conditions d'interprétation pour les définis et indéfinis génériques

Comme nous l'avons vu, une interprétation générique des définis est toujours possible, car celle-ci ne dépend pas d'un contexte spécifique. Pour une interprétation spécifique, en revanche, « *il faut que le contexte d'usage permette de déterminer un domaine d'interprétation dans lequel* N *puisse isoler un individu, soit un particulier, soit une sous-espèce. Cela étant,* le N *est saturé et désigne l'individu en question. La saturation consiste à déterminer un domaine dans lequel N reçoive effectivement fonction désignative.* » (Corblin, 1987 : 116). Il s'agit dans ce cas d'un usage non générique et contextuel qui donne lieu à une interprétation relativement à un domaine contextuel restreint où *le N* isole un *N* particulier des autres objets du domaine. Dans les sections suivantes, nous allons examiner les deux possibilités de saturation référentielle lors d'une interprétation spécifique.

2.3.5 *Mode de saturation pour les définis spécifiques*

La première possibilité de saturation référentielle pour un défini spécifique est donnée par un emploi anaphorique, comme en (31) :

(31) I_1 donc tu vas prendre *le petit rond* pour faire la tête
I_2 et de chaque côté *du petit rond* tu vas mettre des petites barres (C8Eglise)

Le petit rond fonctionne ici en tant qu'anaphore, car il reprend une mention (*le petit rond*), appelée traditionnellement son antécédent, qui la précède dans le texte. Le domaine d'interprétation d'une anaphore est donc donné par des mentions explicites dans le discours précédent. Pourtant, nous défendrons, comme F. Corblin (1987) et B. Gaiffe et al. (1997), le point de vue selon lequel ce n'est pas la saturation standard des descriptions définies : en effet, le premier problème de cette caractérisation est précisément d'enlever toute spécificité des définis par rapport aux démonstratifs et pronoms personnels (cf. l'exemple (27)). Le deuxième problème est qu'elle n'arrive plus à expliquer les emplois génériques du défini qui, eux, ne sont jamais utilisés en tant qu'anaphores. Troisièmement, cette caractérisation ne couvre pas l'usage associatif des définis, auquel elle doit attribuer un statut particulier.

Par la suite, nous étudierons plus en détail cet usage associatif. Cela nous amènera à conclure que la reprise anaphorique n'est qu'un cas particulier d'un principe d'interprétation englobant l'usage anaphorique et associatif des définis spécifiques.

La seconde possibilité de saturation des définis spécifiques est donc celle que l'on appelle, depuis G. Guillaume (1919), l'usage associatif. L'exemple (32) en donne deux illustrations, avec *la tête* (I_2) et *les nattes* (I_4) qui sont interprétées relativement par rapport à (ou « en association » avec) *la petite fillette* (I_1).

(32) I_1 et en dessous de ce soleil on va faire *la petite fillette*
I_2 donc tu vas prendre le petit rond pour faire *la tête*
I_3 et de chaque côté du petit rond tu vas mettre des petites barres
I_4 de façon à faire les... *les nattes* quoi (C8Eglise)

G. Guillaume caractérise cet emploi de la façon suivante : *« A la limite, l'article d'extension est applicable à toute chose qui, étant donné le sujet, s'annonce comme déductivement nécessaire »* (1919 : 165). La conception linguistique moderne (résumée par Kleiber et al., 1994) considère l'anaphore associative comme une configuration discursive contenant une expression référentielle étant anaphorique (qui identifie son référent grâce à des informations présentes dans le texte) sans être coréférentielle (il n'y a pas identité des référents). Des définitions plus étroites imposent des contraintes supplémentaires, par exemple sur la réalisation linguistique de l'antécédent (groupe nominal) ou sur le type des relations entre antécédent et anaphore (méronymique).

Contrairement à l'usage anaphorique en reprise directe, l'ensemble d'interprétation n'est donc pas fourni en extension par un ensemble de mentions explicites. Néanmoins, ces emplois sont prédictibles par l'hypothèse suivante : les définis spécifiques n'exigent pas que le contexte fournisse textuellement le désignatum, mais un domaine d'interprétation limité ayant suffisamment de propriétés empiriques pour permettre à la référence virtuelle d'en isoler un individu. Cette hypothèse inclut les interprétations par reprise d'un élément textuel, car un domaine donné en extension par des mentions explicites permet *a fortiori* d'en isoler un par reprise. Dans le cas des interprétations associatives, ce sont les connaissances sur la composition d'ensembles stéréotypiques qui permettent de extraire un élément : en (32), *la petite fillette* introduit un référent dont on peut déduire qu'il est normalement constitué d'un certain nombre de parties du corps. Celles-ci forment donc un ensemble permettant d'isoler un référent pour la référence virtuelle *tête*.

Définir le principe d'interprétation des définis spécifiques comme une opération d'extraction d'un référent pouvant être isolé sur la base de sa référence virtuelle d'un ensemble contextuel hétérogène (Corblin, 1987 ; Gaiffe et al., 1997) a plusieurs avantages : premièrement, il est compatible avec le fonctionnement général des définis (cf. 2.3.2). Deuxièmement il couvre les emplois spécifiques (anaphoriques et associatifs). Troisièmement, il permet de prédire la distribution du pronom et du défini en (27) : en admettant d'une part que le rôle du pronom est de référer à un objet saillant dans une perspective de continuité discursive (thèse défendue en 2.5.3) et d'autre part, que le calcul référentiel du défini demande plus d'efforts que celui du pronom, la différence entre (a) et (b) s'explique par la mise à disposition d'un ensemble contextuel justifiant l'extraction par un défini en (a), alors que ce n'est pas le cas en (b).

De plus, d'autres arguments viennent appuyer cette hypothèse sur le fonctionnement des définis comme opérateurs d'extraction : d'abord et contrairement à ce que pourraient laisser croire des approches traitant le défini spécifique comme marqueur prioritairement anaphorique (cf. par exemple Bosch et Geurts, 1990), les reprises textuelles ne constituent pas une majorité des emplois du défini. M. Poesio et R. Vieira (1998) ont montré sur un corpus de textes journalistiques en anglais que seulement 30 % des définis sont employés de façon anaphorique stricte[18] (*un chat – le chat*). Les autres emplois se répartissent comme suit : 18 % sont des emplois coréférentiels avec une divergence sur la tête nominale (*un chat – l'animal*) ou des anaphores associatives à antécédent textuel (*une fille – la tête*), les 52 % restant sont des extractions référentielles à partir de la situation de communication ou de connaissances communes. Cela signifie que moins de la moitié des emplois du défini servent à la reprise anaphorique.

Un autre argument en faveur de l'hypothèse non anaphorique est fourni par une observation sur l'orientation du processus référentiel de l'anaphore associative. G. Kleiber et al. (1994) constatent que la plupart des anaphores associatives « roulent » dans le sens « tout – partie », comme dans *une fillette – la tête* ou *un village – l'église*. A partir de là, les auteurs s'interrogent sur la pertinence de cette orientation et défendent l'hypothèse qu'elle est systématique. La raison en est « *l'impossibilité d'avoir une expression anaphorique qui apporte des informations sur le référent non disponibles ou non inférables de l'expression source antérieure* » (Kleiber et al., 1994 : 137). En d'autres termes, l'anaphore associative suppose bien son domaine d'interprétation.

(33) *Le pied* est abîmé, mais *la chaise* est toujours solide. (A. Azoulay, repris dans Kleiber et al., 1994)

(34) *? Un pied* est abîmé, mais *la chaise* est toujours solide.

Des exemples apparemment contraires à ce constat (33) s'expliquent en fait en tant que « cataphores associatives » : comme pour les anaphores associatives classiques, l'impossibilité de mettre la partie (*le pied*) à l'indéfini montre bien qu'il s'agit non pas de l'antécédent, mais d'une mention dépendante d'un domaine, fût-il introduit dans la suite du texte (34). Cela signifie que l'ordre « naturel » des anaphores associatives va du tout à la partie, ce qui confirme l'hypothèse d'une interprétation des définis dans un domaine présupposé.

Enfin, si l'interprétation par reprise était le principe d'interprétation fondamental et l'interprétation associative l'exception, on n'expliquerait pas pourquoi, dans des cas admettant les deux lectures, l'association semble l'emporter. Or, c'est précisément ce qui se passe dans l'exemple (35), comme le note F. Corblin (1987),

(35) *Un garçon* a volé la voiture et *le voleur* a été puni[19]. (Corblin, 1987)

[18] en « reprise fidèle » (même tête nominale, même référent)

[19] L'exemple doit être pris avec prudence : un certain nombre de personnes que nous avons interrogées ne perçoivent pas du tout la possibilité d'une interprétation anaphorique.

Une dernière observation, également apportée par F Corblin (1987), concerne les groupes nominaux modifiés : en cas de disjonction de la propriété modifiante, il est très difficile d'admettre une lecture coréférentielle :

(36) *Cette rose rouge* me gène. Je vais jeter *la rose fanée.* (Corblin, 1987)

Cette observation qui ne s'explique pas par une théorie anaphorique du défini, est compatible avec l'hypothèse du défini vu comme extracteur dans un domaine : ici, *la rose fanée* suppose un domaine d'interprétation constitué de roses pouvant se distinguer selon leur degré de décomposition. Rien ne contraint alors la ré-identification d'un objet introduit en tant que *rose rouge* (malgré les informations prédicatives qui semblent plutôt favoriser une telle reprise). Au contraire, le fait que celle-ci soit présentée comme étant *rouge* implique un effort d'interprétation supplémentaire, car cette information ne semble plus être pertinente pour la suite de l'interprétation.

2.3.6 *Résumé du principe d'interprétation des définis*

L'objectif de cette section était de présenter le fonctionnement référentiel des groupes nominaux définis. Plusieurs propositions, en particulier celle de F. Corblin (1987) et celle de B. Gaiffe et al. (1997) ont pour but de dépasser les inconvénients des approches considérant les définis comme des marqueurs prioritairement anaphoriques. Ces dernières ne prédisent en effet ni les emplois génériques, ni les emplois associatifs et ne réussissent pas à délimiter les définis par rapport à des marqueurs proches, en particulier le pronom personnel.

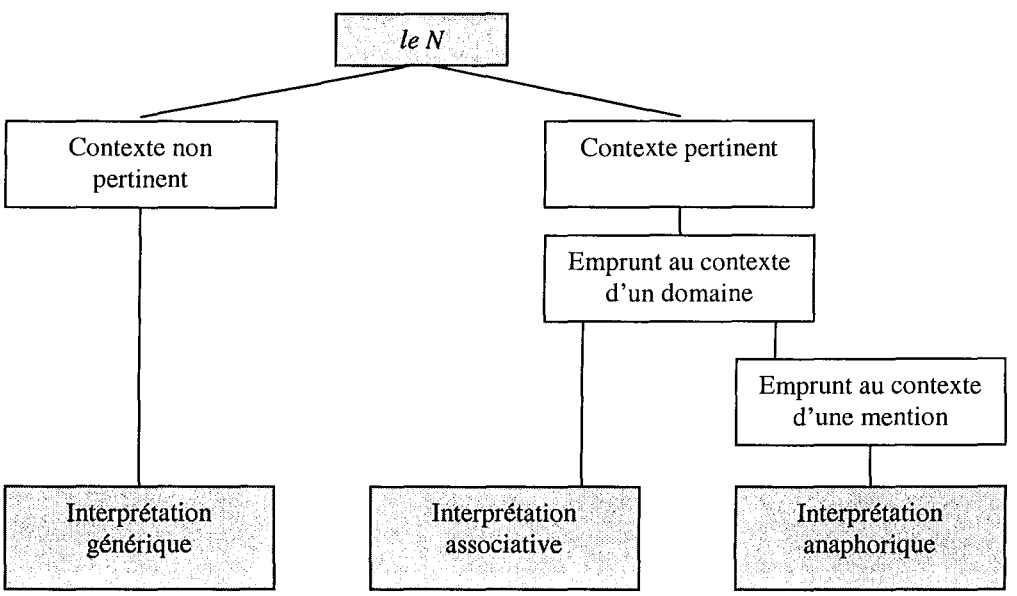


Figure 3 : Principe d'interprétation des groupes nominaux définis

Le principe retenu peut être formulé ainsi : une expression définie *le N* présuppose un domaine d'interprétation hétérogène à l'intérieur duquel un individu peut être isolé de tous les autres en vertu de la référence virtuelle associé à *N*. En dehors de tout contexte, il est toujours possible d'isoler l'espèce *N* des autres espèce (interprétation générique). Cependant, dans un contexte particulier et pertinent, le défini *le N* peut être saturé par emprunt au contexte d'un domaine d'interprétation, à l'intérieur duquel il extrait son référent soit sur la base de connaissances empiriques sur la composition de ce domaine (interprétation spécifique associatif), soit par emprunt au contexte d'un domaine constitué par des mentions explicites (interprétation spécifique anaphorique). La Figure 3 donne un aperçu schématisé de ce principe.

2.4 Les démonstratifs

2.4.1 Identification du référent par reprise

Le fonctionnement référentiel propre des groupes nominaux démonstratifs se définit par le statut définitoire de la reprise (Corblin, 1987). Contrairement à la saturation référentielle des définis – ne passant pas obligatoirement par la reprise d'une mention antérieure, mais par l'isolement, selon la référence virtuelle, d'un élément à l'intérieur d'un domaine d'interprétation (cf. 2.3.2) – le démonstratif demande qu'un objet ait été mentionné récemment dans le contexte d'usage ou que la situation de discours l'ait rendu saillant. Il ne circonscrit donc pas son domaine d'interprétation, mais a **besoin d'un contexte qui doit lui-même fournir un moyen d'isoler l'objet désigné** par le démonstratif, soit par mention récente, soit par un geste de désignation concomitant. Ce recours nécessaire au contexte est par ailleurs exploité dans des textes littéraires : puisque l'usage du démonstratif exige l'existence du référent (est *token-réflexif*, selon W. de Mulder), il peut être employé comme moyen stylistique pour amener le lecteur à reconstruire l'univers du récit (de Mulder, 1998).

2.4.2 La possibilité de reclassification

Le principe d'identification par des critères de saillance extérieurs permettent une redéfinition du rôle de la référence virtuelle du syntagme démonstratif : puisque celle-ci ne participe pas à l'identification du référent, elle est disponible pour une reclassification du référent, identifié auparavant sur des critères externes. Selon F. Corblin (1987), la valeur de cette reclassification n'est pas soumise à des contraintes linguistiques. Il peut s'agir d'une simple répétition (37), d'une généralisation[20], d'une propriété empirique compatible avec le type d'objet antérieurement isolé (38) ou de n'importe quelle autre classification, à condition qu'elle soit motivée, dans la situation de communication, par une analogie perceptible avec l'objet désigné (39).

(37) I_1 au dessus de ça tu vas mettre un rond qui correspondra au soleil
I_2 et *ce rond* il faut qu'il soit au dessus de la deuxième pyramide (C8Egypte)

(38) I_1 donc tu prends le gros rond et tu le mets à droite de l'église
I_2 et en dessous de *ce soleil* on va faire la petite fillette (C8Eglise)

(39) Deux arbres encadraient l'entrée et *ces sentinelles* dormaient. (Corblin, 1987)

Selon F. Corblin, la limite de la possibilité de reclassification ne se traduirait pas par un échec d'identification référentielle de l'objet, mais par une erreur de classification, ce qui soulignerait bien l'indépendance de ces deux opérations. Néanmoins, lorsque plusieurs référents sont possibles, le degré de compatibilité entre la référence virtuelle et les objets disponibles dans le contexte immédiat semble pouvoir influer sur le choix : ainsi, en (38), *ce soleil* sélectionnera, parmi les antécédents possibles, non pas *l'église*, mais *le gros rond.*

2.4.3 Effets interprétatifs

Le principe d'interprétation des syntagmes démonstratifs se résume donc de la façon suivante : *ce* exige de fixer un référent à proximité dans le contexte linguistique ou situationnel. Le contenu linguistique n'intervient pas dans ce processus et est alors libéré pour jouer le rôle d'un classificateur.

[20] Contrairement à F. Corblin qui accepte des exemples comme « *Un animal courait. Ce chien ...* », il nous semble qu'un enchaînement par un démonstratif qui restreint la référence virtuelle de son antécédent est difficilement acceptable. G. Kleiber (1994) défend ce même point de vue.

Par le fait même de classifier un référent comme étant un *N*, le démonstratif *ce N* suppose l'existence d'une classe virtuelle *N*. L'opération de reclassification peut alors être considérée comme opposant le référent courant aux autres éléments ce cette classe. De ce fait, elle instaure un contraste interne à l'intérieur de cette classe (Corblin, 1987), même si ce contraste reste souvent latent (37). Il peut néanmoins devenir essentiel dans la suite du discours, en particulier lors de références à d'autres éléments de la même classe comme en (40) et (41) :

(40) Une voiture était rangée devant la porte. C'est *cette voiture* que j'ai prise et non *la tienne*. (Corblin, 1987)

(41) [plusieurs triangles à l'écran] : Mets *ce triangle* à droite et supprime *les autres*.

Cependant, comme l'ont remarqué G. Kleiber (1986, 1988, 1994) et W. de Mulder (1998), le démonstratif peut s'employer sans qu'il y ait une classe de *N*, par exemple en reprise directe (37) ou même en reprise d'un élément appartenant à une classe hétérogène (42) :

(42) Il y avait la mer de Chine, la mer Rouge, l'océan Indien, le canal de Suez [...]
Mais avant tout il y avait *cet océan.*

(M. Duras, *L'amant*, p.136 ; relevé par de Mulder, 1998)

La condition justifiant l'usage d'un démonstratif n'est donc pas l'isolement d'un élément à l'intérieur d'une classe, mais **la mise en relief du référent isolé** (de Mulder, 1998). Cette mise en relief traduit une rupture qui peut être par exemple une rupture du cadre discursif, un changement thématique ou l'introduction d'un nouvel univers de croyance. G. Kleiber (1986) défend la même idée lorsqu'il définit la particularité du démonstratif en reprise immédiate par une saisie directe du référent, indépendamment de la circonstance d'évolution tracée par la phrase précédente (cf. le chapitre 1, section 1.2.1). La notion de rupture discursive, essentielle pour délimiter les usages du démonstratif de ceux du pronom personnel (2.5.3), est en rapport direct avec le mode de saisie du référent : chaque apparition d'un démonstratif demande un nouvel appariement avec un objet contextuel qui est ensuite renommé ou reclassifié. Cela entraîne donc une séparation du référent de la structure l'ayant rendu saillant et permet de l'insérer dans un nouveau cadre.

(43) I_1 donc le deuxième dessin représente une route euh
M_1 ouais
I_2 donc faite avec des barres verticales et euh donc au bord de *cette route*, il y a deux maisons, donc une maison qui se trouve à gauche de euh de *cette route* et une autre à droite

(C9 Forêt)

Le caractère maladroit de l'exemple (43) illustre à la fois l'hypothèse d'une nouvelle saisie systématique et de la rupture discursive : si la nouvelle saisie lors du premier emploi de *cette route* peut se justifier par un changement du point de vue (composition géométrique vs. figuratif), le deuxième démonstratif ne se justifie que difficilement : il n'y a ni reclassification ni rupture, la perspective *bord de route* étant déjà introduite.

2.5 Les pronoms personnels de 3ème personne[21]

La problématique référentielle du pronom personnel a été résumée par G. Kleiber (1994) par la question suivante : « *Comment une expression sémantiquement réduite à des marques de nombre et de genre peut-elle trouver son référent ?* ». Les réponses à cette question varient selon que l'on se positionne dans une optique textuelle (*il* endophorique) ou mémorielle (*il* mémoriel). Nous exposerons

[21] Nous n'aborderons pas dans cette thèse les pronoms personnels déictiques *je/tu/nous/vous*. Pour une prise en compte de ces marqueurs dans le cadre du calcul référentiel en DHM, cf. par exemple Gaiffe et al. (2000).

brièvement ces deux versions, avant de présenter la proposition de G. Kleiber qui tente de pallier les inconvénients de celles-ci.

2.5.1 La version textuelle

La version textuelle « dure » consiste à considérer le pronom personnel comme un désignateur endophorique qui recrute son référent à travers une expression antécédente présente dans le contexte. Le pronom fonctionnerait comme substitut, en vertu d'un principe d'économie textuelle (cf. par exemple Dubois, 1965). Or, la thèse de la lecture substitutive ne se maintient pas, comme l'attestent des exemples comme (44) :

(44) *Un homme* est entré. **Un homme* / *Il* portait un chapeau. (Kleiber, 1994)

La version textuelle « molle » considère alors que le pronom fonctionne comme un désignateur second qui ne peut par lui-même fournir un principe d'identification référentielle. Par conséquent, il l'emprunte à son antécédent, sans que celui-ci puisse toutefois remplacer le pronom en tout lieu.

Le problème principal de cette version est que l'on n'attribue pas au pronom un principe de fonctionnement – ou mode de donation, comme dirait G. Kleiber – original qui permettrait de le caractériser par rapport aux autres types d'expressions référentielles, définis et démonstratifs en particulier. En relation avec cette interrogation, un certain nombre de questions restent ouvertes : Comment trouver le bon référent lorsqu'il n'y a pas accord grammatical (45) ou lorsque plusieurs antécédents sont possibles (46) ? Comment traiter des *il* sans antécédent textuel, comme dans les glissement au générique (47), des anaphores divergentes (48) , des *ils* collectifs (49) et des référents *in absentia* (50) ?

(45) I_1 et tu vas reprendre encore *une ligne verticale*
M_1 une grande ?
I_2 une grande oui
I_3 et tu vas *le* placer en parallèle avec un petit écart (C5Route)

(46) Versez alors dans la marmite *les dés de carottes, les gousses d'ail écrasées* et le bouquet garni. Mélangez le tout juste un instant pour *les* imprégner d'huile d'olive. (http://www.cuisine.free)

(47) J'ai acheté *une Toyota*, parce qu'*elles* sont robustes. (Kleiber, 1994)

(48) I_1 ensuite on a les pyramides dans le désert alors il faut que tu prennes les deux euh enfin le triangle
I_2 *le gros triangle...* voilà et tu *le* reprends une deuxième fois un peu décalé par rapport au premier (C8Egypte)

(49) *Ils* sont encore augmenté les impôts. (Kleiber, 1994)

(50) [en parlant de l'ordinateur] : *Il* est un peu lent, mais c'est pas grave. (C7Route)

2.5.2 La version mémorielle

Selon la version mémorielle, qui essaie en particulier de répondre aux interrogations sur les référents *in absentia*, un pronom tel que *il* renvoie à un référent saillant, qu'il soit donné textuellement ou par la situation. En même temps, cette version fournit une réponse au problème des poulets brûlés vifs, pommes écrasées et autres référents évolutifs (de Mulder et Tasmowski-De Ryck, 1997). Enfin, invoquer la saillance comme principe de fonctionnement des pronoms permet de marquer leur spécificité par rapport aux autres marqueurs référentiels.

Cependant, les problèmes liés à cette approche sont les suivants : comment traiter un pronom lorsqu'il y a plusieurs référents saillants ? Si on ne renvoie pas à un antécédent textuel, comment sont alors représentés les référents ? Comment expliquer des emplois cataphoriques, puisque le référent est supposé être déjà saillant ? Enfin, la contrainte de saillance préalable à l'interprétation n'est plus compatible avec le traitement des exemples dans lesquels le référent du pronom n'est pas directement donné, mais calculé à partir de ce qui est donné, comme en (47) ou (48). Face à ce dernier problème, A. Reboul (1990) propose une extension du mécanisme inspiré de la théorie de la pertinence (Sperber et Wilson, 1989) : un pronom peut renvoyer à « *à tout ce qui est manifeste au locuteur au moment de l'énonciation et que ce locuteur croit [...] mutuellement manifeste* » (Reboul, 1990 : 224). Le nouveau problème est alors la sur-puissance du mécanisme, car il prédit ainsi des cas non souhaités.

2.5.3 La version de G. Kleiber (1994)

G. Kleiber attribue à *il* le statut d'un marqueur référentiel à part entière, tout en essayant d'éviter l'écueil de la sur-puissance de la version mémorielle par une « re-sémanticisation » du pronom : *il* co-détermine ses référents, il apporte de l'information par la projection d'un point de vue particulier, voire l'introduction d'un nouveau référent dans certains cas, comme en (49) ou (50). De plus, comme tous les autres marqueurs, il possède des instructions sémantiques et procédurales sur sa façon de récupérer le référent, qui son sens descriptif et son sens instructionnel.

- sens descriptif (nombre et genre) : le nombre restreint a priori l'extension du pronom. Des emplois génériques sont cependant possibles sous certaines conditions (mention préalable, prédicat compatible, relation causale,...). Indépendamment de la marque masculin ou féminin du genre, mais parce qu'il indique le genre, le pronom est prévu pour référer aux choses qui sont classifiées, c'est-à-dire qui sont reconnues comme faisant partie de telle ou telle catégorie[22]. Dans le cas des humains, qui sont déjà classifiés en tant que tels, la variation pour le genre est disponible, à condition que cela apporte des informations supplémentaires. Dans le cas des non humains, le référent doit être classifié soit par mention textuelle, soit par saillance situationnelle. Dans ce dernier cas, la classification se ferait sur la base de sa dénomination basique (Rosch, 1978) :

- sens instructionnel : il faut rechercher le référent de *il* dans un cadre saillant, c'est-à-dire présent

- dans le focus de l'interlocuteur. Le référent doit y être impliqué comme un actant principal, c'est-à-dire jouer le rôle d'argument. Il faut que la phrase comportant *il* soit un prolongement de cette structure saillante.

Le rôle spécifique du pronom peut donc être considéré en termes d'économie, non pas en ce qui concerne l'identification des antécédents discursifs – le calcul pour récupérer *il* indirect serait même plus coûteux – mais appliqué à des représentations dans le focus discursif : le pronom **saisit un référent saillant lui-même ou présent dans une situation saillante, dont on va parler en continuité avec ce qui l'a rendu saillant**. C'est cela qui le distingue des autres marqueurs qui opèrent des assignements référentiels plus coûteux : calcul évaluatif justificateur d'unicité dans un domaine donné pour le défini et appariement référentiel par procédure indexicale pour le démonstratif.

[22] Ce point justifie en partie l'opposition de *il* à *ça* (cf. Kleiber (1994) ou Corblin (1995) pour une comparaison de *il* explétif avec *ça*).

2.6 Les groupes nominaux sans nom

2.6.1 Propriétés du paradigme

Dans une série de travaux, F. Corblin (1990, 1995) fait l'hypothèse d'une classe morpho-syntaxique et interprétative particulière, celle des groupes nominaux sans nom. Il est intéressant de noter que D. Creissels (1995) – en remettant en cause des notions syntaxiques issues de la tradition grammaticale scolaire et dans le but de dégager des notions syntaxiques fondamentales et universelles – aboutit à une classe morpho-syntaxique semblable : les syntagmes nominaux issus d'une opération de réduction discursive. Le paradigme de ces syntagmes sans nom comprend, aussi bien dans la classification de F .Corblin que dans celle de D. Creissels :

- les groupes nominaux elliptiques : *le rouge, le grand,...* ;
- les groupes nominaux indéfinis à *en* quantitatif : *j'en veux un bleu, j'en prends un, ...* ;
- les pronoms possessifs : *la sienne, le tien,...* ;
- et les composés de *celui* : *celui-ci, celui qui..., celui de....*

Il s'agit de syntagmes nominaux dont la tête nominale n'est pas fixée *in situ*, mais par emprunt au contexte. D'après D. Creissels, ils sont le résultat d'une suppression « *du terme déterminé, celui-ci pouvant rester sous-entendu au sens où le déterminant peut continuer d'être interprété comme s'appliquant à un terme structurellement présent mais non explicité, dont le contexte permet de rétablir l'identité* » (1995 : 76). Dans la classification syntaxique que propose D. Creissels, cette proposition a l'avantage de regrouper dans une même classe une partie de l'ancienne catégorie hétérogène des pronoms : ceux qui entretiennent une relation systématique avec les adjectifs, c'est-à-dire les pronom indéfinis, démonstratifs et possessifs.

La caractéristique commune de ces formes est la récupération de la tête nominale dans le contexte textuel ou situationnel (51). Cette récupération impose une contrainte d'identité sur le genre, mais pas sur le nombre (52). Au niveau syntaxique, la dislocation à droite en *de N* permet de restituer la source de la tête lexicale (53). En cas d'absence de tête contextuelle, l'interprétation qui s'impose par défaut fait appel à la propriété *humain* (54) (cf. Corblin, 1995 ; Laborde 1999).

(51) tu prends une grande barre verticale et tu *en* mets *une deuxième* à côté (C12Maisons)

(52) tu prends une grande barre verticale et tu *en* mets *une deuxième*/ **un deuxième* à côté

(53) tu prends une grande barre verticale et tu *en* mets *une deuxième, de barre*, à côté

(54) *Les derniers* seront *les premiers.*

La particularité de l'interprétation référentielle de ces formes consiste en une séparation de deux opérations distinctes : d'une part, la récupération de la tête du syntagme et d'autre part le calcul du référent de la totalité du groupe. C'est cette disjonction interprétative en une opération d'anaphore nominale et une opération de recouvrement référentiel qui permet de prédire que dans certaines situations d'emploi les sources de ces deux opérations peuvent être disjointes :

(55) A propos de bateau, *celui-ci* est à lui. (Corblin, 1995)

En (55), l'anaphore nominale récupère sa source à travers le groupe nominal précédent, mais comme celui-ci n'est pas référentiel, la source de *celui-ci* en est forcément différente : il est probable qu'elle soit ici donnée par la situation d'énonciation. Bien que la disjonction des sources ne soit pas obligatoire, elle est fréquente et c'est encore la disjonction des sources, inattendue cette fois-ci, qui provoque l'effet comique en (56) :

(56) J'ai passé une excellente soirée, mais ce n'était pas *celle-ci.* (G. Marx, merci à Hélène...)

2.6.2 *Les références mentionnelles : une classe à part ?*

F. Corblin (1998, 1999) et M.-C. Laborde (1999) proposent de distinguer, parmi les groupes nominaux sans nom, une classe interprétative particulière : les références mentionnelles. Il s'agit d'emplois tels qu'en (57) et (58) :

(57) Mais l'existence de rapprochements possibles entre *ce* et *le* ne me semble pas constituer une raison suffisante d'inclure la signification *du second* dans celle *du premier.* (Gary-Prieur, 1998)

(58) Pierre connaît deux médecins. *L'un* est généraliste. *L'autre* est acupuncteur. (Laborde, 1999)

Cette catégorie est caractérisée de la façon suivante : « *Le phénomène anaphorique joue sur la syntaxe, sur le sens mais aussi sur la matérialité du texte. Il va utiliser l'ordre strict des éléments mentionnés du discours* ». Plus particulièrement, « *une expression comme* la première *a besoin pour être comprise, non seulement d'un antécédent mais d'une position particulière dans le discours.* » (Laborde, 1999 : 12). En plus de leur caractéristique principale – la saturation par le contexte textuel uniquement – les références mentionnelles se distinguent par le fait qu'elles n'acceptent pas la dislocation à droite en *de N*, n'admettent pas d'interprétation humaine par défaut et refusent la disjonction des sources référentielles.

Dans le cadre de notre modélisation (et plus particulièrement au chapitre 8), nous montrerons que les références mentionnelles, bien que possédant des propriétés particulières, ne mettent pas en cause les principes d'interprétation fondamentaux que nous formulerons pour la classe des groupes nominaux sans noms. Leur seule particularité consiste à faire appel à la structuration matérielle du contexte. Les caractéristiques énumérées ci-dessus sont les conséquences directes de cette particularité, mais n'affectent pas fondamentalement la manière de calculer le référent de ces expressions. D'ailleurs, F. Corblin considère lui-même que « *le fonctionnement mentionnel est donné comme une propriété naturelle des GN*[23] *sans nom dont la sémantique est positionnelle* » (1998 : 38).

2.6.3 *Particularités de* celui-ci

Au vu de l'analyse de *celui-ci* en tant que groupe nominal sans nom (Corblin, 1990), G. Kleiber (1994) s'interroge sur le coût interprétatif causé par cette proposition. Il conclut que celui-ci est trop élevé quand il n'y a pas disjonction de sources. Ses arguments reposent sur le constat qu'il existe un certain nombre d'emplois de *celui-ci* qui n'entrent pas dans le paradigme défini par F. Corblin :

(59) Paul se rendit chez le directeur. Celui-ci refusa de le recevoir. (Kleiber, 1994)

(60) Paul se rendit chez le directeur. ? Ce directeur refusa de le recevoir. (Kleiber, 1994)

(61) Paul se rendit chez le directeur. *Celui-ci, de directeur, refusa de le recevoir. (Kleiber, 1994)

En (59) par exemple, l'interprétation s'effectuerait de manière globale comme pour le pronom personnel, soit textuellement, soit situationnellement. Dans ce cas, la reprise nominale et la dislocation à droite en *de N* seraient impossible ((60) et (61)) et selon G. Kleiber, il n'y aurait pas disjonction des sources.

F. Corblin répond doublement à cette observation : en 1995, il propose de considérer qu'il ne s'agit pas d'une reprise, mais de l'interprétation *humain* accessible par défaut On n'emprunterait donc pas nécessairement au contexte un *N* spécifique, même si celui-ci est disponible. Le problème est alors qu'on ne sait plus très bien pas dans quels cas on doit recruter ou non un *N* du contexte.

[23] groupes nominaux

La deuxième proposition consiste à considérer *celui-ci* comme un mentionnel. Il désigne alors ce qui est « *associé à la mention la plus proche de l'occurrence de* celui-ci » (Corblin, 1998 : 37). Le fait qu'ici, contrairement aux démonstratifs (cf. 2.4.3), il n'y a pas opposition interne, mais opposition par rapport au référent d'une mention non contiguë explique en effet les observations de G. Kleiber et en particulier l'impossibilité de dislocation à droite. Par ailleurs, cette proposition fournit une explication pour l'emploi jugé non naturel, lorsque l'antécédent textuel se trouve tout seul :

(62) ? Le ballon a éclaté. Celui-ci était trop gonflé. (Kleiber, 1994)

Contrairement à ces différentes propositions de F. Corblin, le souci de G. Kleiber est plutôt de maintenir l'unité de l'analyse des différents emplois de *celui-ci*. Il propose alors le fonctionnement sémantico-référentiel suivant :

- la composante *lui*, en analogie avec le fonctionnement des pronoms personnels (2.5.3), renvoie à une classe de référents déjà saillante ;

- la composante démonstrative reflète la *token-reflexivité* typique aux démonstratifs (cf. 2.4.1). Elle appelle à un appariement avec un référent identifié par des éléments spatio-temporels. Cela entraîne une séparation du référent de la structure l'ayant rendu saillant et apporte donc du nouveau : la saillance du nouveau référent, une reclassification, une rupture du cadre discursif, un changement de thème, une visée contrastive etc.

Cette approche a l'avantage de couvrir à la fois tous les emplois observés par F. Corblin et d'être prédictive par rapport aux observations apportées par G. Kleiber. Par rapport aux analyses de F. Corblin, la classe saillante peut être rapprochée de l'ensemble des *N* pour ce qui est de l'emploi en anaphore nominale et de l'ensemble des mentions récentes constituant l'ensemble des oppositions pour l'emploi mentionnel. La composante démonstrative couvre l'introduction d'un nouveau référent pour l'anaphorique nominal. En ce qui concerne le mentionnel, elle est tout à fait compatible avec des observations supplémentaires apportées par F. Corblin (1998) : celui-ci constate que *celui-ci* mentionnel est employé rarement, malgré son caractère remarquablement non ambigu. L'explication serait qu'il sert à gérer, comme le démonstratif (2.4.3), des ruptures discursives, se traduisant précisément par l'introduction de chaînes topiques secondaires, plus difficiles à gérer.

2.7 Synthèse

La description séparée du fonctionnement référentiel des expressions indéfinies, définies, démonstratives et des pronominales – sous leur forme pleine ou discursivement réduite – est une étape nécessaire à la mise en évidence des particularités interprétatives propres à chacune de ces classes. En revanche, elle ne doit pas faire oublier que ces expressions constituent un ensemble : l'ensemble des formes intervenant dans des actes référentiels. De ce fait, elles sont étroitement liées à l'organisation des entités contextuelles auxquelles il est possible de faire référence. Ce lien est double : d'une part, **l'usage d'une expression référentielle est conditionné par l'organisation contextuelle**. D'autre part, l'isolement d'un référent par l'emploi d'une expression référentielle a des effets interprétatifs qui entraînent une **réorganisation de l'ensemble contextuel**. Ce sont là des caractéristiques communes à toutes les expressions étudiées. A l'intérieur de ce schéma général, les différents types d'expression se distinguent par la nature des conditions d'organisation contextuelle et par la nature des effets interprétatifs.

A partir des observations des sections précédentes, nous proposons ici une synthèse de l'interprétation des expressions référentielles en français sous forme d'un schéma (Figure 4). Cette synthèse associe à chaque expression référentielle (partie centrale du schéma) les conditions

contextuelles nécessaires pour que l'expression puisse être interprétée (partie gauche) et les effets interprétatifs obtenus (partie droite).

En partant des conditions contextuelles, le schéma fait apparaître la séparation des formes qui supposent un référent déjà saillant (pronom et démonstratif) de celles opérant à l'intérieur d'un ensemble d'interprétation (indéfini et défini). A l'intérieur du premier groupe, la saillance peut être le résultat d'une mention récente ou d'un geste de désignation. Le critère de différenciation entre pronom et démonstratif est essentiellement interprétatif. L'emploi d'un pronom exprime une continuité, alors que l'emploi d'un démonstratif traduit une rupture. Cette rupture peut être un changement du point de vue ontologique (reclassification pour les formes pourvues d'une référence virtuelle), du point de vue discursif (changement du thème, introduction d'un contre-topique) ou l'introduction d'un nouveau référent. Dans tous les cas, le démonstratif a pour effet d'instaurer une opposition (quelquefois virtuelle) entre le référent extrait et les autres éléments de la classe des *N* (*ce N*), des éléments de type *N* contigus (*celui-ci*) ou des mentions antérieures (*celui-ci* mentionnel).

A l'intérieur du second groupe, défini et indéfini se distinguent par la composition de leurs ensembles d'interprétation respectifs. Le défini demande, pour s'interpréter, un domaine hétérogène : soit des éléments de types différents (*le N*), soit des éléments d'un même type, mais se différenciant selon une propriété particulière (*le P*), soit les deux à la fois (*le N P*). Les propriétés différenciantes ne sont pas uniquement des propriétés physiques comme la couleur ou la taille. Il peut s'agir par exemple de l'appartenance ou de propriétés relationnelles comme l'ordonnancement. L'ordonnancement lui-même peut être donné par la perception, par un ordre événementiel ou par la matérialité du discours. Dans tous les cas, une expression définie a pour effet d'isoler son référent à l'intérieur du domaine d'interprétation et ceci par une opération d'extraction en vertu de sa référence virtuelle. Elle opposera alors le référent extrait aux autres éléments du domaine.

L'indéfini *un N* suppose une classe d'éléments de type *N*. L'homogénéité de la classe interprétative est une conséquence du fonctionnement particulier de l'indéfini : il s'agit d'une opération de dénombrement d'éléments de type *N*, ce dénombrement étant effectué indépendamment d'autres critères d'organisation contextuelle. La saillance d'un tel ensemble conditionnera l'utilisation de la forme pleine (*un N*) ou réduite (*en... un*). L'effet interprétatif est dans tous les cas une extraction aléatoire (indépendante du contexte) d'un élément de la classe, celui-ci étant opposés aux autres éléments en vertu de la prédication de la proposition englobante.

Cet aperçu global du système référentiel, inspiré de travaux linguistiques, a l'avantage de faire apparaître la grande variété des facteurs dont il faudra tenir compte lors de la modélisation du calcul référentiel. Une telle modélisation, si elle doit rendre compte de toute la richesse des différents marqueurs référentiels, devra alors intégrer :

- la gestion d'ensembles contextuels avec des informations sur le type et les propriétés des éléments réunis ;

- la gestion de la variété de propriétés différenciantes : propriétés physiques, relations d'appartenance, ordonnancement ;

- la possibilité de calculer ces propriétés en provenance de sources diverses : connaissances conceptuelles, données situationnelles, organisation de la matérialité du discours ;

- la gestion de la notion de saillance en relation avec la récence d'une mention ou la désignation gestuelle d'un référent ;

- la restructuration du contexte en fonction des effets interprétatifs : gestion des extractions, des oppositions résultantes, gestion de la continuité ou des ruptures discursives.

Dans les sections suivantes, nous examinerons un certain nombre de modélisations proposées, dans l'objectif des les comparer en fonction de ces caractéristiques.

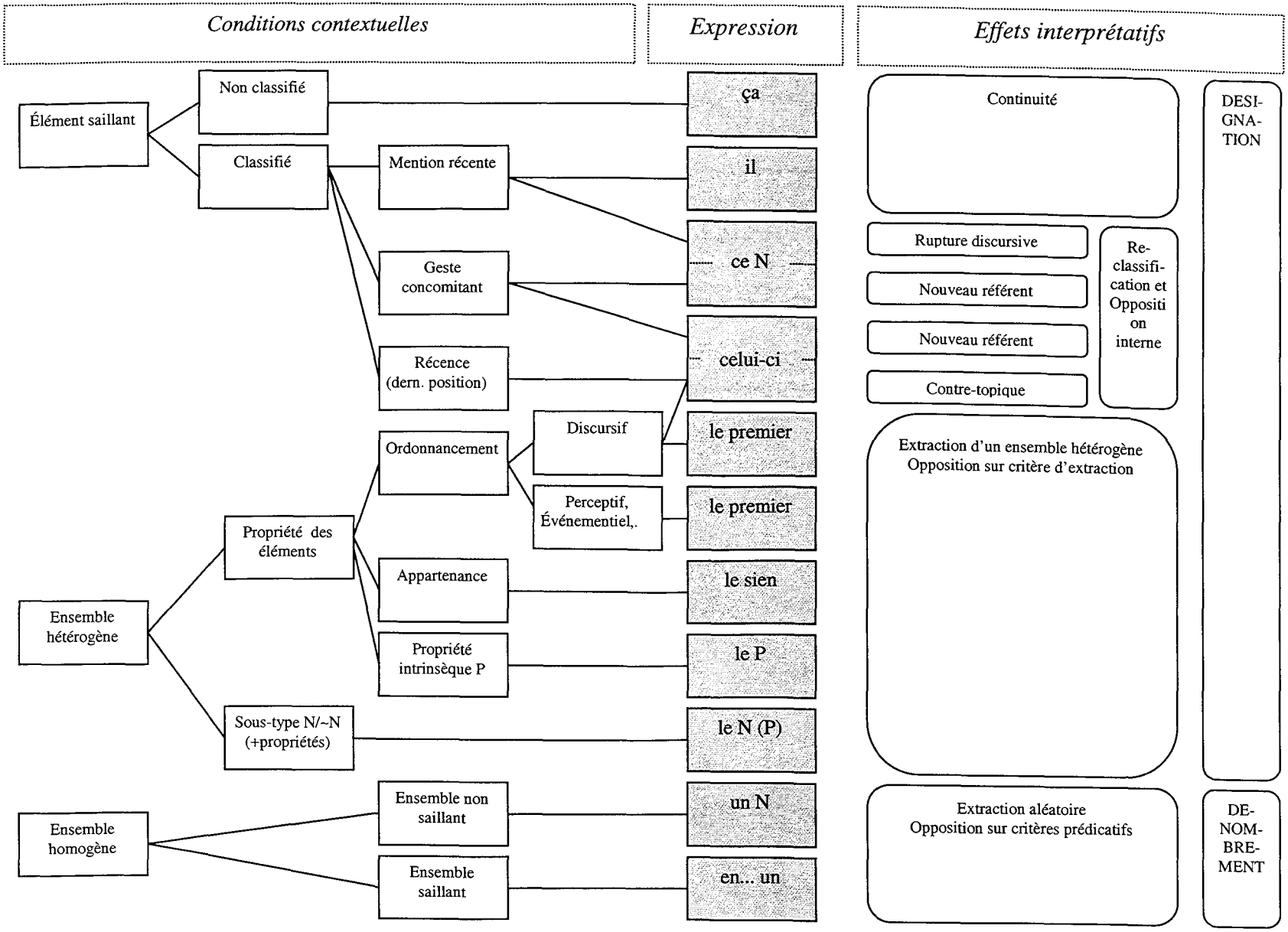


Figure 4 : Synthèse des conditions contextuelles et effets interprétatifs associés aux différents marqueurs référentiels

3 Les approches cognitives : l'optimisation du calcul référentiel

En complément des études linguistiques sur la spécificité de chaque marqueur dans le système référentiel d'une langue particulière, il existe des approches qui essayent d'expliquer l'usage des expressions référentielles par des principes plus généraux, liés au fonctionnement de la mémoire humaine. C'est pour cette raison que nous proposons de les regrouper ici sous le terme de « cognitif ».

Le point commun de ces approches est de partir de la capacité limitée de la mémoire de travail. Cette limitation concerne à la fois la capacité de stockage (Miller, 1956) et la capacité de calcul. L'objectif premier du processus d'interprétation d'une expression référentielle étant la (re-)identification de la représentation d'un référent en mémoire, il s'agit de minimiser le coût de calcul de cette opération. Ce coût dépend de l'emplacement du référent dans la mémoire et donc de sa « familiarité » (Prince, 1981), de son « accessibilité » (Ariel, 1988) ou encore de son « statut cognitif » (Gundel et al., 1993). L'hypothèse fondamentale de ces travaux est alors la suivante : l'identification du référent dans la mémoire est facilité par le type de l'expression (indéfini, défini, pronom, ...) qui fournit une indication sur l'emplacement du référent dans la mémoire, ce qui permet de restreindre l'espace de recherche. L'adéquation entre cette indication et l'emplacement réel du référent dans la mémoire se traduirait alors par une optimisation du coût de traitement.

Par la suite, nous proposons d'examiner cette approche à travers certains travaux marquants : nous avons retenu ceux de E. Prince (1981) pour leur aspect précurseur, ceux de M. Ariel (1988) pour la tentative d'intégration d'une approche cognitive de la référence à la théorie de la pertinence (Sperber et Wilson, 1989) et ceux de J. Gundel et al. (1993) pour la validation des hypothèses sur des données réelles et multilingues. Nous nous arrêterons également sur reformulations des hypothèses en termes de critères opératoires, ayant permis de confirmer certains points par des expérimentations psycholinguistiques (M. Fossard et al., 1999 ; A. Almor, 1999).

3.1 De la « familiarité » au « statut cognitif »

Une des caractéristiques des langues naturelles est de pouvoir référer à un même objet par différentes expressions. Mais toutes les formes ne sont pas utilisables dans les mêmes circonstances. A partir de cette observation s'est posée la question sur la relation entre le statut des représentations associées aux entités désignées et les formes linguistiques employées.

3.1.1 L'échelle de familiarité

E. Prince (1981) constate qu'une simple distinction entre « introduction d'une représentation nouvelle » et « identification d'une représentation faisant déjà partie de l'univers partagé » est insuffisante. Elle propose donc une taxonomie plus élaborée des statuts cognitifs possibles, sans les relier pour autant de façon stricte à des formes linguistiques. Les statuts sont définis en termes de « familiarité » ce qui donne lieu à une « échelle de familiarité ». Cette échelle distingue les entités nouvelles des entités inférables et des entités évoquées ou connues, chacune de ces classes étant décomposée elle-même en sous-classes, détaillées dans la Figure 5.

En fonction du statut cognitif d'une entité désignée, les opérations à effectuer pour interpréter une expression référant à cette entité seraient plus ou moins complexes : création d'une nouvelle représentation dans le modèle du discours, déplacement d'une représentation de la mémoire à long terme vers le modèle du discours, identification d'une représentation dans le modèle du discours,

identification d'une entité saillante dans le modèle du discours. Le Tableau 6 récapitule les degrés de familiarité et les opérations associées et donne un exemple pour chaque cas.

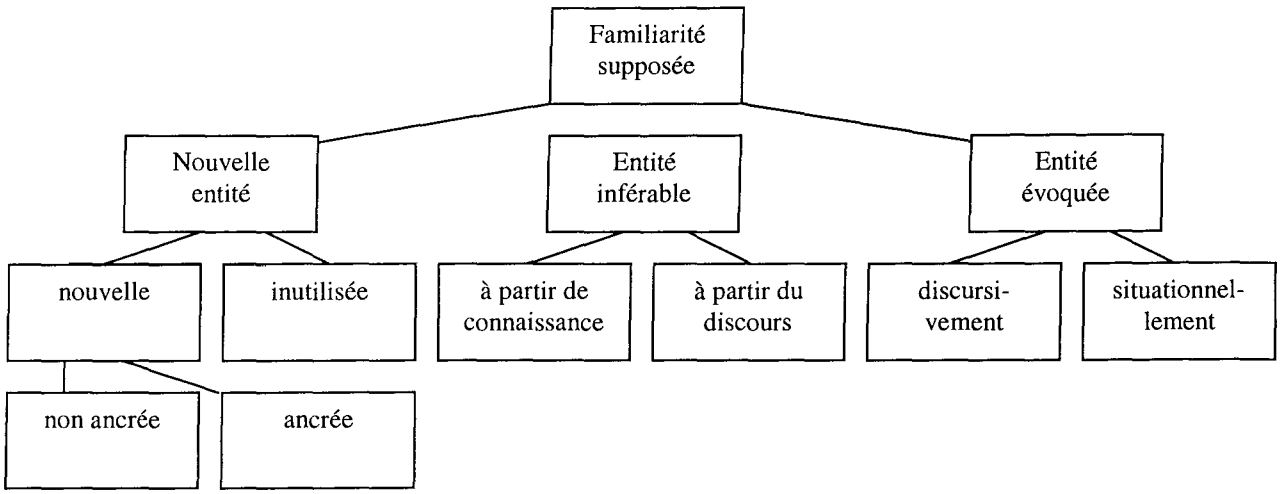


Figure 5 : Échelle de familiarité, d'après E. Prince (1981)

Statut cognitif du référent	***Opération d'identification***	***Exemple***
non ancré	création d'une nouvelle représentation	*J'ai acheté une robe.*
ancré	déplacement d'une représentation de la mémoire à long terme	*Chomsky est connu.*
inférable à partir de connaissances conceptuelles	création d'une nouvelle représentation à partir d'une représentation existante	*J'ai heurté un bus. Le conducteur dormait.*
inférable à partir de connaissances discursives	création d'une représentation à partir d'une représentation nouvellement créée	*Un des œufs n'est pas frais.*
évoqué discursivement	identification d'une représentation dans le modèle du discours	*Chaperon Rouge allait voir sa grand-mère. La brave dame était en train de faire sa sieste.*
évoqué situationnellement	identification d'une représentation pour une entité du contexte	*Je n'ai pas d'autres idées.*

Tableau 6 : Statuts cognitifs et opérations associées, d'après E. Prince (1981)

Les propositions de E. Prince (1981) ne prédisent pas encore directement l'apparition d'une expression en fonction du statut cognitif du référent, mais elles ont l'avantage de souligner la relation entre ce que le locuteur pense être connu par l'auditeur et la réalistion linguistique des syntagmes nominaux, tout en dépassant la vision simpliste d'une bipartition entre « information nouvelle » et « information connue ». L'introduction de cette échelle de statuts cognitifs a donné suite à d'autres travaux qui s'en sont inspirés pour prédire plus directement l'apparition d'une expression en fonction de son « accessibilité » dans la mémoire.

3.1.2 *L'échelle d'accessibilité (Ariel, 1988)*

Parmi ces travaux, ceux de M. Ariel (1988) sont connus pour s'inspirer d'une théorie cognitive plus globale, la théorie de la pertinence (Sperber et Wilson, 1989). L'hypothèse de M. Ariel est que les processus de compréhension respectent un équilibre entre le coût du calcul et l'effet interprétatif. Appliqué au calcul référentiel, l'effet interprétatif consiste en l'identification du « bon » référent, alors que le coût du calcul est lié à la quantité de l'information transmise (et donc à traiter) par l'expression référentielle. Cette hypothèse l'amène à concevoir une « échelle d'accessibilité » qui suggère une optimisation de l'emploi des expressions référentielles : plus une entité est facile à identifier, c'est-à-dire « accessible » en mémoire, moins la quantité d'information transmise par l'expression sera importante. Ceci réduit le coût du traitement et optimise ainsi le processus du calcul référentiel.

Comme le souligne A. Reboul (Reboul et al., 1997), certains aspects de cette hypothèse sont extrêmement séduisants. Tout comme E. Prince (1981) et J. Gundel et al. (1993), M. Ariel vise une théorie cognitive qui considère que les référents sont modélisés par des représentations mentales (et non pas par des antécédents discursifs) et que le choix des expressions référentielles devrait se faire en fonction des croyances attribuées à l'interlocuteur. Intégrer une telle théorie du calcul référentiel dans une théorie plus globale de la pertinence semble plausible et intéressant, d'autant plus que cela devrait conduire à une prise en compte de données hétérogènes, indispensable à la modélisation du calcul référentiel. Par ailleurs, la structuration du contexte autour de la notion-clé d'*accessibilité* devrait amener à réfléchir sur des concepts proches comme *saillance*, *récence*, *focus* ou *point de vue*.

Au vu de ces ambitions, l'application pratique que propose M. Ariel de sa théorie réduit les relations d'accessibilité, que les expressions sont supposées véhiculer, au simple fait de récence textuelle. L'échelle qu'elle propose est en fait le résultat de calculs statistiques sur la distance textuelle entre une expression référentielle et son antécédent, localisé soit dans la même phrase ou la phrase immédiatement précédente, soit dans le même paragraphe, soit dans le même texte. De ce fait, l'échelle se limite à une contrainte de distance textuelle entre expressions coréférentes : les prédictions effectives se réduisent à ce qu'un pronom – indicateur d'une accessibilité forte – trouve souvent son antécédent dans la même phrase ou la phrase précédente, alors qu'un nom propre – indicateur d'une accessibilité faible – le trouve généralement à l'intérieur du même texte (Reboul et al., 1997).

Malgré les critiques sévères adressées par A. Reboul à cette théorie (Reboul et al., 1997 ; Moeschler et Reboul, 1998), nous en retiendrons l'hypothèse que la forme linguistique et plus particulièrement le type d'une expression référentielle est lié au statut cognitif du référent visé. Une facette de ce statut est son accessibilité dans la mémoire, sachant que la récence textuelle n'est probablement qu'une indication parmi d'autres pour mesurer cette accessibilité. Enfin, nous pensons que la mise en œuvre de cette théorie, jugée décevante par certains, traduit une difficulté réelle d'opérationalisation des concepts cognitifs sous-jacents : toutes les approches cognitives présupposent effectivement qu'il est possible de calculer l'activation d'une représentation indépendamment de l'expression utilisée pour y référer. Or, cela peut s'avérer extrêmement difficile, comme l'illustrent également les travaux de J. Gundel et al. (1993).

3.1.3 *Statuts cognitifs et expressions référentielles (J. Gundel et al., 1993)*

Toujours dans l'objectif de fournir une explication d'ordre cognitif à la distribution des différents types d'expressions référentielles, les travaux de J. Gundel et et al. (1993), connu sous le terme de « *Givenness Hierarchy* », partent d'une hypothèse proche de celle de M. Ariel : « *Different determiners and pronominal forms conventionally signal different cognitive statuses (information about location in memory and attention state), thereby enabling the addressee to restrict the set of*

possible referents »[24] (1993 : 274). La hiérarchie ainsi que les exemples des auteurs sont réunis dans le Tableau 7 :

Niveau	***Statut cognitif du référent***	***Conditions mémorielles***	***Forme linguistique associée et exemple***
1	focalisé	accès à la représentation d'une entité susceptible de servir de topique	*My neighbor has a dog. It kept me awake.*
2	activé	accès à la représentation d'une entité présente dans la mémoire à court terme	*My neighbor has a dog. This dog kept me awake.*
3	familier	accès à la représentation d'une entité présente dans la mémoire à long terme	*I couldn't sleep last night. That dog kept me awake.*
4	identifiable de façon unique	accès à la représentation d'une entité existante ou pouvant être construite sur la base du contenu nominal de l'expression	*I couldn't sleep last night. The dog kept me awake.*
5	référentiel	accès à la représentation ou construction d'une représentation sur la base de l'expression et du contenu prédicatif	*I couldn't sleep last night. This dog (next door) kept me awake.*
6	possibilité d'identifier un type	accès à la représentation du type de l'objet désigné par le contenu nominal	*I couldn't sleep last night. A dog kept me awake.*

Tableau 7 : « Givenness Hierarchy », d'après Gundel et al., 1993

Contrairement aux statuts mutuellement exclusifs que proposent M. Ariel et E. Prince, les statuts cognitifs sont ici reliés par une relation d'implication. Chaque statut implique, par définition, tous les statuts inférieurs : un référent focalisé, par exemple, est obligatoirement activé, familier et ainsi de suite.

Selon les auteurs, le statut cognitif d'une entité est une condition nécessaire et suffisante pour l'emploi de la forme linguistique correspondante. Sur la base de cette hiérarchie, il est alors possible de faire deux prédictions : premièrement, une forme particulière est inappropriée si le référent ne possède pas le statut cognitif nécessaire. Deuxièmement, il est possible d'utiliser une forme linguistique correspondant à un statut inférieur que le statut effectif du référent. Un des mérites de ce travail est d'avoir testé ces deux prédictions sur des extraits de dialogues de cinq langues différentes : le chinois, l'anglais, le japonais, le russe et l'espagnol.

Concernant la première prédiction, les auteurs ont observé que la très grande majorité des expressions réfèrent à des entités possédant effectivement au moins le statut cognitif nécessaire. Ils remarquent néanmoins que si cela n'est pas le cas, deux solutions se présentent : soit le référent n'est pas identifié par l'interlocuteur, soit il est accommodé, comme dans l'exemple (63), où *they* réfère au couple que forment Barb et son compagnon :

(63) A : Barb has it. I suspicion she was a cat in some previous life.
Oh, did I tell you that *they* have a cat ? (Gundel et al., 1993)

Or, cette possibilité d'accommodation est délicate à manipuler et fait apparaître un certain nombre de questions. D'abord, ne doit être accommodé que ce qui n'est pas déjà focalisé. Comme les auteurs ne donnent pas d'indications claires sur la façon de calculer le focus (ni les autres statuts cognitifs), on

[24] « Différents déterminants et formes pronominales signalent de façon conventionnelle différents statuts cognitifs (informations sur la localisation dans la mémoire et sur l'état d'attention) et permettent ainsi de réduire l'ensemble des référents potentiels. »

ne sait pas pourquoi le référent de *they* n'est pas déjà focalisé. Ensuite, en admettant que cela ne soit pas le cas, il serait nécessaire de spécifier plus avant les conditions qui permettent une accommodation de ce type, faute de quoi tout pourrait être accommodé et le rôle des statuts cognitifs dans la détermination de la forme linguistique serait caduque.

Concernant la deuxième prédiction – selon laquelle il est possible d'utiliser une forme linguistique correspondant à un statut inférieur que le statut effectif du référent – les deux observations retenues sont les suivantes :

Cette prédiction ne semble pas se confirmer pour l'utilisation des démonstratifs (rarement utilisés à la place d'un pronom pour désigner des entités focalisées), ni pour celle des indéfinis (jamais utilisés à la place d'un défini pour des référents étant identifiables de façon unique). Pour expliquer cette observation, les auteurs invoquent la maxime de quantité de P. Grice (1979) selon laquelle une communication réussie exige que l'on dise ce qui est nécessaire, mais sans dire plus que ce qui est nécessaire. Étant donné que les auteurs considèrent qu'un indéfini donne moins d'informations qu'un défini sur la localisation de son référent, dire autant que nécessaire signifie ici qu'utiliser un indéfini à la place d'un défini n'est pas une attitude coopérative. Le même type de raisonnement s'applique au démonstratif et expliquerait qu'il n'est que rarement utilisé pour référer à une entité focalisée.

Contrairement à ce que prédit l'échelle des statuts cognitifs du Tableau 7, les démonstratifs ne sont employés que rarement pour référer à une entité familière. En revanche, ce rôle est joué dans 85 % des cas par les définis. Ici aussi, les auteurs trouvent une explication sur la base de la maxime de quantité de P. Grice. En partant du constat qu'une entité est le plus souvent identifiable de façon unique sur la base de sa familiarité, ils estiment qu'utiliser un démonstratif contredit le principe selon lequel il ne faut pas donner trop d'information. Et comme le démonstratif se trouve ainsi « dépourvu » de conditions d'emploi propres sur la base d'un statut cognitif associé, son emploi se justifierait finalement par l'ajout de nouvelles informations sur les référents.

Nous pensons qu'ici aussi, tout comme pour la première prédiction, les choses sont délicates car les résultats d'observation reposent sur un classement manuel non élucidé des statuts cognitifs. Mais ce qui est plus inquiétant encore, c'est qu'une partie non négligeable des prévisions sur la possibilité de « remplacement » par des formes linguistiques correspondant à un statut cognitif inférieur ne semble pas correspondre aux observations. De ce fait, il est légitime de s'interroger sur la plausibilité de la relation d'implication présumée entre les statuts cognitifs qui est à l'origine de cette vision « échangiste » des marqueurs référentiels. Il est par ailleurs à noter que cette vision est à l'opposé des travaux linguistiques présentés dans le chapitre précédent, essayant précisément de souligner le caractère propre et irréductible de chaque marqueur référentiel. Le fait que les observations ne se justifient plus par rapport aux statuts cognitifs présumés, mais par rapport à des principes extérieurs laisse plutôt penser que l'échelle des statuts cognitifs n'arrive pas, à elle seule, à rendre compte de l'usage des expressions référentielles.

3.2 Données expérimentales

Les développements théoriques présentés ci-dessus ont incité certains chercheurs à étudier de façon empirique les effets du statut cognitif sur le coût du traitement de différents marqueurs référentiels. L'apport de ces études expérimentales est double : d'une part, elles obligent à expliciter les concepts manipulés de façon à les transformer en variables expérimentales indépendantes (statut cognitif) et dépendantes (coût du calcul et effets interprétatifs) qui soient opératoires et calculables. D'autre part, elles fournissent des données qui contribuent à confirmer ou à infirmer les hypothèses avancées. En

revanche, les inconvénients des ces études sont évidemment liés à l'opérationalisation, qui risque d'entraîner une certaine simplification : en particulier, les expériences rapportées ci-dessous abandonnent la notion de représentation mentale du référent au profit de la notion d'antécédent textuel, beaucoup plus facile à manipuler.

3.2.1 Pronom contre groupe nominal : 1 – 0

Certaines études sont consacrées à l'influence de la focalisation d'un référent sur l'emploi de pronoms et d'anaphores nominales (Gordon et al. (1993) pour l'anglais, Fossard et al. (1999) pour le français). Les deux hypothèses testées par M. Fossard (1999) s'inspirent directement des travaux de M. Ariel (1988) et de J. Gundel et al. (1993). Une première hypothèse prédit que, étant donné le statut cognitif requis pour l'emploi d'un pronom – la focalisation – une expression anaphorique pronominale devrait demander un coût de traitement plus élevé lorsqu'elle réfère à une entité non focalisée. A l'inverse, une deuxième hypothèse prédit que, étant donné le statut cognitif requis pour l'emploi d'une expression nominale – la non-focalisation – une expression anaphorique nominale (matérialisée ici par nom propre répété) devrait demander un coût de traitement plus élevé lorsqu'elle réfère à une entité focalisée.

Dans les expériences menées par M. Fossard, l'opérationalisation des variables s'est faite de la façon suivante : la variable indépendante, c'est-à-dire la focalisation d'un référent, est liée à une variation des fonctions syntaxiques (sujet vs. objet oblique) de l'antécédent concerné. La variable dépendante, c'est-à-dire le coût du traitement, a été mesurée par le temps de lecture de textes. Ces textes se présentent comme dans l'exemple (64), reproduit ici d'après M. Fossard (1999) :

(64)	a.	Les patrons de l'usine faisaient une grande réunion.	(phrase d'introduction)
	b.	*Alice* exigeait le rapport de Patrice pour donner un avis.	(introduction de l'antécédent)
	c.	Pendant une heure, *elle* parla sans arrêt.	(reprise pronominale/nominale)
	d.	La femme est bavarde.	(question de compréhension)

Les résultats confirment la première hypothèse : dans tous les cas, la focalisation de l'antécédent facilite la lecture, mais l'utilisation d'un pronom pour référer à l'élément focalisé réduit encore le temps de lecture par rapport à l'utilisation du nom propre répété et coûte donc moins cher. Ce fait a déjà été observé par ailleurs. Il est connu sous le terme de « pénalité de nom répété » (Gordon, 1993). Cette « pénalité » viendrait du fait que l'on prive le locuteur d'une information importante, à savoir que le référent est focalisé.

La deuxième hypothèse, selon laquelle un nom répété devrait demander plus de temps de traitement lorsqu'il réfère à un élément focalisé, n'est pas confirmée. Il n'y a pas de différence significative entre la référence, par un nom propre répété, à une entité focalisée ou non. Cela amène M. Fossard à penser que les noms propres et, plus largement les expressions nominales, identifient leur référent sur d'autres critères que la focalisation. Cette hypothèse est entièrement compatible avec les propositions issues de la linguistique descriptive (cf. notre chapitre 2), soucieuses d'attribuer à chaque marqueur référentiel son fonctionnement propre. En revanche, elle va plutôt à l'encontre d'une détermination de la distribution des marqueurs référentiels sur l'unique critère d'accessibilité ou de focalisation.

3.2.2 Où il sera rendu justice au groupe nominal

Les travaux comme ceux que nous venons de présenter ci-dessus ont tendance à conclure à une supériorité systématique du pronom sur le groupe nominal lors de la référence à des entités focalisées. Mais, comme le remarque A. Almor (1999), un test qui repose uniquement sur des anaphores

répétitives sous forme de nom propre ne permet pas de tirer des conclusions sur le comportement général des anaphores nominales. Premièrement, les noms propres sont susceptibles d'avoir un comportement spécifique par rapport aux descriptions définies. Deuxièmement, travailler uniquement sur des reprises fidèles risque de faire oublier qu'une anaphore peut aussi apporter des informations nouvelles, ce qui justifie éventuellement un coût de traitement plus élevé.

L'objectif de A. Almor (1999) est donc de confirmer de façon expérimentale une hypothèse qui soit valide pour toutes les anaphores nominales (y compris le nom répété). Son hypothèse est proche de celles de M. Ariel (1990) et de J. Gundel et al. (1993) et préconise, elle aussi, l'emploi de la forme linguistique la moins complexe pour servir le but communicatif. Plus précisément, elle associe le coût du traitement d'une anaphore à sa charge informationnelle (nous traduisons ainsi le terme « *informational load* »). La charge informationnelle résulte de la distance sémantique entre l'antécédent et l'anaphore, sachant que la généralisation et la spécification se traduisent respectivement par une charge informationnelle négative et positive, comme le montre la Figure 6. Selon l'auteur, une charge informationnelle élevée (en pratique zéro ou positive) n'est justifiée que par l'identification d'un antécédent jusqu'alors non focalisé ou par l'ajout d'une information nouvelle sur un référent focalisé.

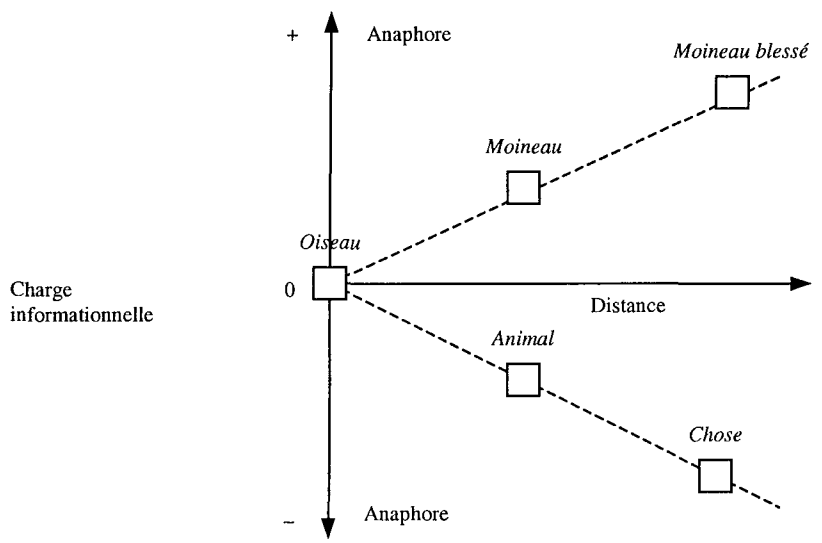


Figure 6 : Charge informationnelle et distance sémantique entre antécédent (oiseau) *et anaphore, d'après Almor (1999)*

Différentes séries expérimentales confirment cette hypothèse : d'abord, les anaphores à charge informationnelle haute mais sans apport d'informations (les répétitions pures) sont lues plus rapidement lorsque l'antécédent n'est pas focalisé que lorsqu'il est focalisé. Cette observation – qui est en contradiction avec ce qu'observe M. Fossard (1999) – rejoint la « pénalité du nom répété » et s'explique par le fait que la charge informationnelle du nom répété est élevée, mais non justifiée : elle ne sert ni à ajouter de l'information (il s'agit d'une répétition pure), ni à détourner l'attention sur une entité jusqu'alors non focalisée. Mais si cette « pénalité du nom répété » venait uniquement, comme le pense P. C. Gordon (1993), du fait de priver le lecteur de l'indice focal apporté par le pronom, elle devrait alors être valable pour toutes les anaphores nominales. Or, A. Almor montre par d'autres séries expérimentales que ce n'est pas le cas : si le nom répétitif est effectivement pénalisé, ce n'est pas le cas pour des reprises non répétitives.

En effet, les anaphores à charge informationnelle basse (généralisations) ou à charge informationnelle haute avec apport de nouvelles informations (spécifications) sont lues plus rapidement si l'antécédent est focalisé. Concernant les généralisations, cette observation se justifie par

le fait qu'une charge informationnelle basse implique un coût de traitement faible, optimal pour des anaphores n'apportant pas d'informations nouvelles. A. Almor observe d'ailleurs que plus une anaphore est générale, plus elle est traitée rapidement. Ce résultat est compatible avec les travaux accordant la préférence aux pronoms (cas extrême de généralisation) face aux répétitions nominales pour référer à l'entité focalisé. Il confirme également des résultats obtenus sur le français : M. Ferréol et M. Fayol (1994) ont observé qu'une reprise par hyperonyme était traitée plus facilement qu'une reprise par nom répété. En ce qui concerne les spécifications, l'explication est la suivante : une spécification implique une charge informationnelle haute, mais celle-ci est justifiée par un apport d'informations supplémentaires.

La conclusion principale à tirer des travaux de A. Almor est la relativisation du rôle prépondérant accordé traditionnellement au pronom. Or, un pronom n'est pas intrinsèquement « meilleur » qu'un groupe nominal pour la reprise d'un référent focalisé. Tout dépend de l'équilibre entre le coût du traitement et l'effet interprétatif : s'il s'agit d'une reprise pure et simple, une description définie ne se justifie pas. En revanche, une anaphore nominale peut se justifier soit par un apport d'informations supplémentaires, soit par un coût de traitement faible lors d'une reprise par généralisation.

3.3 Synthèse

L'apport essentiel des approches cognitives concerne la structuration du contexte qui est ici inspirée du fonctionnement cognitif et plus particulièrement du fonctionnement de la mémoire de travail. Le contexte est supposé contenir un nombre limité de représentations mentales pour des entités du monde, ordonnées selon leur activation. Cette conception a deux avantages : d'une part, elle permet d'asseoir certaines notions dégagées par l'analyse linguistique (saillance, récence) sur des facteurs cognitifs. D'autre part, puisque ces facteurs sont supposés relever du fonctionnement de la mémoire humaine, ils peuvent constituer un point de départ pour des études multilingues : les interrogations concernent alors la réalisation linguistique des différents statuts cognitifs à travers différentes langues et peuvent aboutir à la question de l'universalité des marqueurs référentiels, ou, au moins, de certaines oppositions.

En revanche, une telle investigation demanderait d'abord de résoudre ce que nous considérons comme le problème majeur des approches cognitives : le calcul des différents statuts cognitifs. Alors que tous les développements théoriques distinguent au moins six statuts différents, nous n'avons pu trouver aucune indication permettant de déterminer avec un minimum de cohérence le statut cognitif d'une représentation mentale pour une entité donnée. Plusieurs indices confirment d'ailleurs cette impression de flou autour de ce concept. Du côté théorique, J. Gundel et al. (1993) font, par exemple, une distinction entre les statuts « identifiable de façon unique » et « familier », mais cette distinction ne s'avère pas pertinente dans la suite de la théorie. D'où vient alors cette distinction et qu'est-ce qui justifie son maintien ? D'un point de vue pratique, M. Ariel n'a pas pu trouver d'autres solutions que la distance textuelle entre antécédent et anaphore pour mesurer l'accessibilité d'une représentation. Enfin, en ce qui concerne les validations expérimentales, les statuts examinés se limitent à une opposition entre focalisation et non-focalisation. Toujours est-il que ces expériences ont mis en lumière quelques facteurs importants pour le calcul focal : la récence de la mention (matérialisée dans le discours par la distance), les fonctions syntaxiques et certaines constructions comme les clivées.

En admettant que ce premier problème puisse trouver une solution, une deuxième interrogation porte sur le caractère (in)suffisant des statuts cognitifs pour la prédiction de la distribution des marqueurs référentiels. Un examen des théories proposées fait apparaître qu'en fin de compte, d'autres facteurs que le seul statut cognitif jouent un rôle important. Un premier point non résolu est le statut imprécis attribué à des processus d'accommodation qui semblent se superposer aux statuts cognitifs.

Ensuite, J. Gundel et al. (1993) remarquent que la distribution des expressions référentielles, en particulier celle des démonstratifs, ne correspond pas toujours aux prédictions. La solution passe alors par une redéfinition de la fonction du démonstratif en termes d'ajout d'informations, c'est-à-dire en termes d'effets interprétatifs. Cette tendance apparaît encore plus clairement lors de l'interprétation des résultats expérimentaux : M. Fossard (1999) explique la distribution entre pronom et anaphore nominale en partie par des différences d'effets discursifs (continuité *versus* rupture) et A. Almor (1999) prend explicitement en compte une fonction discursive d'apport informationnel qui explique la distribution des différents types d'anaphore nominale. Il semble donc le statut cognitif seul n'explique pas tout et que le recours aux effets interprétatifs spécifiques de chaque marqueur s'impose *a posteriori*. Or, c'est précisément l'idée largement défendue par les travaux linguistiques (présentés dans le chapitre précédent) qui essaient de prédire la distribution des expressions référentielles en fonction de la structuration du contexte *et* en fonction des effets interprétatifs.

Enfin, il n'est pas tout à fait clair ce qu'est un statut cognitif. S'agit-il juste d'un taux d'activation, comme le pourrait laisser croire M. Ariel, ou s'agit-il d'une structuration plus complexe ? C'est probablement cette deuxième hypothèse qui est sous-jacente aux propositions de J. Gundel et al. (1993), car la possibilité de calculer l'unicité d'une représentation nécessite la représentation d'un ensemble contextuel à l'intérieur duquel le référent est à isoler. Or, cet aspect, laissé implicite, mériterait un développement pour au moins deux raisons : d'une part, cela pourrait constituer le début d'une prise en compte de certaines expressions elliptiques qui se justifierait ainsi par rapport à l'activation d'un ensemble contextuel. D'autre part, cela pourrait être une ouverture vers la prise en compte d'ensembles contextuels fournis non seulement par l'univers discursif, mais aussi par l'environnement perceptif, ou plus largement situationnel.

4 Modélisations du contexte

4.1 Introduction

Nous nous intéresserons ici à des approches plus globales des processus de compréhension discursive. Ces modélisations mettent donc moins l'accent sur la spécificité de tel ou tel marqueur référentiel (chapitre 2) ou l'accessibilité (chapitre 0). En revanche, elles sont complémentaires des travaux des chapitres précédents, puisqu'elles situent notre thématique – la résolution référentielle – dans le contexte plus large de l'interprétation d'un énoncé. Toutes les approches présentées ci-dessous reposent sur une hypothèse commune et relativement générale : l'interprétation d'un énoncé est contrainte par le contexte d'énonciation. Parmi les informations contextuelles, il y a bien sûr des informations linguistiques. Cependant, d'autres informations – perceptives ou encyclopédiques, par exemple – contribuent également à la compréhension d'un énoncé. L'objectif commun aux modélisations contextuelles présentées est alors d'étudier la nature de ces informations et la façon dont elles se combinent, afin de pouvoir faire des prédictions sur l'interprétation de la suite du discours. Cela implique en particulier de structurer les informations de façon les rendre plus ou moins accessibles. Une façon de modéliser le contexte consiste alors à effectuer des opérations de regroupement ou d'abstraction, qui mènent à des structures arborescentes sur lesquelles peuvent être formulées des contraintes d'accessibilité.

Par la suite, nous comparerons différentes modélisations du contexte, avec leur processus de mise à jour et les contraintes qui en découlent sur la suite de l'interprétation. Nous commencerons par deux modèles ayant connu un grand succès : la théorie du Centrage (Grosz et al., 1995) et la théorie des représentations discursives (DRT, Kamp et Reyle, 1993). Bien qu'elles n'aient pas les mêmes objectifs, elles partagent le souci de modéliser certaines contraintes d'accessibilité pour l'interprétation pronominale. La théorie du Centrage propose des mécanismes sur la base d'un contexte très local, issu de la phrase précédente. La DRT permet de construire un contexte plus global, composé de l'ensemble des référents du discours. L'opposition entre « contexte local » et « contexte global » nous amènera à une interrogation sur la nécessité de structurations contextuelles intermédiaires.

Des solutions intermédiaires ont en particulier été envisagées par des approches consacrées aux structures discursives. Celles-ci considèrent que l'organisation thématique, rhétorique ou intentionnelle d'un discours pourrait fournir des indices pour la structuration du contexte. Nous en présenterons alors les travaux fondateurs : la théorie des structures rhétoriques (RST, Mann et Thompson, 1988) et les travaux de B. Grosz et C. Sidner (1986) sur les relations entre structures intentionnelle et attentionnelle. Parmi les développements plus récents, nous retenons une confirmation empirique des intuitions issues de la RST sur l'accessibilité référentielle (théorie des veines, Cristea et al., 1998), la proposition d'une structure discursive arborescente avec la définition de l'accessibilité le long de la frontière droite (Webber, 1990) et l'intégration des structures rhétoriques dans le cadre de la DRT (S-DRT, Asher, 1993 ; Lascarides et Asher, 1993).

La comparaison s'articulera autour de trois questions auxquelles nous répondrons dans une synthèse en fin de chapitre :

Quelle est la nature des entités composant le contexte ?

Comment la combinaison de ces entités permet-elle de faire des prédictions sur l'interprétation de l'énoncé suivant ?

Ces modélisations tiennent-elles compte du fonctionnement spécifique des différents marqueurs référentiels ?

4.2 Deux extrêmes : théorie du Centrage et DRT

4.2.1 Une approche locale : la théorie du Centrage

Nous avons vu au chapitre précédent (chapitre 0) que les pronoms personnels sont couramment considérés comme référant à une entité ayant un statut cognitif particulier, c'est-à-dire étant focalisée. Cependant, un des problèmes était précisément le calcul des statuts cognitifs. La théorie du Centrage (Grosz et al., 1995) apporte un début de réponse à cette question : son objectif est d'établir une relation entre la structure attentionnelle locale (« focus ») et le choix des expressions référentielles, sachant que cette relation régit la cohérence locale d'un discours. Initialement, la théorie du Centrage est conçue comme le complément de la théorie de la structure discursive (Grosz et Sidner, 1986; cf. 4.3.3 ci-dessous), dans le sens où cette dernière étudie la cohérence d'un discours à travers les structures globales formées par les segments du discours et leurs intentions associées. Néanmoins, l'autonomie acquise par la théorie du Centrage et le succès qu'elle a rencontré très rapidement dans le domaine de l'intelligence artificielle justifient qu'on la considère à part.

La notion centrale – la focalisation – doit être considérée comme une fonction permettant de limiter les inférences nécessaires à la compréhension des énoncés d'un discours. Plus ces inférences demandent de l'effort à l'interlocuteur, moins un discours est considéré comme cohérent. A partir de là, la théorie du Centrage prédit les conditions d'une pronominalisation dans un segment S_{i+1} en fonction de la nature des « centres » du segment précédent S_i. Les « centres » sont les concepts de base de la théorie du Centrage. Est définie comme centre toute entité d'un énoncé qui peut servir à relier cet énoncé à d'autres énoncés du discours. En théorie, il doit s'agir d'objets sémantiques et non pas d'entités linguistiques (phrases, syntagmes,...), mais dans la pratique, les centres restent très liés aux groupes nominaux.

Étant donné une suite de deux segments S_i et S_{i+1} (le plus souvent, deux énoncés), on distingue entre :

> « centres en avant » (CAv_i) : il s'agit de la liste partiellement ordonnée des entités discursives évoquées en S_i. L'ordonnancement se fait sur la base de critères grammaticaux (sujet > objet > autre). Les auteurs soulignent explicitement que d'autres critères comme des structures parallèles ou la notion de thème ne sont pas pertinents. Parmi les centres en avant, est défini comme « centre préféré » (CP_i) le premier élément de la liste CAv_i ;
>
> « centre en arrière » (CAr_{i+1}) : l'élément de CAv_{i+1} ayant le rang le plus élevé en CAv_i.

En fonction de l'évolution du centre en arrière et du centre préféré du segment S_{i+1}, il existe trois types de transitions :

> « continuation » : le centre en arrière de S_{i+1} est le même que celui de S_i (s'il existe) et constitue en même temps le centre préféré de S_{i+1} ;
>
> « rétention » : le centre en arrière de S_{i+1} est le même que celui de S_i (s'il existe), mais il ne constitue plus en même temps le centre préféré de S_{i+1} ;
>
> « changement » : le centre en arrière de S_{i+1} n'est plus le même que celui de S_i.

Le pouvoir prédictif de la théorie du Centrage vient de la définition d'un ordre de préférence sur les transitions entre deux segments (continuation > rétention > changement) et d'une règle postulant que s'il y a pronominalisation en S_{i+1}, alors le centre en arrière de S_{i+1} doit être pronominalisé. L'application de ces principes produirait des discours cohérents.

L'exemple (65), extrait du corpus Ozkan, respecte ces principes : comme la montre l'analyse de l'évolution des centres (Tableau 8), la règle de pronominalisation n'est pas violée et les transitions « moins cohérentes » en début du dialogue viennent d'un changement de thème allant des pyramides (P) vers les égyptiens (E) :

(65) I_1 il faut mettre *les pyramides* après *le soleil* et à la fin *l'horizon*
M_1 tu comprends pourquoi *les égyptiens* en ont bavé pour *les* faire
M_2 s'*ils les* déplaçaient en plus comme ça régulièrement
M_3 et au dernier moment *ils* s'aperçoivent que *l'horizon* n'allait pas sur *les pyramides*
M_4 il a fallu qu'*ils les* déplacent toutes
I_2 c'est vrai quand même ces pyramides (C11Egypte)

	I_1	M_1	M_2	M_3	M_4
CAv	{P, {S,H}}	{E, P}	{E, P}	{E, (H,P}}	{E, P}
CAr	Ø	P	E	E	E
CP	P	E	E	E	E
Transition		rétention	changement	continuation	continuation

Tableau 8 : Calcul des centres et transitions pour l'exemple (65) (P=Pyramides, E=Egyptiens, S=Soleil, H=Horizon)

L'exemple (66) est une modification volontaire de ce même exemple, afin d'illustrer les effets discursifs de la violation de la règle de pronominalisation (M_1 – M_2) et de la préférence de continuation (M_3 – M_4) :

(66) I_1 il faut mettre *les pyramides* après *le soleil* et à la fin *l'horizon*
M_1 tu comprends pourquoi *les égyptiens* en ont bavé pour *les* faire
M_2 si ***les égyptiens*** *les* déplaçaient en plus comme ça régulièrement
M_3 et au dernier moment *ils* s'apercoivent que *l'horizon* n'allait pas sur *les pyramides*
M_4 il a fallu qu'***elles*** soient toutes déplacée par ***les égyptiens***
I_2 c'est vrai quand même ces pyramides (C11Egypte, modifié)

L'apport principal de la théorie du Centrage est sa simplicité calculatoire, permettant des implémentations relativement rapides. Cette facilité de formalisation a été à l'origine de son succès dans le domaine de la compréhension automatique (Brennan et al., 1987 ; Strube, 1998 ; Beaver, 2000), comme dans celui de la génération d'expressions référentielles (McCoy et Strube, 1999)[25]. Par ailleurs, les principes de cette théorie ont été testés dans des études multilingues et certains aspects sont confirmés par des études psycholinguistiques (Gordon et al., 1993).

Il s'agit, en revanche, d'une approche extrêmement locale, puisque l'on reconstruit un nouveau contexte sous forme de liste de centres pour chaque phrase et ce contexte est limité à un seul segment. Si cette approche peut éventuellement se justifier empiriquement pour le traitement des pronoms[26], elle n'est pas généralisable à notre objectif – la modélisation contextuelle pour la résolution de tout type d'expressions référentielles. Bien plus, la théorie du Centrage ne couvre même pas tous les emplois pronominaux : les contextes proposés sont trop restrictifs pour couvrir des emplois collectifs, des

[25] Les travaux cités sont présentés au chapitre consacré aux algorithmes et implémentations (chapitre 5).
[26] cf. les travaux de J. Hobbs (1978) qui ont montré pour l'anglais que 98 % des pronoms pouvaient se résoudre par un algorithme « naif » cherchant l'antécédent le plus proche en accord avec le pronom dans la même phrase ou la phrase précédente (cf. la section 5.2.1)

glissements au générique ou tout simplement des reprises pronominales dépassant le cadre de la phrase précédente, comme c'est le cas de l'exemple (67) :

(67) a. Gisela aime bien la choucroute.
b. Helmut préfère les saucisses.
c. Mais *ils* s'entendent tout de même très bien.

4.2.2 Une approche globale : la théorie des représentations discursives (DRT)

La DRT (Kamp et Reyle, 1993) propose une modélisation dynamique de la compréhension d'un discours qui passe par la construction d'une représentation discursive d'un texte. L'objectif consiste à fournir une structure permettant de faire un certain nombre de prédictions sur l'interprétation de la suite du discours : l'interprétation aspectuelle, l'interprétation temporelle et l'interprétation référentielle. Dans la DRT, les informations véhiculées par le discours sont sauvegardées dans des DRS (*Discourse Representation Structure*), formées de deux composantes : l'univers du discours avec des variables pour les objets et les événements introduits par le discours d'une part et des conditions portant sur ces variables, d'autre part.

La mise à jour d'une DRS s'effectue de façon dynamique. L'intégration d'un nouvel énoncé se fait sur la base de sa structure syntaxique. En fonction de cette structure s'appliquent des règles d'introduction de nouveaux référents et de conditions. Un indéfini est par exemple considéré comme introduisant une nouvelle variable et une nouvelle condition, correspondant à la prédication introduite par la tête nominale sur cette variable. Pour résoudre une référence pronominale, on introduit une nouvelle variable qui sera liée, par une condition d'égalité, à un référent existant dans l'univers du discours, à condition que celui-ci soit accessible et « approprié ». Les contraintes d'accessibilité sont modélisées en fonction de la structure syntaxique des énoncés. Ainsi, un référent introduit sous la portée d'un opérateur de négation ou dans une construction conditionnelle ne sera plus pleinement accessible pour la suite du discours. Les prédictions de la DRT sur l'interprétation référentielle et plus particulièrement sur l'interprétation pronominale, sont donc plutôt formulées en termes d'inaccessibilité qu'en termes d'accessibilité. En effet, parmi les référents non inaccessibles sur critères syntaxiques, la DRT n'indique pas comment choisir le référent le plus approprié.

Le traitement des groupes nominaux définis pose également problème, car ceux-ci ne peuvent être traités qu'en s'inspirant du traitement soit des syntagmes nominaux indéfinis (introduction d'un nouveau référent), soit des pronoms (résolution sur un référent existant). Ceci empêche, entre autres, un traitement approprié des anaphores associatives et des descriptions définies opérant une extraction à partir d'un groupement de référents (Gaiffe et al., 1997). Une extension possible pour la prise en compte des définis a été proposée par J. Bos et al. (1995). Nous y reviendrons lors de la synthèse de ce chapitre (4.4.3), mais nous pouvons déjà faire remarquer que cette proposition ne résout pas tous les problèmes, dans la mesure où elle ramène le problème de la référence à une relation dont le prototype est le « liage » *(linking)* entre un référent accessible et une expression référentielle. Ceci revient à ramener la fonctionnalité première du défini à celle du pronom, point de vue contraire aux descriptions linguistiques (cf. le chapitre 2).

Le traitement de l'exemple (68), repris de la section précédente, illustre la mise en œuvre de la DRT (Figure 7) et étaye quelques-unes des réflexions ci-dessus[27] :

[27] Nous avons omis, pour des raisons de clarté de la présentation, le traitement des prédications verbales.

(68) I_1 il faut mettre *les pyramides* après *le soleil* et à la fin *l'horizon*
M_1 tu comprends pourquoi *les égyptiens* en on bavé pour *les* faire
M_2 s'*ils les* déplaçaient en plus comme ça régulièrement
M_3 et au dernier moment *ils* s'aperçoivent que *l'horizon* n'allait pas sur *les pyramides*
M_4 il a fallu qu'*ils les* déplacent toutes
I_2 c'est vrai quand même ces pyramides (C11Egypte)

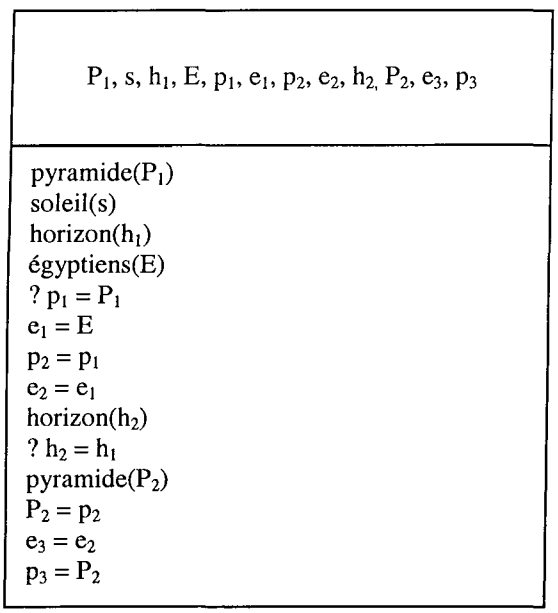


Figure 7 : Traitement de l'exemple (68) à la DRT

Les définis de la première intervention (P_1, s, h_1) peuvent être considérés comme renvoyant à des référents discursifs mentionnés au début du dialogue (avant l'extrait) et ne posent donc pas de problème. Il n'en est pas de même pour le défini *les égyptiens* en M_1 dont le référent n'a pas été mentionné auparavant et qui doit donc être accommodé, sans que l'on sache exactement à partir de quelles connaissances.

Le pronom *les* (M_1) est doublement problématique : il renvoie à un référent pluriel, la DRT ne se donne pas les moyens de décider entre P_1 (*les pyramides*) et E (*les égyptiens*). Même en supposant une décision en faveur de P_1[28], il faudra s'interroger si, dans ce contexte, on n'a pas à faire à une anaphore divergente – les référents de P_1 et de p_1 ne sont effectivement pas les mêmes – et quelles sont alors les conséquences d'une mise en égalité dans le cadre du dialogue de commande, où il s'agira, *in fine*, d'identifier des objets à manipuler.

Le premier de ces problèmes liés au pronom – le choix du référent du discours – se retrouve de la même façon pour e_1, p_2, e_2, e_3 et p_3. Le second problème – l'hypothèse des anaphores divergentes – se pose à nouveau pour h_1 et P_2. Enfin, l'extrait pose aussi le problème de la transposition des prédictions formulées pour un discours « bien formé » par rapport à une norme de l'écrit. On a en effet le début d'une structure conditionnelle en M_2, mais cette structure étant déviante par rapport à la norme de l'écrit et il nous semble que des prédictions dérivées d'une telle structure doivent être manipulées avec prudence dans des dialogues oraux.

L'apport de la DRT par rapport à nos objectifs concerne deux points. D'une part, la DRT nous propose un mécanisme dynamique et calculable pour la construction d'un contexte. D'autre part, elle

[28] Cette décision pourrait s'appuyer sur les principes du gouvernement et du liage (cf. la section 5.2.1).

intègre la notion d'accessibilité qui permet de déduire des contraintes sur la suite de l'interprétation du discours. Les limites de cette approche reposent essentiellement sur ce que nous appellerons son aspect « global », en opposition à la localité de la théorie du Centrage : modulo des contraintes d'inaccessibilité sur critères syntaxiques, une DRS maintient l'accès à tous les référents discursifs introduits préalablement. Le fait qu'elle s'appuie exclusivement sur la syntaxe des énoncés pour limiter l'accès à cette liste, la rend insuffisante pour un modèle de la résolution de la référence, d'autant plus que les structures syntaxiques de l'oral peuvent ne pas respecter les normes définies pour l'écrit (Blanche-Benveniste, 1990).

4.2.3 Et s'il y avait un juste milieu ?

Comme nous venons de le voir, la théorie du Centrage peut être considérée comme une approche extrêmement locale. Le contexte d'interprétation est reconstruit pour chaque nouveau segment, entraînant ainsi la perte de l'accès à tous les référents des segments précédents et par conséquent, la possibilité de traiter des expressions référentielles demandant d'autres connaissances que celles issues directement du segment précédent. Son avantage par rapport à la DRT est une structuration des éléments du contexte sous la forme d'une liste ordonnée. Ceci permet de pallier, en partie, le problème mentionné ci-dessus de la DRT qui, elle, peut être considérée comme une approche globale. Ici, le contexte est en effet le réservoir non structuré de tous les référents introduits dans le discours, en dehors de ceux qui sont inaccessibles sur la base de critères syntaxiques. Un des points critiques est alors la surabondance de référents disponibles. Ceci s'exprime par exemple par des contraintes insuffisantes sur la résolution des pronoms, la seule règle étant le liage du pronom à un « *suitable discourse referent chosen from the universe of the DRS* »[29] (Kamp et Reyle, 1993 : 70), ce qui ne permet pas réellement de choisir parmi plusieurs candidats. Le problème de la DRT est en quelque sorte qu'elle est trop globale. A part la syntaxe, elle ne considère pas d'autres facteurs susceptibles d'influer sur l'accessibilité des référents, comme par exemple les informations encyclopédiques ou perceptives, la prise en compte des intentions des locuteurs ou la structuration thématique du discours.

La nécessité d'un niveau de structuration intermédiaire a en effet été envisagé dans le cadre des sémantiques formelles. L'idée selon laquelle l'identification d'un référent ne se fait pas uniquement à partir de la description linguistique donnée par l'expression à interpréter, mais aussi relativement à un sous-ensemble contextuel, est par exemple défendue par Westerstahl (1984) et développée, entre autres, par J. Groenendijk et al. (1995) et P. Dekker (1998). Ainsi, J. Groenendijk et ses collègues considèrent les descriptions indéfinies et définies comme des quantificateurs dynamiques, éventuellement restreints à des ensembles contextuels. Étant donné un état d'information, ils définissent l'ensemble contextuel relatif à cet état comme le sous-ensemble des individus du domaine de discours ayant été liés à un référent du discours. Ils considèrent ensuite que les descriptions indéfinies et définies imposent certaines contraintes sur le contenu de cet ensemble contextuel. Si les contraintes ne sont pas remplies, l'interprétation échoue. Dans le cas contraire, l'ensemble contextuel est mis à jour en fonction de la quantification et du contenu descriptif de l'expression. Ce principe permet en particulier de traiter des expressions comme *another N* et *the one and the other* dont l'interprétation dépend effectivement d'un ensemble fourni par le contexte et qui ne sont pas traitées dans la version standard de la DRT (Kamp et Reyle, 1993).

Pourtant, cette approche se heurte à deux problèmes : le premier réside dans le fait que dans certains cas, la compatibilité du contenu descriptif n'est pas suffisante pour identifier un élément de l'ensemble contextuel. Ainsi, dans « *A man came to the doctor. The man said...* », *man* s'applique aussi bien à *man* qu'à *doctor*. Il est alors nécessaire de sauvegarder les informations sur la manière

[29] « un référent approprié choisi dans l'univers de la DRS »

dont l'entité contextuelle a été introduite. Cela signifie d'ores et déjà que les ensembles contextuels doivent être structurés selon certains critères liés au discours et nous verrons dans la suite que d'autres critères peuvent être nécessaires. Un deuxième problème est soulevé : l'ensemble contextuel tel que défini par les auteurs correspond toujours à la totalité des entités introduites discursivement. Si cela ne pose pas de problème pour les exemples donnés par les auteurs (constitués maximalement de trois phrases), il est probable que les domaines de quantification dans des textes ou dialogues réels correspondent le plus souvent à des réels sous-ensembles contextuels. Mais le problème de la création de ces sous-ensembles est relégué (silencieusement !) à une composante pragmatique : « *We will henceforth, silently, assume some pragmatic mechanism which serves to restrict natural language quantifiers to appropriate domains of quantification* »[30] (Dekker, 1998 : 9).

Restent donc à déterminer les critères de formation et de (dés-)activation de ces sous-ensembles. Une idée largement explorée consiste à former de tels ensembles sur la base de la structuration thématique du discours. Dans le chapitre suivant (4.3), nous présenterons donc quelques travaux marquants, toujours en nous demandant comment ces propositions peuvent répondre à notre interrogation sur des structures contextuelles intermédiaires entre un réservoir pour la totalité des référents du discours (DRT) et les seuls référents introduits dans le dernier segment (théorie du Centrage).

4.3 Structures hiérarchiques du discours

4.3.1 Relations rhétoriques et organisation textuelle hiérarchique

La *Rhetorical Structure Theory* (RST, Mann et Thompson, 1988) propose une approche descriptive de l'organisation hiérarchique textuelle. Son objectif est l'identification des relations fonctionnelles (relations rhétoriques) entre les parties d'un texte. L'analyse d'un texte est précédée par sa division en unités, qui sont des segments de texte adjacents. L'organisation hiérarchique du texte est ensuite représentée par une structure arborescente qui fait apparaître à la fois les parties du texte et les relations entre celles-ci. Selon les auteurs, cette structuration hiérarchique et rhétorique traduirait l'organisation communicative textuelle basée sur les intentions du locuteur.

A partir de l'observation d'une asymétrie entre les segments d'une même relation (possibilité de suppression d'un segment sans perte d'informations essentielles), la RST introduit la notion de nucléarité. L'interprétation fonctionnelle de la nucléarité peut se faire, selon les auteurs, en considérant la fonction de la communication comme constructrice de mémoire. Dès lors que la structure rhétorique d'un texte représente un schéma d'accès à cette mémoire, la nucléarité permettrait d'expliquer l'accès aux satellites (informations secondaires) à partir des noyaux (informations principales).

Bien qu'il soit en pratique assez difficile de transposer les principes de la RST à l'analyse de dialogues oraux, nous avons essayé de reconstituer, dans la Figure 8, ce qui pourrait être la structure rhétorique des énoncés I_5 à M_3 de l'exemple (69) :

(69) I_1 maintenant, il faut que tu prennes un gros cercle
I_2 pour faire le soleil
I_3 tu le mets en haut à droite
M_1 ouais + [geste de manipulation d'un cercle]
I_4 voilà
I_5 maintenant, tu vas prendre un petit carré
I_6 et le mettre en bas à gauche

[30] « Nous supposerons, silencieusement, à partir de maintenant un mécanisme pragmatique qui sert à restreindre les quantifieurs des langues naturelles à des domaines de quantification appropriés. »

I_7 pour faire une maison
M_2 [geste de manipulation d'un carré]
I_8 maintenant tu en prends un autre
M_3 avec un toit aussi ?

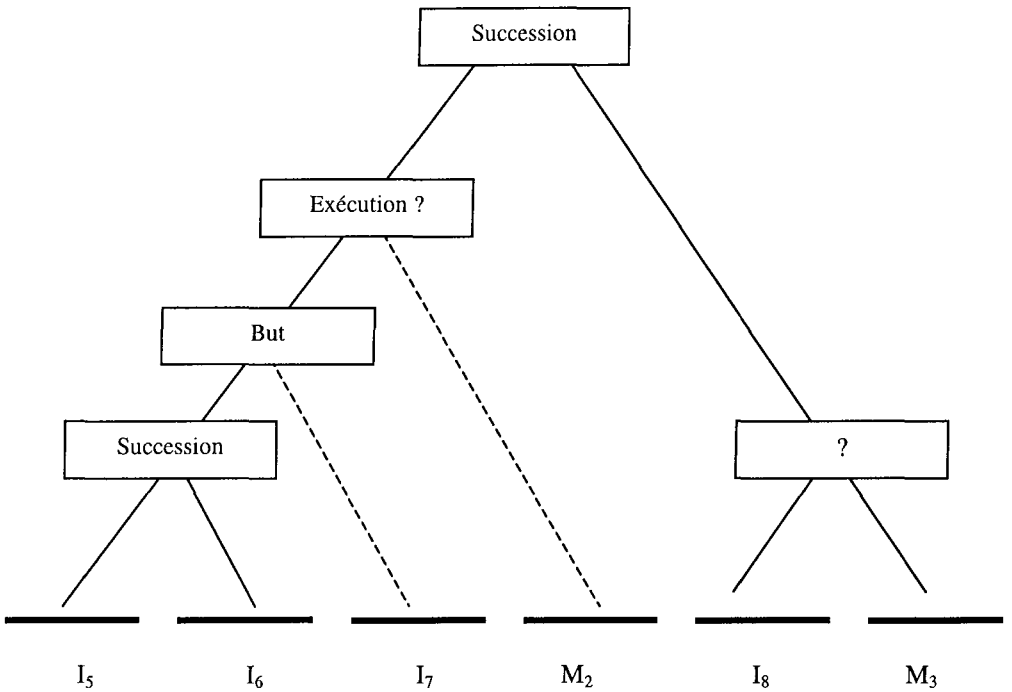


Figure 8 : Structure rhétorique des énoncé I_5 à M_3 de l'exemple (69)

Les feuilles de l'arbre correspondent aux énoncés et chaque nœud indique une relation binaire (ou d'arité supérieure) entre segments élémentaires. Nous avons traduit la notion de nucléarité par l'utilisation des lignes pleines ou pointillées. Toutes les lignes pointillées remontent à partir de constituants satellites. Ainsi, la relation qui lie le segment constitué par les énoncés I_5 et I_6 (eux-mêmes reliés par une relation de succession) à l'énoncé I_7 correspond à une relation de but dans laquelle le noyau est constitué des énoncés I_5 et I_6 et le satellite de l'énoncé I_7.

Selon W. Mann et S. Thompson, cette structure suit les intentions des locuteurs. Cette affirmation est intéressante dans la mesure où l'objectif *in fine* de la compréhension dans le dialogue homme-machine est précisément d'interpréter les intentions du locuteur humain. Dans le cadre de la RST, cette interprétation s'avère néanmoins difficile, pour deux raisons : premièrement, la RST propose une analyse purement descriptive et *a posteriori* d'un texte. Or, analyser les intentions du locuteur après coup est incompatible avec les impératifs du DHM, où il faut construire des structures contextuelles de façon dynamique. Deuxièmement, même en supposant que l'on dispose d'un mécanisme de construction dynamique, on constate que les relations proposées sont inadaptées au dialogue : comment qualifier, par exemple, la relation entre $I_5 - I_7$ et M_2, ou celle entre I_8 et M_3[31]? Bien plus, il nous semble qu'elles sont même insuffisantes pour des monologues (savoir que I_5 et I_6 sont liés par une relation de succession n'aide pas réellement à la détermination d'une intention quelconque).

[31] Il nous est apparu au cours de la discussion avec des personnes autour de nous qu'il pouvait y avoir une ambiguïté sur l'attachement de M_3 : tel que nous l'avons compris, M_3 est une élaboration de I_7 (dans le sens « faire une maison avec un carré (pour les étages) et aussi un toit ». Or, d'autres personnes l'ont compris comme le début d'un sous-dialogue relatif à I_8. Étant donné qu'il n'y pas de manipulation d'un premier triangle pour le toit de la maison à construire en I_7, cette variante nous semble néanmoins moins probable. Ceci dit, le problème de l'adaptation des relations rhétoriques au dialogue ne serait pas pour autant résolu : même si l'on voulait attacher M_3 à I_8 de cette façon, il n'y aurait pas de relation appropriée.

La deuxième notion qui nous intéresse par rapport à notre objectif est celle de nucléarité, dans la mesure où elle pourrait éventuellement aider à formuler des contraintes sur l'accessibilité de certaines informations du contexte et par là, à constituer des hypothèses anticipatoires sur la suite du dialogue. En l'appliquant à notre exemple, on pourrait supposer que les informations secondaires (celles véhiculées par les satellites) deviennent, sous certaines conditions, inaccessibles pour des reprises dans la suite des énoncés. Or, il n'en est rien, comme le montre l'énoncé M_3. Pour comprendre correctement l'expression *un toit* (M_3), il faut disposer de l'information véhiculée dans le satellite I_7 (« pour faire une maison »), mais cette information n'est plus accessible.

Compte tenu du fait que l'analyse proposée est purement descriptive, ne s'applique qu'aux textes écrits et ne peut se faire qu'*a posteriori*, l'approche n'est pas transposable à nos objectifs. Par rapport à nos interrogations sur la constitution d'un contexte et en faisant un parallèle entre la structure rhétorique proposée et la modélisation du contexte, nous retenons pourtant deux idées. D'une part, les auteurs introduisent, à travers la notion de nucléarité, une pondération de l'importance des informations et abordent par là la question de l'accessibilité des entités du contexte. Ce point a été repris récemment par la théorie des veines (Cristea et al., 1998), ayant proposée une validation expérimentale de la notion de nucléarité (cf. la section 4.3.2). D'autre part, l'idée d'un lien entre les intentions et la structuration du discours demande à être approfondie par une réflexion sur la nature des relations à considérer et la façon de calculer ces relations. Des réflexions allant dans ce sens sont développées en particulier par B. Grosz et C. Sidner (1986), par B.Webber (1991) et par N. Asher (1993). Elles font l'objet des sections 4.3.3, 4.3.4 et 4.3.5, respectivement.

4.3.2 Structure rhétorique et accessibilité référentielle : la théorie des veines

D. Cristea et ses collègues (1998) proposent un modèle dont l'objectif est d'étendre les principes de la théorie du Centrage à des segments plus larges allant jusqu'au discours. L'idée principale est de déterminer des domaines d'accessibilité pour résoudre des expressions référentielles sur la base de la structure rhétorique, telle que proposée par la RST et de valider ces domaines d'accessibilité par une expérimentation trilingue (anglais, français, roumain). La segmentation du discours est donc comparable à celle effectuée par la RST. Le problème du choix automatique des relations n'a pas été résolu, puisque la segmentation, l'annotation relationnelle et le choix des noyaux et des satellites ont été entièrement effectués à la main.

A partir des informations données par la structure rhétorique arborescente et plus particulièrement de la distinction entre noyau et satellite, les auteurs proposent de calculer le domaine d'accessibilité pour les référents des expressions de l'énoncé courant. L'hypothèse fondamentale est que ce domaine d'accessibilité est fourni par un sous-ensemble de segments nucléaires dans le discours antérieur, correspondant en quelque sorte aux « veines » du discours.

Concernant notre objectif, l'apport principal de la théorie des veines est une validation empirique d'une notion d'accessibilité basée sur la structure rhétorique. En effet, dans le corpus d'expérimentation – constitué de 176 segments (textes littéraires anglais, français et roumain) et contenant 317 expressions référentielles – plus de 87 % des expressions trouvent leur antécédent directement sur une « veine nucléaire », comme prédit par la théorie. Cependant, l'application de l'algorithme proposé reposant sur un texte pré-segmenté et annoté manuellement, le problème essentiel réside dans l'automatisation de ce processus.

4.3.3 Structures intentionnelle et attentionnelle

L'objectif de la théorie proposée par B. Grosz et C. Sidner (1986) consiste à fournir un modèle de la structure discursive de textes ou de dialogues reposant sur l'intentionnalité. Le modèle poursuit en partie les travaux B. Grosz (1981) sur la notion de focalisation pour la compréhension de dialogues finalisés et ceux de C. Sidner (1979, 1983) sur l'interprétation des anaphores définies[32]. L'idée de base est celle d'une structure discursive – composée de trois niveaux (linguistique, intentionnel et attentionnel) – qui fournit des contraintes sur la manière dont un énoncé s'intègre dans l'ensemble du discours : à partir de là, il serait possible d'expliquer par exemple les interruptions, l'emploi des connecteurs discursifs ou encore l'interprétation des expressions référentielles. Étant donné nos objectifs, la composante centrale du modèle est la structure intentionnelle, dans la mesure où c'est elle qui régit l'évolution de la structure attentionnelle et donc l'accessibilité des entités pour des actes de référence.

La structure linguistique comprend la suite des énoncés regroupés en segments du discours. A chaque segment du discours est associée une intention locale. Chacune de ces intentions contribue à l'accomplissement de l'intention globale du discours. L'organisation hiérarchique de ces intentions se traduit par des relations de dominance et de succession temporelle. Par rapport à la RST, l'essentiel pour la détermination de la structure du discours serait non pas la nature particulière des intentions (succession, but etc.), mais la nature de leurs relations. Ces relations, contrairement aux relations rhétoriques, seraient très peu nombreuses et se ramènent, pour B. Grosz et C. Sidner, à une relation de dominance et une relation de succession.

Parallèlement à la structure intentionnelle, l'espace attentionnel est modélisé par une structure focale. Celle-ci consiste en une pile d'espaces focaux, régie par des règles de destruction ou de création. Chaque espace est associé à un segment du discours et contient les objets, propriétés ou relations introduits dans ce segment. L'évolution de la structure focale est étroitement liée aux relations hiérarchiques qu'entretiennent les intentions véhiculées par les segments du discours. En fonction de celles-ci, des informations apparaissent ou disparaissent de la pile. L'objectif de cette structure est de fournir le contexte nécessaire à l'interprétation des énoncés suivants, y compris à la résolution de la référence. Le principe consiste à rendre accessibles toutes les entités présentes dans l'espace focal associé au segment courant ou dans les espaces focaux inférieurs de la pile (correspondant à des segments véhiculant des intentions dominantes par rapport au segment traité).

La notion de structure attentionnelle ou focale répond en partie au besoin d'une structure contextuelle capable de contraindre l'intégration et l'interprétation d'un nouvel énoncé dans une suite discursive ou dialogique. Les auteurs proposent en particulier de lier la hiérarchie de la structure intentionnelle – les objectifs et les sous-objectifs des locuteurs – à la hiérarchie des entités saillantes ou accessibles dans la structure attentionnelle. En somme, c'est la relation de dépendance des intentions qui régit l'accessibilité des entités dans le focus. Cette vision des choses n'est finalement pas très éloignée de ce que propose la RST : si l'on fait abstraction de la variété des relations proposées par la RST et en n'en retenant que le concept de nucléarité, les deux théories se rejoignent par une correspondance entre la relation noyau – satellite (RST) et celle de dominance (Grosz et Sidner, 1986)[33].

[32] Les travaux de C. Sidner et plus particulièrement son algorithme pour le traitement des anaphores définies, font l'objet d'une présentation dans la section 5.2.3.

[33] Ce rapprochement est d'ailleurs discuté plus amplement par M. Moser et J. Moore (1996).

La première critique que l'on peut adresser à cette approche concerne les liens entre la structuration d'une tâche donnée et la structure intentionnelle. Si on suppose, dans le cadre du dialogue finalisé, que les objectifs et sous-objectifs des locuteurs sont guidés par la tâche à accomplir, on aboutit alors à une structuration de l'espace focal parallèle à celle de la tâche. Lorsque la tâche n'est pas entièrement spécifiée à l'avance, chose que les auteurs considèrent pourtant expressément, il nous semble que la reconnaissance des intentions et de leur relations – qui est pourtant la base pour la construction de l'espace focal – reste un problème majeur de cette approche.

Par ailleurs, les critères de segmentation et de reconnaissance proposés par les auteurs (marqueurs linguistiques, intentions phrastiques et leurs combinaisons) restent flous et ne s'appliquent en aucun cas systématiquement. Devant cette inopérationnalité, nous avons renoncé au traitement de l'exemple (69). De plus, les critères linguistiques pour la segmentation n'échappent pas à un problème de circularité, comme le note G. Sabah (1988). Ainsi, certains phénomènes référentiels (la reprise par un groupe nominal plutôt que par un pronom) sont supposés marquer une frontière de segment et donc un changement dans la hiérarchie intentionnelle, alors que le but de la théorie est justement d'expliquer ce type de phénomènes sur la base de la structure focale, elle-même calculée à partir de la structure intentionnelle.

4.3.4 Arbres et oracles

B. Webber (1991) développe une structure discursive afin d'interpréter ce qu'elle appelle des « déictiques discursifs ». Il s'agit de trouver les référents pour les pronoms démonstratifs anglais *this*, *that* et *it*. Comme le note B. Webber, les référents de ces expressions ne sont ni des segments textuels, ni des groupes nominaux, mais des événements ou propositions introduits par les segments discursifs précédents. Dans ce cadre, l'auteur constate que ce ne sont pas tous les segments précédents qui peuvent fournir des antécédents d'événements ou de propositions. Cette observation mène à l'hypothèse selon laquelle l'accessibilité des informations nécessaires à la résolution de la référence déictique discursive est régie par la structure du discours considéré.

Cette structure du discours est une structure dynamique et récursive, incrémentée au fur et à mesure de la compréhension du discours. Ses unités minimales, les propositions, se combinent pour former des segments discursifs d'un niveau hiérarchique supérieur. La structure discursive est donc un arbre dont les nœuds correspondent à des segments et les arcs à des relations de parenté (père ou frère). L'abandon de la grande variété des relations discursives de la RST au profit d'une simple notion de succession (frère) ou de dominance (père) est donc comparable à ce que proposent la théorie des veines (Cristea et al., 1998) et le modèle de la structure intentionnelle (Grosz et Sidner, 1986). L'intégration d'une nouvelle proposition à l'arbre discursif se fait par l'une des deux opérations de mise à jour que sont l'attachement et l'adjonction. L'attachement permet d'intégrer la proposition en tant que dernier fils sous un nœud père existant, alors que l'adjonction crée un nouveau nœud non terminal sous lequel se regroupent plusieurs autres segments et le segment à intégrer.

Le traitement des énoncés I_5 à I_8 de l'exemple (70), déjà traités à la RST (4.3.1), illustre l'incrémentation de la structure discursive :

(70) I_1 maintenant, il faut que tu prennes un gros cercle
I_2 pour faire le soleil
I_3 tu le mets en haut à droite
M_1 ouais + [geste]
I_4 voilà
I_5 maintenant, tu vas prendre un petit carré
I_6 et le mettre en bas à gauche
I_7 pour faire une maison
M_2 [geste]
I_8 maintenant tu en prends un autre
M_3 avec un toit aussi ?

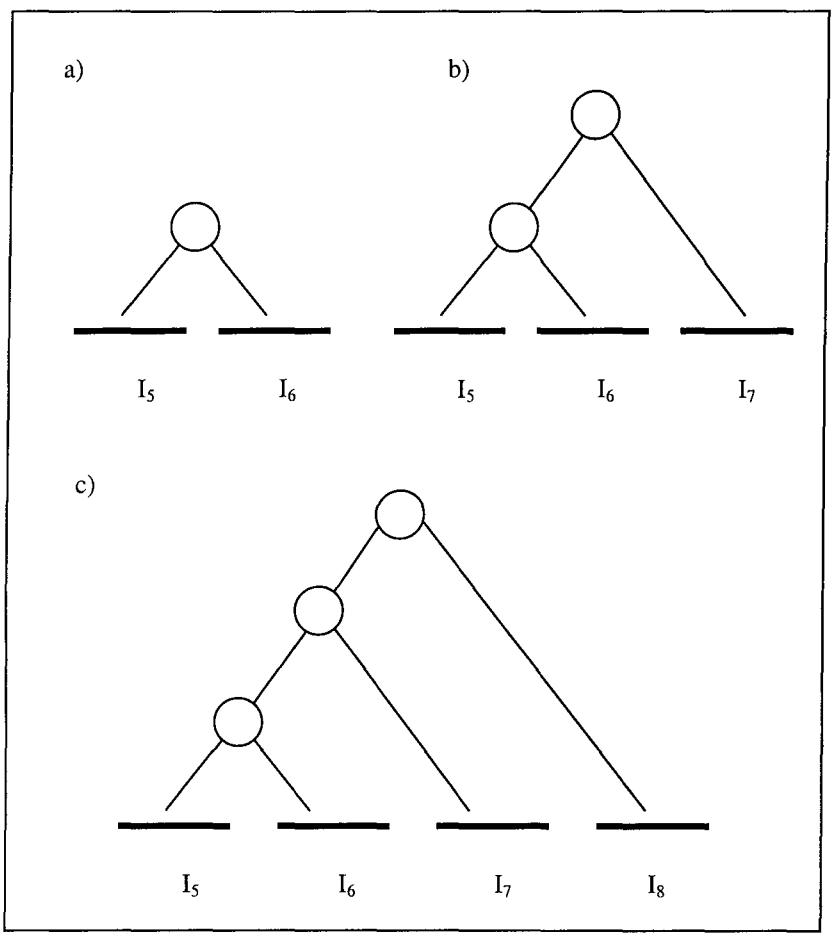


Figure 9 : Analyse de l'exemple (70) selon Webber (1991)

Le segment I_6 est lié à I_5 par une adjonction (a). Pour ce qui est de I_7, on peut supposer qu'il s'attache non pas directement en tant que frère à I_6, mais à l'ensemble formé par I_5 et I_6, ce qui donne encore lieu à une adjonction (b). Nous laissons de côté le traitement de M_2 et supposons que I_8 s'ajoute également par adjonction, c'est-à-dire par création d'un nouveau nœud non terminal regroupant I_5 à I_7 (c).

Sur la base de cette structure discursive, B. Webber définit les segments qui sont accessibles pour la résolution de la référence des pronoms déictiques discursifs : il s'agirait de tous les segments situés sur la frontière droite de l'arbre. Cette définition qui a déjà été suggérée par d'autres auteurs, en particulier par R. Scha et L. Polanyi (1988), est d'ailleurs comparable à ce que proposent B. Grosz et C. Sidner (1986) : le parcours de la frontière droite de l'arbre de la structure discursive proposé par B. Webber correspond au parcours de la pile d'espaces focaux et les opérations d'attachement et

d'adjonction sont comparables aux opérations d'ajout ou de retrait (suivies d'ajout) de la pile. L'utilisation d'une seule structure arborescente assortie d'un algorithme d'insertion pour les nouvelles entités permet donc de modéliser en même temps l'évolution de la structure discursive (ou intentionnelle) et l'évolution des espaces focaux. D'ailleurs, B. Grosz et C. Sidner font remarquer elles-mêmes que « *the focusing structure is parasitic upon the intentional structure* »[34] (1986 :180).

Cette proposition appelle pourtant deux remarques critiques. Premièrement, B. Webber ne donne pas de critères ni pour déterminer l'endroit d'ancrage d'une nouvelle proposition, ni pour le choix de la relation et donc de l'opération de mise à jour à effectuer (elle suppose que cela se fait par un « oracle »). Le problème de l'automatisation de ce processus n'est donc toujours pas résolu. Deuxièmement, tant que le contenu des nœuds non terminaux n'est pas clairement défini, le principe de l'accessibilité de la frontière droite reste vague. Si l'on suppose par exemple que le segment regroupant deux sous-segments `a` et `b` contient la somme plurielle des contenus propositionnels de ces sous-segments `a` et `b` (ce que pourrait suggérer la notation `(a,b)` utilisée par B. Webber), alors tous les contenus propositionnels, bien qu'à des niveaux hiérarchiques différents, seraient finalement accessibles à tout moment. Et si cela n'est pas le cas, comment calcule-t-on alors le contenu des nœuds non terminaux ?

4.3.5 DRT + relations rhétoriques = S-DRT ?

N. Asher propose, en collaboration avec A. Lascarides, d'augmenter la DRT (4.2.2) par des mécanismes qui rendent compte de la structure discursive (Asher, 1993 ; Lascarides et Asher, 1993). Cette proposition va précisément dans le sens d'une introduction de structures plus locales remplaçant une seule DRS « globale ». Elle est motivée par plusieurs observations. D'abord, il semble tout à fait possible de reprendre par des anaphores non seulement des objets correspondants, dans la DRT, à des variables de l'univers du discours, mais aussi des objets abstraits comme les événements, les propositions ou les faits (cf. Webber, 1991). Or, modéliser la reprise d'un complexe de propositions suppose que l'on dispose d'un mécanisme capable de représenter ce complexe de propositions. Cela est impossible dans la DRT, où toutes les propositions se fondent dans les variables et conditions d'une même DRS. De plus, N. Asher soulève un problème lié à l'interprétation temporelle telle qu'elle est traitée dans la DRT. Sur la base de la seule structure syntaxique, les événements e_1 et e_2, décrits dans les deux énoncés de l'exemple (71), ne reçoivent qu'une seule interprétation temporelle : e_1 précède e_2.

(71) Gisela se lève. Le soleil se couche.

Or, la prise en compte d'autres connaissances (par exemple, le fait que Gisela se lève de son transat, parce que le pâle soleil d'hiver de Wanne-Eickel se couche et ne la réchauffe plus...) pourrait mener à une interprétation alternative : e_2 précède e_1.

La solution à ces problèmes passe, pour N. Asher, par une intégration des structures rhétoriques (4.3.1) dans des représentations du discours. Au lieu de construire une seule DRS pour le discours entier, il propose de construire des S-DRS (*Segmented Discourse Representation Structures*) pour chaque énoncé et de lier celles-ci par des relations discursives comme *Continuation*, *Elaboration*, *Background* ou *Cause*.

L'ancrage d'une nouvelle S-DRS sur la structure discursive se fait sur la base d'un ensemble de connaissances syntaxiques, sémantiques, pragmatiques et extralinguistiques : relations partie-tout, relations d'inclusion, relations causales, parallélisme ou contraste entre deux événements,

[34] « La structure focale est parasitaire par rapport à la structure intentionnelle. »

connaissances sur un regroupement possible sous un même topique ou connaissances sur la structure du discours jusqu'alors traité. Ces connaissances sont représentées dans une base de connaissances sous forme de faits et de règles d'inférence. L'ancrage d'une nouvelle S-DRS se fait en deux étapes : dans un premier temps, on calcule l'ensemble des points d'ancrage accessibles et l'on choisit le point d'ancrage effectif. Ensuite, on détermine la relation discursive entre l'énoncé à traiter et l'énoncé de la structure sur lequel il sera greffé.

Les contraintes sur les points d'ancrage suivent des intuitions telles que celles formulées par R. Scha et L. Polanyi (1988) et B. Webber (1991) qui introduisent les notions de relation de coordination et de subordination entre segments discursifs. Pour A. Lascarides et N. Asher, une nouvelle phrase peut s'ancrer soit sur la phrase précédente du discours (en coordination), soit sur une phrase qui est élaborée ou expliquée (en subordination). Dans une structure hiérarchique représentée sous forme arborescente, les points d'ancrage ouverts correspondent alors, tout comme pour B. Webber et d'autres, aux nœuds situés sur la frontière droite de cette structure. La notion de point d'ancrage ouvert est liée étroitement à la notion de cohérence et à la notion d'ambiguïté discursive : lorsqu'il est impossible de déterminer un point d'ancrage pour l'énoncé à traiter, cet énoncé ne peut être intégré dans la structure discursive et on parlera d'un discours incohérent. Lorsqu'il est impossible de choisir un point d'ancrage parmi tous les points accessibles, le discours est ambigu.

L'ancrage en coordination correspond à la relation de narration. Celle-ci est considérée comme relation par défaut et s'instancie à chaque fois qu'on ne dispose pas d'informations supplémentaires permettant d'inférer un autre type de relation. Ce principe traduit, selon les auteurs, la maxime dite « de manière » de Grice[35]. Une seule condition contraint cette relation discursive : il doit être possible de regrouper sous un même topique les deux énoncés et ce sans que ce topique soit égal au contenu d'un des énoncés. La notion de topique reste relativement imprécise : un topique commun à deux énonces est subordonné soit à une égalité des agents et/ou patients impliqués dans les événements relatés, soit à une relation de parallélisme ou de contraste entre les deux événements. Cette contrainte expliquerait l'incohérence des discours tels que (72) par l'impossibilité de trouver un topique commun :

(72) My car broke down. The sun set.

Lorsque l'on dispose d'informations supplémentaires susceptibles de mettre en cause la relation de narration (par exemple, des connecteurs linguistiques ou des connaissances sur des relations causales entre deux événements), le principe de spécificité s'applique et la relation discursive est alors déterminée sur la base de ces informations[36]. En ce qui concerne les connaissances extralinguistiques, la base des connaissances contiendrait par exemple des lois de « cause à effet », permettant d'inférer des relations *Explanation* ou *Result*, des lois sur les sous-événements d'un événement, permettant d'inférer une relation *Elaboration* et des lois sur la possibilité de chevauchement d'états et d'événements, permettant d'inférer une relation *Background*.

L'application des principes de la S-DRT aux énoncés I_5 à M_3 de notre exemple (73), donnant lieu à la structure reproduite par la Figure 10, pose évidemment le problème des relations spécifiques au dialogue. Dans notre cas, nous introduisons une relation *OrderExec* qui sera instanciée lorsqu'un ordre est suivi de l'exécution de l'action demandée. Afin de traiter un maximum de problèmes en un minimum d'espace, nous avons changé l'ordre entre I_6 et I_7[37] :

[35] « *Donnez les informations dans le bon ordre !* »

[36] Une illustration de ce principe a été donnée par B. Gaiffe et L. Danlos (2000) pour l'exemple de la coréférence événementielle.

[37] Ce changement n'affecte pas, d'après nous, l'acceptabilité du dialogue. C'est une configuration qui se trouve dans le corpus examiné, comme par exemple au début du dialogue d'exemple (cf. I_1 à I_3). La seule gêne dans l'exemple modifié pourrait venir de l'accord en genre du pronom.

(73) I_1 : maintenant, il faut que tu prennes un gros cercle
I_2 : pour faire le soleil
I_3 : tu le mets en haut à droite
M_1 : ouais + [geste]
I_4 : voilà
I_5 : maintenant, tu vas prendre un petit carré
I_6 : pour faire une maison
I_7 et le mettre en bas à gauche
M_2 : [geste]
I_8 : maintenant tu en prends un autre
M_3 : avec un toit aussi ?

I_5 introduit une première S-DRS avec un seul constituant. I_6 entretient, selon nous, une relation du type *Result* avec I_5. Comme il ne s'agit pas d'une relation de continuation, nous ne construisons pas de topique commun. Vient ensuite I_7. Intuitivement, nous voudrions considérer I_7 comme *Narration* relié à I_5 (il s'agit clairement de deux événements qui se suivent temporellement). Or, la relation de *Narration* ne peut exister qu'entre deux constituants adjacents. Autrement dit, I_5 n'est plus accessible. Ce problème ne vient pas du découpage ou de la spécificité de notre dialogue. En considérant un exemple comme le texte de narration suivant (74), on retrouve le même problème :

(74) Johannes répara le tracteur. Celui-ci remarchait comme s'il avait été neuf. Ensuite, le brave paysan s'en alla vérifier l'état de santé de ses vaches.

Il semble qu'il est impossible dans la S-DRT de relier par *Narration* des segments non strictement adjacents. La seule possibilité consisterait à remplacer la relation *Result* par une relation d'*Elaboration* demandant la construction d'un topique commun T_1. Si l'on accepte cette solution, le topique commun pourrait être le contenu de I_5 (sur la base de l'hypothèse de nucléarité, cf. 4.3.1), ce qui permettrait ensuite de rattacher I_7 à ce topique par *Narration*. Par conséquent, il faut trouver un topique commun T_2 entre T_1 et I_7. Les indications sur la construction du topique sont minces, mais on pourrait supposer quelque chose comme « mettre le petit carré en bas à gauche ».

Nous arrivons alors au changement de locuteur. Dans un dialogue, il serait imaginable que le changement de tour de parole influe aussi d'une manière ou d'une autre sur le regroupement d'énoncés d'un même locuteur. Dès lors que l'on admet qu'un locuteur prend la parole avec une certaine intention et donc la notion de force illocutoire, cette force devrait être représentée dans la structure dialogique (Grisvard, 2000). Nous ajoutons alors la force illocutoire *Order* au topique précédent. Ceci nous permettra d'attacher l'intervention M_2 à ce topique et d'inférer la relation *OrderExec*. Il reste à décider si cette relation demande la construction d'un topique commun. D'un point de vue interactionnel, il s'agit effectivement d'une unité qui peut être considérée comme autonome. Si nous introduisons un topique commun T_3, il nous faut réfléchir aux informations à mettre dans ce topique. Compte tenu du faible apport informationnel de l'intervention M_2, il semble raisonnable de sauvegarder des informations provenant du topique précédent T_2.

Cette décision se révélera judicieuse lors de l'intégration de I_8. L'interprétation de l'expression référentielle *en...un autre* demande effectivement qu'on ait accès au participant *petit carré* qui est accessible dans le topique T_3. I_8 s'ancre donc sur ce topique, mais quelle pourra être la relation entre ce topique et I_8 ? La structure obtenue à ce stade est celle de la Figure 10 :

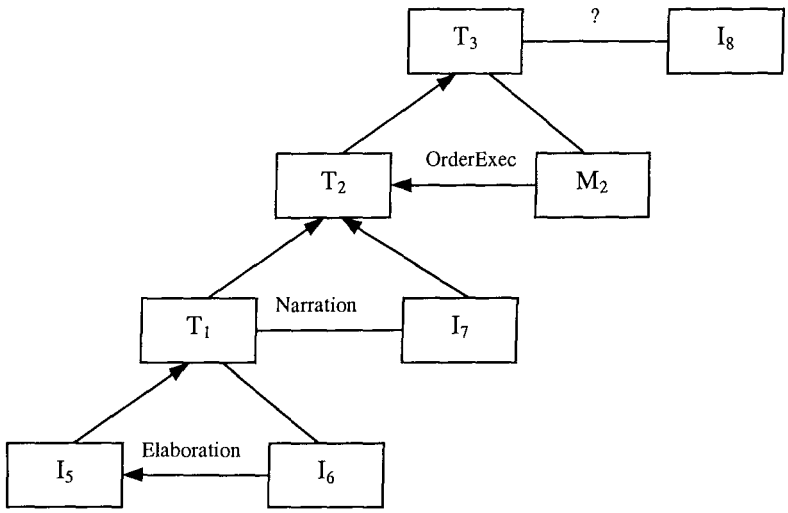


Figure 10 : Traitement de l'exemple (73) à la S-DRT

Il nous faut ensuite traiter M_3. L'interprétation de l'expression référentielle *un toit* pose un nouveau problème car *une maison* en I_6 est absolument inaccessible selon les principes de la SDRT. I_6, qui fournit la clé pour interpréter cette anaphore associative, n'est ni un constituant élaboré, ni le dernier intégré et ne se trouve nullement sur la frontière droite de la structure discursive. Du point de vue de l'enchaînement dialogique, il s'agit en M_3 d'une intervention incidente (Luzzati, 1995) qui ne peut pas être expliquée proprement sans prendre en compte un modèle de la tâche et l'élaboration, par M, d'un plan pour la construction d'une maison – but indiqué par I en I_6. Cependant, cette interprétation va largement au-delà de ce que propose la S-DRT.

En résumé, la S-DRT établit donc une relation entre une représentation du discours à base de relations discursives, l'étude de l'ordonnancement temporel des événements et l'étude de l'accessibilité des référents. Par rapport à nos interrogations, elle propose un contexte arborescent, structuré localement en S-DRS. La construction de ce contexte se fait sur la base de connaissances linguistiques et extralinguistiques, modélisées par des règles et des faits. Même si la possibilité de construction de cette base de connaissances doit être étudiée plus en détail, c'est une avancée par rapport aux autres propositions qui laissent implicite cet aspect. La structuration hiérarchique du contexte devrait permettre de faire des prédictions sur l'accessibilité des référents du discours, mais le problème essentiel est le suivant : cette structuration étant étroitement liée à la notion de topique commun entre deux énoncés, tout dépend de la formulation de ce topique et nous avons vu que ce point, tout comme chez B. Webber (1991), n'est pas suffisamment développé.

4.3.6 Regard critique sur les approches discursives arborescentes

Les approches présentées ci-dessus possèdent toutes un point commun : la proposition d'une structure contextuelle arborescente dans l'objectif de contraindre l'interprétation des énoncés ultérieurs. Si nous en retenons la nécessité de la modélisation hiérarchique du contexte, nous sommes pourtant amenée à nous poser certaines questions.

Notre première interrogation porte sur la mise en œuvre réelle de la modélisation proposée et en particulier sur l'algorithme de construction des représentations hiérarchiques. Indépendamment du problème que pose le choix des relations, il nous semble que ce processus est très difficilement calculable. Tous les auteurs mentionnent des indices linguistiques pouvant éventuellement guider le choix des relations (en particulier, les connecteurs pragmatiques), mais ils soulignent tout aussi

unanimement que ces indices ne sont aucunement obligatoires. Le constat de B. Webber (1991 : 116) « *I will have to assume, [...] the existence of an oracle that can decide with which, if any, existing segment the next utterance in a text shares a common subject, viewpoint etc.* »[38] nous paraît tout à fait symptomatique à cet égard. Les réels critères pour la construction de la structure discursive sont les buts communicatifs plausibles (RST), les intentions basées sur la structuration de la tâche (Grosz et Sidner, 1986) ou l'inférence à partir de connaissances générales (SDRT). Or, ces critères supposent l'amorce d'une interprétation intentionnelle de l'énoncé avant même de pouvoir l'intégrer effectivement dans la structure contextuelle (et donc de pouvoir calculer la référence), ce qui ne nous semble pas facilement automatisable.

Notre deuxième interrogation porte sur le caractère véritablement adapté de structures arborescentes à la modélisation de l'accessibilité des informations du contexte. En plus de la question sur la nature et le contenu des nœuds (cf. la définition imprécise du topique par N. Asher ou les problèmes soulevés à propos de la modélisation proposée par B. Webber), il semble en effet qu'une telle structure – fût-elle calculée par un oracle – n'arrive pas à rendre compte de tous les phénomènes d'accessibilité référentielle. Premièrement, elle risque de prédire l'inaccessibilité d'informations auxquelles il est possible de référer (cf. les problèmes que pose l'interprétation de l'expression référentielle *un toit* de notre exemple). Deuxièmement, des informations accessibles dans la structure contextuelle ne sont pas toujours réellement disponibles pour une reprise. Ainsi, si l'on essaye de traiter à la S-DRT l'exemple (75) :

(75) Helmut mangea du saucisson.
Ensuite, il mangea du pâté.
Enfin, il mangea du jambon.

on devra inférer un topique commun qui résume les trois événements et leurs participants. Ce topique pourrait raisonnablement ressembler à « *Helmut mange de la charcuterie* ». Selon la S-DRT, on est alors en droit de reprendre les référent du discours pour *Helmut* et *la charcuterie*. Pourtant, la reprise pronominale da la variante (a) de l'exemple (76) nous paraît pour le moins curieuse, alors que la possibilité de reprise par *la charcuterie* subsiste :

(76) a. ? *Elle* [la charcuterie] venait du traiteur d'en face.
b. La charcuterie venait du traiteur d'en face.

Il semble donc que l'on ne puisse pas tirer des conclusions sur l'accessibilité sans considérer le type de l'expression référentielle employée (défini vs. pronom) et la manière dont le topique a été construit (explicite vs. inféré).

Enfin, la troisième interrogation porte sur l'adaptation de ces approches discursives au traitement du dialogue. Les tentatives de traitement de notre exemple dialogique à la RST et à la S-DRT ont fait apparaître un problème qui nous semble crucial : les relations discursives sont insuffisantes pour traiter correctement les relations dialogiques, même au niveau le plus simple des paires adjacentes (Sacks et al., 1974). L'introduction *ad hoc* de relations de type *Question-Answer* (R. Scha et L. Polanyi, 1988 ; N. Asher, 1997) pose plus de problèmes qu'elle n'en résout. Ces auteurs traitent en effet au même niveau des relations discursives monologiques et des relations dialogiques. Or, si l'on admet qu'un dialogue se construit par l'enchaînement de parties monologiques énoncés, composées potentiellement de plus d'un énoncé (Roulet et al., 1985), il devient difficile de traiter les relations discursives entre ces énoncés au même niveau que les relations qui lient les différentes interventions (cf. les problèmes rencontrés lors de l'analyse de notre exemple à la S-DRT). R. Scha et L. Polanyi (1988 : 573) remarquent pourtant cette différence de niveaux d'analyse : « *A speaker engaged in a discourse may*

[38] « Je suis amenée à supposer [...] l'existence d'un oracle qui puisse décider, le cas échéant, avec quel segment l'énoncé suivant d'un texte partage un sujet, point de vue etc. commun. »

perform speech acts whose illocutionary force has scope over complex propositions which are built up out of individual sentence meanings and the rhetorical relations between them »[39].

(77) A Pourquoi Helmut est-il si gros ?
B Parce qu'il mange trop de charcuterie !

Par ailleurs, cet ajout de nouvelles relations interpersonnelles introduit des problèmes d'ambiguïtés. Dans l'exemple (77), la relation entre la question et la réponse serait certainement de type *Question-Answer*, mais est-ce que cela veut dire qu'elle ne serait plus de type *Explanation* comme cela serait le cas dans un monologue ? Ou est-ce qu'il peut y avoir plusieurs relations entre des énoncés ? Dans ce cas, il faut vérifier que cette décision ne conduise pas à des incohérences, par exemple en ce qui concerne la nécessité d'un topique commun ou la définition des conditions d'accessibilité. Le problème qui se pose ici nous semble relever d'une confusion entre relations de niveaux différents : les relations sémantiques ou conceptuelles d'une part et les relations pragmatiques (qui rejoignent la notion de force illocutoire) d'autre part.

Ce problème se complique encore dans la mesure où les auteurs ayant travaillé sur les relations discursives à un niveau monologique ont essayé de prendre en compte certaines relations qui nous semblent proches de la notion de force illocutoire, notamment dans ce que E. Hovy et al. (1993) qualifient de relations interpersonnelles. Cette idée se trouve déjà chez W. Mann et S. Thompson (1988) qui proposent une classification de leurs relations en « *subject matter* » (objectif : reconnaître la relation) et en « *presentational* » ou « *pragmatic* » (objectif : augmenter la disposition de l'interlocuteur à accepter ce qui vient d'être dit). Nous nous demandons s'il est judicieux de mélanger ainsi ces deux types de relations. Comme le font remarquer E. Hovy et al. (1993), les relations sémantiques devraient pouvoir se calculer sans prise en compte d'un modèle de l'utilisateur et uniquement sur la base des connaissances encyclopédiques, alors que cela ne semble pas être le cas pour les relations pragmatiques qui, elles, ne peuvent se calculer sans la prise en compte des croyances et intentions des interlocuteurs.

4.4 Synthèse sur les modélisations du contexte

4.4.1 Quelles entités composent le contexte ?

Un première question transversale à adresser aux modèles que nous avons passés en revue est la nature des unités entrant dans le contexte. Qu'il s'agisse de *discourse constructs* (théorie du Centrage), de *discourse referents* (DRT), de *complexes prédicatifs* (théorie des veines) ou de *propositions* (S-DRT), il n'en reste pas moins que ces entités restent de nature discursive, dans le sens où leur introduction dépend d'une mention discursive. Une exception est peut-être donnée par l'espace attentionnel qui est supposé contenir « *contextual information needed to process utterances at each point of the discourse* »[40], sachant qu'il peut s'agir d' « *objects, properties and relations that are most salient [...] and has links to relevant parts of both the linguistic and the intentional structure* »[41] (Grosz et Sidner, 1986 : 182). Cette définition peut recevoir une interprétation très large, laissant penser qu'elle inclut des objets autres qu'introduits linguistiquement, à condition qu'ils soient saillants ou impliqués d'une façon ou d'une autre dans la structure intentionnelle. Mais cela nous ramène aux problèmes liés à la détermination de cette structure.

[39] « Un locuteur engagé dans un discours peut effectuer des actes de langage dont la force illocutoire porte sur des propositions complexes, composées de phrases individuelles et de relations rhétoriques entre celles-ci. »
[40] « de l'information contextuelle nécessaire au traitement des énoncés à n'importe quel moment du discours »
[41] « objets, propriétés et relations qui sont les plus saillants [...] et en relation avec des parties pertinentes des structures linguistique et intentionnelle »

Lier l'introduction des entités contextuelles à des mentions discursives est effectivement nécessaire : cela permet en particulier de résoudre la référence à des entités fictives ou hypothétiques. Ce dernier point est précisément un des points forts de la DRT qui traite d'une façon élégante les restrictions d'accès à des entités introduites sous le champ d'une construction conditionnelle. Mais l'introduction d'une entité contextuelle sur les seuls critères de mention textuelle laisse en suspens la question de l'ancrage des entités discursives dans le monde réel. Or, dans le cadre du DHM, il est nécessaire de faire le lien entre des entités linguistiques et les représentations des objets de l'application. Le calcul de la référence ne peut donc s'arrêter au calcul de la coréférence (c'est-à-dire d'une relation entre deux entités textuelles), mais doit mener à l'identification effective des objets de l'application. Il apparaît alors qu'un contexte composé exclusivement d'entités introduites discursivement est insuffisant. Il néglige en particulier l'environnement perceptif qui fournit, au même titre que le discours précédent, des référent possibles.

L'illustration de la nécessité de l'ancrage perceptif peut se faire à partir des postulats de la DRT elle-même : dans un cadre vériconditionnel, connaître le sens d'un énoncé revient à savoir dans quelles circonstances cet énoncé est vrai. Cependant, une forme logique, telle que construite par la DRT pour un discours, n'est pas toujours suffisante pour tester les conditions de vérité (Gaiffe et al., 2000). Alors que la formule logique correspondant à l'énoncé (78) traduit les conditions de vérité de celui-ci, ce n'est pas le cas pour l'énoncé (79) :

(78) Une femme marche dans la rue.

$\exists e \exists x$ femme(x) $\wedge$ marcher_dans_la rue (e,x)

(79) [deux barres visibles à l'écran...]
alors il va falloir que tu prennes *les deux grandes barres* (C8Route)

a. $\exists e \exists x$ 2_grandes_barres(x) $\wedge$ prendre(e, Toi, x) $\wedge$ e > now

b. $\exists e \exists x$ 2_grandes_barres(x) $\wedge$ prendre(e, Toi, x) $\wedge$ e > now avec <x, B>

Le traitement de (79) par la DRT est le suivant : Le défini *les deux grandes barres* n'a pas d'antécédent discursif. Un référent de discours *2_grandes_barres(x)* est donc accommodé (van der Sandt, 1992). Mais la formule en (a) ne peut pas traduire les conditions de vérité de (79), car il n'y a pas une manière unique de rendre vraie cette formule : rien n'interdit par exemple à l'interlocuteur de supprimer les deux barres visibles à l'écran, d'en créer deux nouvelles, puis de les manipuler. Or, ceci n'est pas dans l'intention du locuteur. Rendre l'énoncé vrai demande donc d'abord d'identifier l'individu en question, c'est-à-dire de résoudre la référence pour *les deux grandes barres* en liant la variable quantifiée *x* à une constante *B*, comme en (b). Cette constante *B* est précisément une entité contextuelle qui n'a pas été introduite discursivement, mais perceptivement (ou, par ce que B. Grosz et C. Sidner considèrent peut-être comme une condition de saillance basée sur l'intention du locuteur).

Cette idée d'ancrage perceptif pourrait être rapprochée de la notion d'ancre de la DRT. Les ancres y ont été introduites pour rendre compte de phénomènes particuliers liés à l'usage des noms propres. A partir du constat qu'un nom propre β n'est jamais utilisé comme synonyme d'un syntagme nominal « *une personne s'appelant* β [42] », l'ancrage externe a été défini comme une fonction qui lie un référent du discours à un individu réel (Kamp et Reyle, 1993 : 248). Calculer la référence pour le dialogue homme-machine reviendrait alors à systématiser l'usage de la fonction d'ancrage, dans la mesure où l'objectif est de déterminer systématiquement la relation entre une expression référentielle et une ancre, cette ancre étant la représentation identifiante d'un objet de l'application. Cela demande en particulier de réfléchir à une intégration de ces ancres dans le modèle contextuel, ce implique que celui-ci ne sera plus constitué uniquement d'éléments introduits par le discours.

[42] Un locuteur visant une telle interprétation utiliserait probablement un indéfini : *"Un « John » a appelé pour toi."*

4.4.2 Comment le contexte est-il structuré en domaines d'accessibilité ?

Il apparaît à travers ce chapitre et les chapitres précédents que des facteurs très variés limitent, sous certaines conditions, l'accès aux entités du contexte. Il s'agit de facteurs :

linguistiques : la DRT, en particulier, étudie en détail les conséquences de certaines constructions syntaxiques sur des reprises : négations, conditionelles, modalisations. La théorie du Centrage s'appuie sur les fonctions syntaxiques (sujet, objet) afin d'ordonner les centres en avant d'un segment donné. Enfin, dans certaines expérimentations psycholinguistiques, des constructions spécifiques (topicalisations et clivées) ont été utilisées pour focaliser une entité ;

cognitifs : nous avons vu au chapitre précédent que l'apport essentiel des approches cognitives consiste en l'établissement d'une relation de dépendance entre statuts cognitifs et formes linguistiques. Les statuts cognitifs se traduisent essentiellement par différents degrés de focalisation ou de familiarité des entités contextuelles. Les facteurs qui déterminent le degré de focalisation d'une entité peuvent être d'origine linguistique (cf. ci-dessus), mais il faut aussi garder en mémoire la notion de récence et la possibilité d'une focalisation ou familiarisation par des critères non linguistiques (présence dans la situation de communication) ;

discursifs : pratiquement toutes les approches présentées dans la section 4.3 ont pour caractéristique commune de construire une structure discursive sur la base de l'agencement thématique ou rhétorique des unités du discours. Cet agencement repose sur des relations diverses (RST), ou des relations de dominance (théorie de veines, B. Webber (1991), S-DRT) seulement ;

intentionnels : la particularité de la proposition de B. Grosz et de C. Sidner (1986) est de baser la structure limitant l'accès aux entités contextuelles sur la structure intentionnelle d'un texte ou d'un dialogue.

L'agencement dynamique des entités contextuelles doit donc avoir pour objectif de modéliser les contraintes d'accessibilité liées à ces facteurs. Dans toutes les modélisations présentées, l'incrémentation du contexte s'opère segment par segment : elle consiste en général en l'ajout de nouvelles entités et en une restructuration éventuelle du contexte. La différence essentielle entre ces approches est liée à la manière dont elles prennent en compte les facteurs mentionnés ci-dessus. En fonction de cela, la structuration du contexte en domaines d'accessibilité sera plus ou moins fine. Nous avons distingué entre des structurations :

globales : la DRT peut être considérée comme une approche « globale », dans la mesure où l'univers du discours met à disposition tous les référents introduits dans le discours. Les seules restrictions d'accessibilité sont imposées sur la base de certaines constructions syntaxiques. Un des points critiques est alors la surabondance de référents disponibles et l'absence de critères pour choisir parmi plusieurs candidats ;

locales : la théorie du Centrage, en revanche, est une approche « locale », puisque l'on reconstruit un nouveau contexte pour chaque phrase. L'ordonnancement des éléments de ce contexte a l'avantage de fournir un critère de choix, lorsque plusieurs référents sont disponibles. Cependant, contrairement à la DRT, les contextes sont trop restrictifs pour tenir compte de reprises pronominales dépassant le cadre de la phrase précédente ;

intermédiaires : à la lumière de la comparaison de la DRT avec la théorie du Centrage, il apparaît que le calcul référentiel repose à la fois sur la modélisation d'un contexte global et des

structurations locales plus fines. Les différentes théories présentées en 4.3 peuvent être considérées ouvrant cette persecptive. La S-DRT, par exemple, élargit l'appareil formel de la DRT de façon à introduire des représentations contextuelles intermédiaires entre le discours et l'énoncé. Sur la base d'axiomes par défaut, dérivés de relations rhétoriques telles que proposées par la RST, elle définit des conditions de regroupement de DRS locales sous un même topique. Les problèmes principaux de toutes les approches reposant sur une structuration discursive arborescente sont le calcul des relations entre les segments, la définition du contenu des nœuds non terminaux (problème du « topique ») et un pouvoir prédictif limité.

A la lumière de cette revue – déjà riche en informations importantes pour la structuration contextuelle – il apparaît pourtant qu'un certain nombre de facteurs n'ont pas été pris en compte. Si toutes les approches essaient de structurer le contexte en fonction d'informations discursives, il n'en est pas de même pour d'autres informations. Là encore, tout comme pour les entités composant le contexte, nous posons le problème de l'environnement perceptif. Celui-ci est en effet susceptible, non seulement d'introduire des entités contextuelles, mais également de fournir des indications pour la structuration contextuelle.

L'exemple (80) montre d'une part que les entités contextuelles peuvent être reliées par des relations extralinguistiques et d'autre part, que la connaissance de ces relations s'avère quelquefois indispensable à l'interprétation des expressions référentielles :

(80) a. Marquez un cercle sur la droite/ la gauche !
b. Marquez *le cube* !

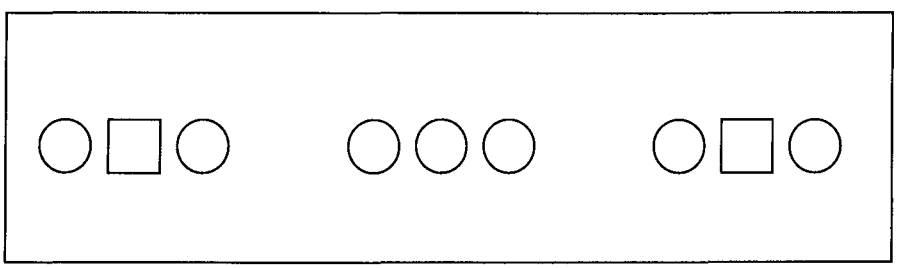


Figure 11 : Scène d'ambiguïté référentielle (Kessler et al.,1996)

Cet exemple est tiré d'une expérimentation psycholinguistique (Kessler et al., 1996) qui a confirmé l'hypothèse selon laquelle le choix du cercle en (a) – droite vs. gauche – est déterminant pour l'identification d'un référent pour *le cube* en (b). Cela signifie que la relation spatiale entre un premier référent (pour *cercle*) et un deuxième référent (pour *cube*) permet de résoudre l'ambiguïté potentielle de l'expression référentielle *le cube* en (b). Or, cette relation est précisément une relation non linguistique, déterminée uniquement à partir de facteurs perceptifs, comme ici la proximité (Wolff et al., 1998).

Une autre source d'informations non linguistiques sont les connaissances conceptuelles, ou, plus largement, les connaissances encyclopédiques. Comme nous l'avons déjà vu à plusieurs reprises, il y a en effet des expressions qui ne réfèrent pas directement à une entité introduite, mais à des entités accessibles à travers des connaissances supplémentaires (conceptuelles ou encyclopédiques) sur les entités mentionnées, comme en (81) :

(81) La saucisse était si délicieuse que Helmut en a même mangé *la peau*.

La question est alors de savoir sous quelle forme et à quel moment on introduit ces informations dans le contexte. La théorie du Centrage, la théorie des veines et celle des arbres discursifs de

B. Webber ne peuvent pas répondre à cette question, étant donné qu'elles se préoccupent essentiellement des pronoms.

La proposition sur les composantes de la structure attentionnelle de B. Grosz et C. Sidner n'exclut pas une prise en compte des informations inférables : ceci d'autant plus que B. Grosz (1981) envisageait déjà dans ses travaux antérieurs sur la représentation du focus un « focus implicite » donnant accès, pour les objets du focus « explicite », aux parties composantes des ces objets. Cette vision est compatible avec la caractérisation très large (« information contextuelle ») des entités composant la structure attentionnelle, mais on ne sait pas exactement, comment ces informations entrent dans le contexte et à quel moment elles sont activées pour une reprise.

En ce qui concerne la S-DRT, A. Lascarides et N. Asher se servent de connaissances pragmatiques afin de déterminer des relations discursives. L'accès à ces bases de connaissances leur permet également de traiter les anaphores associatives : à partir d'une critique des travaux de R. van der Sandt (1992) et de J. Bos et al. (1995) – qu'ils estiment sur-puissants dans la mesure où il n'y a plus de « *presupposition failure* »[43] – ils proposent un traitement reposant sur l'hypothèse d'une interaction entre relations rhétoriques et calcul du contenu sémantique d'une description définie (Asher et Lascarides, 1998).

4.4.3 Quel rôle pour les expressions référentielles ?

Une troisième question transversale concerne le rôle accordé aux expressions référentielles. Il semble en effet qu'un problème commun aux théories présentées consiste en la prise en compte insuffisante de la nature de l'expression sur les opérations de structuration contextuelle. D'une façon un peu schématisée, les opérations proposées consistent soit en l'ajout d'une entité (pour une description indéfinie), soit en la ré-identification d'une entité (pour une expression pronominale). Le traitement des démonstratifs (hors déictiques discursifs, cf. Webber, 1991) n'est pas mentionné. En ce qui concerne les définis, excepté la S-DRT (cf. ci-dessus), leur fonctionnement n'est pas abordé (théorie du Centrage, théorie des veines et arbres discursifs de B. Webber), ou assimilé à celui des indéfinis ou des pronoms (DRT classique).

Certains travaux en sémantique formelle ont pourtant été consacrés spécifiquement au traitement des descriptions définies. Leur traitement repose généralement sur le raisonnement suivant : une description définie comporte la présupposition d'existence de son référent (Karttunen, 1974 ; Heim, 1992 ; Bosch et Geurts, 1990). Or, le traitement des présuppositions est étroitement lié à l'opération d'accommodation (Karttunen, 1974 ; Stalnaker, 1978) qui consiste à ajouter les informations présupposées au contexte si elles n'y figurent pas déjà. Les travaux de R. van der Sandt (1992) reprennent cette vision des choses en précisant que les présuppositions sont des anaphores et impliquent donc un lien avec le contexte. A partir de là, l'auteur introduit un mécanisme dans la DRT qui accommode une variable pour un défini, s'il n'y pas de variable existante dans l'univers de discours. Or, étant donné les mécanismes de la DRT, cette solution revient précisément à ne pas accorder un fonctionnement spécifique aux définis : ils introduisent un nouveau référent du discours, comme les indéfinis, ou ils ré-identifient un référent du discours existant, comme les pronoms.

Les conséquences de ce traitement sont les suivantes : premièrement, comme l'introduction d'un nouveau référent se fait sans la prise en compte de liens avec les objets existants, la théorie est sur-puissante parce qu'elle ne prédit plus de différence entre (82) et (83), alors que (83) semble plus

[43] « échec présuppositionnel »

difficile à interpréter, justement en raison de l'absence d'un lien avec ce qui précède (« *presupposition failure* ») :

(82) I_1 et en dessous de ce soleil on va faire *la petite fillette*
I_2 donc tu vas prendre le petit rond pour faire *la tête* (C8Eglise)

(83) I_1 et en dessous de ce soleil on va faire *la petite fillette*
I_2 donc tu vas prendre le petit rond pour faire *le ballon* (C8Eglise, modifié)

Deuxièmement, lier le déclenchement d'une opération d'accommodation uniquement à des définis (considérés comme des « *presupposition trigger* »[44]) ne permet pas de traiter ce genre d'inférences lors de l'emploi d'autres expressions : comme le font remarquer N. Asher et A. Lascarides (1998), on rencontre des indéfinis pour lesquels il semble se passer des choses comparables :

(84) John voulait se suicider. Il prit *une corde.* (N. Asher et A. Lascarides, 1998)

J. Bos et. al. (1995) proposent alors une solution partielle concernant le premier des deux problèmes, en utilisant les connaissances issues de la structure *Qualia* (Pustejowsky, 1995). Très brièvement, il s'agit d'une structure du lexique qui associe à une entrée lexicale un certain nombre de connaissances lexicales et conceptuelles, par exemple sur les composantes et l'utilisation possible d'un artefact. La résolution anaphorique se fait alors de la façon suivante : étant donné un défini, on essaye d'abord de le lier à une variable de l'univers du discours. Si cette opération ne réussit pas, on essaye de trouver un antécédent parmi les *Qualia* des référents du discours. C'est ainsi que serait par exemple traité (82). Et seulement si cette deuxième opération échoue, on accommode un nouveau référent de discours. Par rapport à la proposition de R. van der Sandt (1992), les auteurs tiennent donc compte de la différence entre (82) et (83), mais n'évitent toujours pas l'impossibilité de traiter les « indéfinis associatifs », ni la sur-puissance d'une accommodation comme dernier recours pour éviter un « échec présuppositionnel ».

Le point commun (et problématique) de toutes ces approches est de dériver le fonctionnement du défini d'une relation prototypique de liage (« *linking* »), ce qui revient à attribuer au défini par défaut le même rôle qu'au pronom. Cette vision est particulièrement explicite dans le travail de J. Bos et. al. (1995) : « *Note that this definition prefers linking to bridging and bridging to accommodation, which we assume is right* »[45]. Contrairement à ces auteurs, nous ne pensons pas que cela soit systématiquement vrai et nous avons donné au chapitre 2 (section 2.3.5) plusieurs arguments linguistiques qui contredisent cette vision des choses. Ces arguments sont, pour mémoire : l'incompatibilité avec le fonctionnement générique des définis, la dissociation fonctionnelle des emplois anaphorique et associatif, l'impossibilité de prédire la distribution du pronom et du défini, l'impossibilité de prédire dans certains cas ambigus la préférence pour une lecture associative et, enfin, l'impossibilité de prédire la non-coréférentialité en cas d'adjonction de modifieurs.

Malheureusement, la tendance à ne pas distinguer le fonctionnement des descriptions définies de celui des pronoms est très répandue, (cf. par exemple Kempson, 1986)[46]. Ceci est particulièrement dommage, car les contraintes classiquement associées aux référents d'une description définie – celle de la présupposition existentielle et celle de l'unicité – demandent justement à s'interpréter dans un cadre référentiel limité contextuellement (cf. entre autres Westerstahl, 1984 ; Groenendijk et al., 1995) et devraient donc constituer un bon moyen de mettre à l'épreuve les modélisations du contexte. Or, les

[44] « déclencheurs présuppositionnels »
[45] « Notez que cette définition préfère le liage à l'association et l'association à l'accommodation. »
[46] « I propose that the concept of definiteness associated both with pronouns and definite NPs is simply that of guaranted accessibility. If a speaker uses a pronoun or a definite NP, then he or she is indicating to the hearer that a representation of a NP type is immediatly accessible in the sense specified – either from the scenario of the utterance itself, from the preceding utterance, from preceding parts of the same utterance or from concepts expressed by what precedes the anaphor. » (Kempson, 1986 : 214).

propositions ci-dessus ne prédisent ni le point commun qu'il pourrait y avoir parmi les divers usages du défini (anaphorique, générique, associatif ou évolutif), ni la différence entre, d'un côté, pronom et défini et de l'autre côté, indéfini et défini.

5 Algorithmes et implémentations

5.1 Introduction

Après la présentation de certaines approches théoriques de la référence, nous terminerons ce parcours par un examen de ce que ces approches ont pu apporter à la mise en œuvre pratique de systèmes de traitement automatique de la langue naturelle. Étant donné le nombre de systèmes et d'algorithmes, nous avons dû faire un choix. Pour cela, nous nous sommes fixé plusieurs critères : d'abord, nous avons tenu à présenter certains travaux anciens, mais fondamentaux (Winograd, 1972 ; Hobbs, 1978 ; Sidner, 1979 ; Brennan et al., 1987). Ces travaux nous servent de référence pour évaluer des implémentations plus récentes. Pour ce faire, nous avons choisi de présenter, lorsque cela a été possible, des systèmes récents (Mitkov, 1998 ; Beaver, 2000 ; Krahmer et Kievit, 2000). Nous avons donné la priorité soit à de systèmes déjà évalués, soit à des systèmes mettant en évidence un problème particulier. Enfin, nous avons essayé d'intégrer quelques systèmes français dans cette revue (Gaiffe, 1992 ; Popescu-Belis, 1999).

Dans un premier temps, nous examinerons les algorithmes sous-jacents au traitement référentiel dans des systèmes de compréhension. Le fait que ces algorithmes soient souvent limités quant à la prédiction de la distribution réelle des marqueurs référentiels nous amène à nous interroger sur leur adaptation à des systèmes de dialogue, où il faudra gérer à la fois la compréhension et la génération d'expressions référentielles. Nous nous intéresserons alors de plus près à des algorithmes de génération. Comparés aux algorithmes d'analyse, leur caractéristique marquante est le rôle important que joue la modélisation du contexte pour prédire, d'une part, la distribution entre pronoms et descriptions définies et d'autre part, les constituants de la description.

Une gestion affinée des ensembles contextuels s'avère également nécessaire pour le traitement de la référence dans des systèmes de dialogue. Ce traitement ne peut plus se limiter à la recherche d'un antécédent textuel, comme c'est le cas pour certains systèmes de compréhension : effectuer une action appropriée suite à la requête de l'utilisateur requiert l'identification réelle de l'objet de l'application sur lequel porte cette requête. Or, cet objet peut avoir été mentionné auparavant, mais il peut aussi être accessible en raison de facteurs non discursifs, comme un geste de désignation, ou plus simplement sa visibilité sur l'écran. Les contextes nécessaires à la modélisation du fonctionnement référentiel dans des dialogues sont alors multiples : connaissances partagées et privées des interlocuteurs, historique du dialogue et environnement perceptif.

Les conclusions que nous tirerons de cette revue seront mitigées : il apparaîtra clairement une séparation des systèmes en deux grands groupes : d'une part, des systèmes robustes, mais « pauvres en connaissances », d'autre part des systèmes à couverture réduite, mais plus proches des modèles théoriques de la référence. Les systèmes de dialogue font généralement partie de ce deuxième groupe, mais nous verrons que des systèmes ont pu être développés et implémentés, malgré les difficultés que pose la modélisation de connaissances sémantiques et pragmatiques.

5.2 Algorithmes pour l'analyse d'expressions référentielles

5.2.1 *Le pronom ou « The latest in anaphora resolution : going robust, knowledge-poor and multilingual »*[47] *(Mitkov, 1998)*

La conception de systèmes de compréhension pour des applications comme la recherche d'information est subordonnée avant tout à l'impératif de robustesse : cela signifie que l'on préfère un résultat faux ou approximatif à l'absence de réponse. Les décisions doivent être calculables facilement à partir d'informations disponibles, sans demander des investissements humains jugés démesurés, comme c'est le cas par exemple pour la modélisation des connaissances conceptuelles ou spécifiques à un domaine. Étant donné ces contraintes, les seules connaissances disponibles sont donc celles que l'on peut obtenir automatiquement avec un seuil de fiabilité acceptable : cela concerne, en l'état actuel des choses, des analyses morphologiques, des analyses syntaxiques partielles et certains types d'analyses sémantiques (désambiguïsation de sens, par exemple). Pour cette raison, les systèmes de résolution pronominale robustes fonctionnent essentiellement sur la base de critères syntaxiques, certains intégrant des heuristiques supplémentaires.

- *« L'algorithme naïf » (Hobbs, 1978)*

Un des premiers algorithmes dédié au calcul des antécédents pour les pronoms de 3^{ième} personne est celui de J. Hobbs (1978). Connu sous l'appellation « algorithme naïf », il est devenu une référence pour évaluer la performance de systèmes plus récents. Il est qualifié de « naïf » en raison de son approche purement syntaxique : la recherche de l'antécédent d'un pronom se fait par le parcours de l'arbre syntaxique, en tenant compte de l'accord grammatical entre pronom et candidats pour l'antécédent. La recherche est effectuée de gauche à droite avec un parcours en largeur d'abord. La contrainte « gauche-droite » traduit le principe de préférence du sujet à l'objet, si l'on considère qu'en anglais, le sujet précède le plus souvent l'objet. La contrainte « en largeur d'abord » traduit la préférence des constituants immédiats aux constituants emboîtés. Lorsqu'aucun antécédent n'est trouvé dans la phrase courante, on parcourt les arbres syntaxiques des phrases précédentes, en commençant par la phrase immédiatement précédente, ce qui traduit la préférence pour des antécédents proches plutôt qu'éloignés.

L'intérêt linguistique de l'algorithme tient aux fait qu'il intègre deux principes issus de la grammaire générative transformationnelle (Chomsky, 1964) et formulés par R. Langacker (1969) : selon le premier principe (principe A), un pronom non-réflexif ne peut pas trouver son antécédent dans la même phrase simple. Cela prédit par exemple l'impossibilité de coréférence entre *Kheops* et *le* en (86) par rapport à la coréférence entre *Kheops* et *se* en (85) :

(85) Il y a Ramses qui n'est encore pas couché.
Kheops *se* sent un peu à l'étroit en ce moment (C12Egypte)

(86) Il y a Ramses qui n'est encore pas couché.
Kheops *le* sent un peu à l'étroit en ce moment (C12Egypte, modifié)

Le deuxième principe (principe B), reformulé par Langacker (1969) postule que l'antécédent d'un pronom (non réfléchi et non réflexif) doit précéder ou c-commander le pronom. La notion de c-commande est définie comme suit : un nœud N_1 commande un nœud N_2 si N_1 ne domine pas N_2 et N_2 ne domine pas N_1 et si le premier nœud branchant dominant N_1 domine aussi N_2. Ce principe prédit par exemple la possibilité de coréférence en (87) et l'impossibilité de coréférence en (88) :

[47] « Le dernier cri de la résolution anaphorique : robuste, pauvre en connaissances et multilingue ».

(87) After he robbed the bank, John left the town. (Hobbs, 1978)

(88) *Jean le* lave.

L'algorithme « naïf » de J. Hobbs (1978) intègre donc des contraintes syntaxiques sur la coréférence pronominale intraphrastique et traite les anaphores interphrastiques selon une préférence dérivée de la linéarité du discours (ordre des phrases et ordre des syntagmes nominaux). L'algorithme a été évalué sur 300 pronoms non déictiques, issus de trois types de textes différents (technique, journalistique, littéraire) : le taux de succès est de 88,3%. Par ailleurs, J. Hobbs observe que 90% des pronoms trouvent leur antécédent dans la même phrase et 98% dans la même phrase ou la phrase précédente.

- *« RAP – Resolution of Anaphora Procedure » (Lappin et Leass, 1994)*

Au vu des performances de son algorithme, J.Hobbs annonce lui-même le défi pour les propositions à venir : « *Computationally speaking, it will be a long time before a semantically based algorithm is sophisticated enough to perform as well, and these results set very high standards for any other approach to aim for.* » [48](Grosz et al., 1990 : 345). S. Lappin et H. Leass (1994) ont relevé ce défi avec un algorithme baptisé « RAP » (*Resolution of Anaphora Procedure*). Là aussi, il s'agit d'un algorithme identifiant les antécédents des pronoms de 3^{ième} personne. Il repose sur une mesure de saillance à partir de la structure syntaxique et d'un modèle attentionnel, mais sans faire appel à des connaissances complexes. Les auteurs annoncent une performance qui dépasse de 4% celle de l'algorithme de J. Hobbs.

Les composantes de l'algorithme « RAP » sont les suivantes : un filtre syntaxique pour exclure la coréférence intraphrastique, selon le même principe que chez J. Hobbs (cf. ci-dessus) ; un outil morphologique pour tester l'accord en nombre, genre et personne ; un outil d'identification, par comparaison avec des constructions typiques, de pronoms impersonnels (« *Il est possible que...* ») ; un outil pour la détection, sur critères lexicaux, des pronoms réflexifs ou réciproques (« *Il se lave lui-même* ») ; une composante de calcul de saillance des groupes nominaux à partir d'un seuil initial qui varie en fonction de différents facteurs syntaxiques (Tableau 9) ; une composante de calcul de saillance pour des classes d'équivalence (ensemble des groupes nominaux coréférant à une même entité) par addition des scores de saillance de tous les éléments de la classe.

Type du Facteur	***Seuil initial***
Récence phrastique	100
Emphase sur le sujet	80
Emphase existentielle	70
Accusatif	50
Objet indirect	40
Tête d'un NP	80
Emphase non adverbiale	50

Tableau 9 : Scores de saillance en fonction de critères syntaxiques (Lappin et Leass, 1994)

L'algorithme proprement dit est le suivant :

[48] « D'un point de vue computationnel, il faudra attendre longtemps avant qu'un algorithme reposant sur des critères sémantiques ne soit aussi performant et ces résultats posent des standards très élevés pour toutes les autres approches. »

```
1. créer une liste pour tous les NP de la phrase courante
2. pour chaque NP :
      classification selon le type (indéfini, défini, pronom pléonastique,
      pronom personnel, pronom réflexif, indéfini…)
      si NP = {indéfini | défini} alors
            calcul de la saillance
      sinon si NP = pronom réflexif alors
            calculer une liste de paires {NP ; pronom} pour lesquelles il
            peut y avoir coréférence : si plusieurs possibilités,
            l'antécédent est choisi sur le critère de saillance des NP
          sinon
            si NP = pronom personnel de 3ième personne alors
                  calculer une liste de paires {NP ; pronom} pour
                  lesquelles il ne peut pas y avoir co-référence
                  créer la liste des antécédents possibles : celle-ci
                  contient le référent discursif le plus récent pour
                  chaque classe d'équivalence avec son facteur de
                  saillance
                  modifications locales de la saillance : si le NP
                  antécédent se trouve après le pronom, la saillance
                  décroît (pénalisation des cataphores). Si le NP
                  antécédent a le même rôle grammatical que le pronom, sa
                  saillance augmente
                  définir un seuil de saillance et filtrage
                  appliquer le filtre morphologique
                  si plusieurs candidats alors
                    choix sur la saillance du NP antécédent
                     si égalité de saillance alors
                          choix selon proximité
                     fsi
                  fsi
            fsi
          fsi
      fsi
fin pour chaque
```

Figure 12 : Algorithme « RAP » (Lappin et Leass, 1994)

Les principales caractéristiques de cet algorithme sont de bons filtres syntaxiques et morphologiques (plus fins que ceux de J. Hobbs), une prise en compte des pronoms pléonastiques, un calcul de la saillance basé sur d'autres critères syntaxiques que la seule linéarité, le choix sur des critères de proximité en cas d'ambiguïté sur la saillance, la préférence des anaphores intraphrastiques sur les anaphores interphrastiques et la préférence des anaphores sur les cataphores. Mais tout comme l'algorithme de J. Hobbs, « RAP » n'exige pas des connaissances sémantiques ou pragmatiques. L'évaluation sur 354 phrases extraites d'un texte technique a donné un taux de succès de 86% pour un total de 360 pronoms. Ce taux dépasse de 4% celui obtenu par l'algorithme de J. Hobbs sur les mêmes données (cf. Tableau 10). Cette performance est essentiellement due à un meilleur traitement des anaphores intraphrastiques qui constituent, selon les auteurs, 80% des pronoms du corpus. La moindre performance au niveau interphrastique a son origine, selon les auteurs, en l'absence de prise en compte des enchâssements syntaxiques (principale vs. subordonnée) pour le calcul de la saillance, ce qui fournit des antécédents non pertinents.

	Total	***Interphrastique***	***Intraphrastique***
Occurrences pronoms	360	70	290
Cas corrects (RAP)	310 (86%)	52 (74%)	258 (89%)
Cas corrects (Hobbs)	295 (82%)	61 (87%)	234 (81%)

Tableau 10 : Comparaison des algorithmes de Hobbs (1978) et de Lappin et Leass (1994)

- *Les travaux de R. Mitkov (1998)*

Le score de « RAP » (86%) est encore amélioré par R. Mitkov (1998). Avec un algorithme qui ne demande même plus d'analyse syntaxique complète, l'auteur dit obtenir, toujours sur des textes techniques, un taux de succès de 89%. L'algorithme s'applique sur un texte passé à l'analyse morphologique et à l'analyse syntaxique de surface (« *shallow parsing* »). Afin de trouver l'antécédent d'un pronom, il établit la liste des syntagmes nominaux (NP) des deux phrases précédentes, exclut ceux qui ne s'accordent pas en nombre ou genre, attribue des scores aux NP restants en appliquant des règles de préférence et sélectionne comme antécédent le NP ayant le score le plus élevé. Les règles de préférence portent sur différents indicateurs (Tableau 11) supposés favoriser (par exemple, l'occurrence dans la phrase précédente) ou défavoriser (par exemple, NP l'indéfini) le fonctionnement comme antécédent à un pronom.

Indicateur	***Valeur***	***Score***
Définitude	Indéfini	-1
Information connue	Premier NP de la phrase	1
Verbes	NP suivant certains verbes (*discuss, illustrate, examine,...*)	1
Répétition	NP répété deux fois dans le même paragraphe	2
	NP répété une fois dans le même paragraphe	1
Titre	NP inclus dans le titre de la section	1
Phrase prépositionnelle	NP tête d'une PP	-1
Collocation	NP combiné au même verbe	2
Parallélisme	NP dans une phrase précédente avec connecteurs (*and, or, after,...*)	1
Distance	Phrase précédente immédiate	2
	Décrémentation (-1) par phrase	
Terminologie	Terme du domaine technique	1

Tableau 11 : Indicateurs de saillance (Mitkov, 1998)

L'algorithme a été testé sur deux textes techniques comportant au total 294 pronoms, dont 190 étaient non anaphoriques ou déictiques (nous supposons donc qu'ils étaient exclus de l'analyse, même si ce n'est pas précisé par l'auteur) et le taux de résolution correcte se situe à 89%. Par ailleurs, des tests du même algorithme sur le polonais et l'arabe ont donné des résultats légèrement meilleurs : 93% et 95% respectivement.

L'approche prônée par R. Mitkov, si on peut la considérer comme caricaturale, est pour autant symptomatique des modélisations « pauvres en connaissances » : s'il y a effectivement, à court terme, un besoin en systèmes opérationnels, baser un modèle qui se veut scientifique sur un faisceau d'heuristiques dont l'auteur ne prend même plus la peine de donner ne-serait-ce que le début d'une justification théorique (Pourquoi un indéfini a-t-il moins de chances d'être l'antécédent d'un pronom ? D'où vient la liste des verbes et des connecteurs ? Comment justifier la pondération des scores ? Pourquoi la limitation aux deux phrases précédentes ?) comporte un double risque pratique et théorique : d'un point de vue pratique, de tels systèmes sont difficilement maintenables et peu évolutifs, car ils fournissent toujours une réponse, ne signalent jamais de problèmes et sont incapables de fournir des indications sur la source éventuelle des problèmes. D'un point de vue théorique, ils ne valident aucune hypothèse scientifique et ne contribuent en rien à l'explication des processus linguistiques et cognitifs sous-jacents à la résolution des anaphores. Un des symptômes révélateurs est précisément l'exclusion des usages pronominaux non anaphoriques et déictiques (bien que R. Mitkov ne précise pas ce qu'il entend par là...) : le fait que plus de la moitié des pronoms du corpus semblent

échapper à l'algorithme met pourtant profondément en cause soit l'unicité de la catégorie linguistique du pronom personnel de 3^{ième} personne, soit la validité de la modélisation proposée.

5.2.2 *La compréhension pronominale à base de connaissances plus élaborées*

Parallèlement aux approches présentées ci-dessus, un certain nombre de systèmes et d'algorithmes s'efforcent d'intégrer des connaissances autres que purement syntaxiques. Mais contrairement aux algorithmes « pauvres en connaissances », les propositions s'arrêtent souvent à la spécification des architectures et/ou algorithmes, ce qui rend plus difficile leur évaluation numérique.

Parmi les approches « à base de connaissances », on peut distinguer entre les « approches mixtes » et les « approches discursives ». Les premières essayent d'intégrer les apports de différentes théories en proposant des mesures multi-critères venant de la syntaxe, de la sémantique et du discours. Nous renvoyons ici aux travaux de E. Rich et S. Luperfoy (1988). Les secondes, présentées plus en détail, considèrent que les facteurs primaires pour trouver les antécédents des pronoms sont la cohérence discursive et des critères de focalisation. Nous retrouvons dans ce courant différentes versions algorithmiques de la théorie du Centrage (Brennan et al., 1987) ainsi qu'une reformulation récente de cette théorie en termes d'optimalité (Beaver, 2000).

- *Versions algorithmiques de la théorie du Centrage : l'algorithme « BFP »* [49]

Contrairement aux « approches mixtes », les « approches discursives » mettent l'accent sur le rôle de la cohérence discursive et les critères de focalisation pour calculer l'antécédent d'un pronom. Nous nous concentrons ici sur deux versions algorithmiques de la théorie du Centrage : la première, l'algorithme « BFP », a été proposée par S. Brennan et al. (1987) et a fait l'objet de différentes modifications, en particulier par M. Walker et al. (1994). La seconde est une reformulation de la théorie du Centrage dans le cadre de la théorie d'optimalité, proposée récemment par D. Beaver (2000).

L'algorithme « BFP » utilise les concepts de base de la théorie du Centrage, définis dans la section 4.2.1 du chapitre précédent. Pour rappel, il s'agit, pour une phrase donnée U_i, du centre en arrière CAr_i et de la liste ordonnée des centres en avant CAv_i, avec un centre préféré CP_i. L'algorithme reprend également la distinction des différents types de transitions. En ce qui concerne la transition de type *changement*, M. Walker et al. (1994) introduisent en plus une distinction entre *changement majeur* et *changement mineur*. Le Tableau 12 résume les transitions en fonction de l'évolution des centres entre deux phrases U_i et U_{i+1}. L'algorithme « BFP » redéfinit l'ordonnancement des transitions ainsi : *continuation* > *rétention* > *changement mineur* > *changement majeur*. La préférence est donc donnée à la stabilité thématique qui se traduit par une transition de type continuation.

	$CAr_i = CAr_{i+1}$ *(ou pas de* CAr_i*)*	$CAr_i \neq CAr_{i+1}$
$CAr_{i+1} = CP_{i+1}$	continuation	changement mineur
$CAr_{i+1} \neq CP_{i+1}$	rétention	changement majeur

Tableau 12 : Définition des transitions par l'algorithme « BFP »

[49] « BFP » correspond aux initiales des auteurs.

L'algorithme proprement dit est le suivant :

1. Pour deux phrases données, calculer toutes les combinaisons CAr_i / CAv_{i+1} qui respectent l'accord en nombre et genre.
2. Filtrer les combinaisons obtenues
   par des contraintes sur les anaphores intraphrastiques (cf. les deux principes issus de la grammaire générative et utilisés par l'algorithme de Hobbs, 1978),
   par des contraintes sur les restrictions sélectionnelles (compatibilité sémantique)
   par la règle de centrage sur la pronominalisation : « Si un élément de la liste des centres en avant de la phrase U_i est pronominalisé dans la phrase U_{i+1}, alors le centre en arrière de U_{i+1} doit être pronominalisé. »
3. Classifier les solutions restantes selon les types de transition.
4. Choisir selon les préférences sur les types de transition.

Figure 13 : Algorithme « BFP » (Brennan et al., 1987)

Pour l'exemple (89), cet algorithme prédit correctement la résolution de *he* et *him* en (d) sur *Terry* et *Tony*, respectivement. Cette solution correspond effectivement à une *continuation*, alors que l'envers résulterait en une *rétention*. Il prend aussi en compte le caractère curieux de (e), puisque résoudre *he* sur *Tony* donne lieu à un *changement mineur*, alors qu'une résolution de *he* sur *Terry* résulterait en une *continuation* et serait donc préférable.

(89) a. Terry really goofs sometimes.
b. Yesterday was a beautiful day and *he* was exciting about trying out his new sailboat.
c. He wanted Tony to join him on a sailing expedition.
d. *He* called *him* at 6 a.m.
e. *He* was sick and furious at being woken up so early. (Kehler, 1997)

L'algorithme a pourtant sollicité un certain nombre de critiques, dont deux sont clairement formulées par A. Kehler (1997). La première concerne l'objectif initial de la théorie du Centrage, qui semble vouloir modéliser une partie des processus cognitifs du traitement du discours. Or, en dépit de cette ambition, l'algorithme « BFP » ne prédit pas systématiquement la préférence d'un locuteur pour la résolution pronominale, comme le montre l'exemple (90) :

(90) a. Terry really gets angry sometimes.
b. Yesterday was a beautiful day and he was exciting about trying out his new sailboat.
c. He wanted Tony to join him, and left him a message on his answering machine.
d. Tony called him at 6 a.m. the next morning.
e_1. *He* was furious with him at being woken up so early.
e_2. *He* was furious at being woken up so early. (Kehler, 1997)

En ce qui concerne (e_1), l'algorithme « BFP » prédit la résolution de *he* sur *Tony*. La prédiction est correcte et résulte du fait que cette solution implique comme transition un changement mineur, alors que l'inverse (*he* = *Terry*) impliquerait un changement majeur. Mais en (e_2), *he* est résolu sur *Terry* (relation de continuation), car une résolution sur *Tony* donnerait lieu à un changement mineur. Or, cette prédiction est mise en cause par une majorité de locuteurs, qui préfèrent, ici comme en (e_1), une interprétation par co-spécification du sujet de (d). Le problème est alors l'absence du deuxième pronom qui provoque un lien entre les deux formes pronominalisées, plutôt qu'une co-spécification du sujet de (d). D'un point de vue théorique, ces différentes prédictions basées sur la seule présence ou absence d'un autre pronom sont difficilement justifiables et ne semblent pas toujours tenir compte des préférences réelles des locuteurs.

La deuxième critique avancée par A. Kehler concerne l'insuffisance des seuls critères liés à la fonction syntaxique et la réalisation des groupes nominaux. Il y a en effet d'autres facteurs qui jouent sur la résolution pronominale, comme par exemple les relations discursives (Asher, 1993). L'exemple (91) illustre cette possibilité :

(91) a. The three candidates had a debate today.
b. Bob Dole begun by bashing Bill Clinton.
c. He criticized him on his opposition to tobacco.
d_1. Then Ross Perot reminded *him* that most Americans are also anti-tobacco.
d_2. Then Ross Perot slammed *him* on his tax policies. (Kehler, 1997)

La théorie du Centrage prédit aussi bien pour (d_1) que pour (d_2) la résolution de *him* sur *Bob Dole*. Mais, comme le remarque A. Kehler (1997), toutes conditions égales par ailleurs, le parallélisme discursif en (d_2) est à l'origine d'une résolution du pronom sur *Bill Clinton* et la théorie du Centrage seule n'arrive pas à prendre en compte ce fait.

Une troisième critique – ne s'adressant pas au fond théorique, mais plutôt à la formulation de la théorie et de ses variantes algorithmiques – a été à l'origine de la proposition de D. Beaver (2000). Il reproche à la théorie du Centrage d'être conceptuellement difficile, en raison d'une trop grande diversité des règles et contraintes comprenant des règles de Centrage, des filtres, des schémas de transition et des contraintes sur l'ordonnancement des transitions. Aussi bien dans la théorie que dans les algorithmes, le tout est exprimé de façon procédurale, ce qui rend plus difficile le contrôle des différents paramètres de la résolution pronominale. La proposition de D. Beaver consiste alors en une expression déclarative de la théorie du Centrage, permettant à la fois une vision plus claire de l'interaction des contraintes et une réversibilité facile pour appliquer les principes à la génération automatique d'expressions référentielles.

- *Théorie du Centrage et théorie de l'optimalité (Beaver, 2000)*

Premièrement, D. Beaver redéfinit les différentes règles et contraintes de la théorie du Centrage par un ensemble de contraintes non monotones :

AGREE : le pronom et l'antécédent s'accordent en nombre et genre ;

DISJOINT : les arguments d'un même prédicat sont référentiellement disjoints ;

PROTOP : le topique est pronominalisé ;

COHERE : le topique d'une phrase est le topique de la phrase précédente ;

ALIGN : le topique d'une phrase est dans la position sujet.

Ensuite, il utilise la méthode des tableaux de la théorie d'optimalité (Prince et Smolensky, 1993). Appliquée à la résolution pronominale, celle-ci consiste à générer toutes les possibilités combinatoires entre antécédent et pronom et de choisir comme solution, parmi ces possibilités, la variante qui respecte le maximum de contraintes. Le Tableau 13 illustre cette méthode pour (c) de l'exemple (92) : la colonne de gauche propose les différentes combinatoires, les colonnes suivantes correspondent chacune à une contrainte ; une étoile signifie la violation de la contrainte pour la combinatoire donnée. Étant donné le nombre d'étoiles dans chaque ligne, on retient comme solution la première combinatoire (associant *she* à *Jane* et *the young women* à *Mary*.

(92) a. Jane$_i$ likes Mary$_j$.
b. She$_i$ often goes around for tea with her$_j$.
c. She$_k$ chats with the young women$_l$ for ages. (Beaver, 2000)

Exemple (92)c	*AGREE*	*DISJOINT*	*PROTOP*	*COHERE*	*ALIGN*
$k = i, l = j$					
$k = j, l = i$			*		*
$k = i, l = i$		*			

Tableau 13 : Tableau d'optimalité pour (c) de l'exemple (92)

L'avantage de cette reformulation est de faire les mêmes prédictions que la théorie du Centrage, tout en basant la procédure de résolution sur un ensemble de contraintes explicites et en rendant les transitions épiphénoménales. Les paramètres, formulés par des contraintes non-monotones, peuvent être contrôlés et modifiés facilement et de façon explicite. Selon D. Beaver, « BPF » compare uniquement des phrases, mais ne dit plus rien sur la façon dont la structure discursive est utilisée pour faciliter la compréhension. Cette reformulation rétablit certaines des intuitions originales, par exemple le fait qu'un changement de topique est signalé généralement par la position sujet du nouveau topique. Une illustration est donnée par (d) et (e_2) de l'exemple (90): contrairement à l'algorithme « BFP », on prédit ici correctement l'interprétation effective, car pour toutes choses égales par ailleurs, il y a violation de la contrainte COHERE si *he* est résolu sur *Terry*. Par ailleurs, comme le souligne D. Beaver, il est possible d'améliorer facilement les performances de l'algorithme par une pondération des contraintes. Enfin, comme les contraintes sont formulées de façon déclarative, la méthode n'est pas centrée spécifiquement sur l'interprétation du discours et peut s'adapter à la génération. Il est d'ailleurs dans les ambitions de l'auteur d'en faire un premier pas vers un programme plus général intégrant des théories sur l'anaphore, des formes linguistiques et la cohérence du discours. Il rejoint par là les préoccupations des modélisations algorithmiques consacrées non plus uniquement à l'analyse du pronom personnel, mais plus largement à la compréhension de différents types d'expressions référentielles : nous en présenterons deux dans la section suivante.

5.2.3 La compréhension au-delà du pronom

- *Modélisation du focus pour la compréhension d'expressions anaphoriques*

C. Sidner (1979, 1983) propose un des premiers algorithmes utilisant la notion de *focus* pour la résolution des anaphores définies. Elle considère comme anaphores définies les pronoms personnels, les groupes nominaux définis et les déictiques du discours *that/this*. L'idée de base qu'elle défend est une relation entre continuité thématique et anaphorisation. Elle fait donc l'hypothèse que s'il y a anaphore, celle-ci fait préférentiellement référence à l'élément thématisé qui est dans le focus. Dans le cas contraire, l'anaphorisation signale un changement de thème. D'un point de vue algorithmique, cela l'amène donc à la gestion d'une pile de « thèmes-antécédents-focus ». L'objectif de l'algorithme est alors d'identifier, pour une expression définie donnée, la représentation de l'antécédent en utilisant des critères syntaxiques, sémantiques et inférentiels. C. Sidner se distingue clairement des approches présentées ci-dessus par deux points : d'une part, elle cherche un traitement unifié pour différents types d'expressions et d'autre part, elle se détache de la tendance qui consiste à rechercher un antécédent textuel au profit d'une recherche d'un antécédent « mémoriel ».

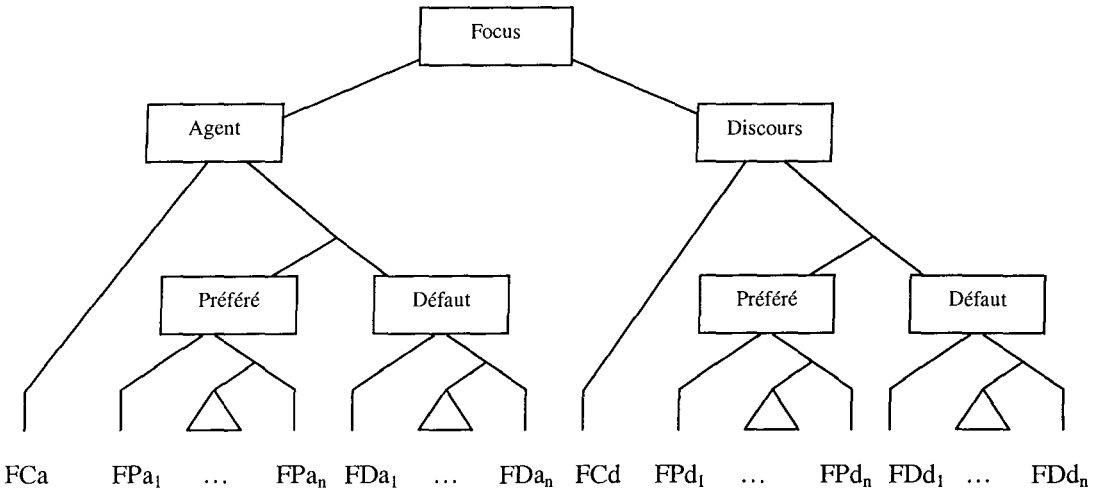


Figure 14 : Structure des registres focaux de C. Sidner (1979)

La partie centrale du modèle proposé est la représentation du focus sous forme d'un ensemble des registres focaux (Figure 14). Comme nous l'avons déjà souligné, les entités entrant dans cette structure ne sont pas directement les antécédents textuels, mais des modèles mentaux ou structures cognitives, incluant en particulier des liens associatifs, par exemple vers les participants et les conditions spatio-temporelles d'un événement, ou la hiérarchie de types pour un objet. La première distinction dans la structure se fait entre le focus discursif et le focus des agents. C. Sidner justifie cette distinction par la nécessité de gérer à part l'évolution thématique des agents et des autres participants des événements. Pour chacune de ces sous-structures, le focus se compose du focus courant (FC) et d'une liste de foci alternatifs, elle-même composée d'une liste de foci par défaut (FD) – issus du discours précédent – et d'une liste de foci préférentiels (FP) – issus de la dernière phrase. Les éléments de ces listes sont ordonnés selon des critères liés aux θ-rôles, cette intuition étant comparable au parcours de gauche à droite de l'algorithme de J. Hobbs (1978) et à l'ordonnancement des centres de la théorie du Centrage.

```
1.  initialiser le focus du premier énoncé par un algorithme particulier
2.  pour chaque phrase P :
       si P contient une anaphore alors
          tester la compatibilité avec le focus courant
          si compatibilité alors
                choisir le focus courant comme antécédent
          sinon chercher un antécédent parmi les foci préférentiels
                si compatibilité alors
                   choisir le focus comme antécédent et l'empiler comme focus
                   courant sur l'ancien focus
                sinon
                   chercher dans la pile des foci par défaut un antécédent
                   qui     convient,  le  promouvoir  comme  focus  courant  et
                   dépiler tous les foci au-dessus dans la pile
                fsi
          fsi
       fsi
    fin pour chaque
```

Figure 15 : Algorithmes pour la résolution des anaphores définies (Sidner, 1979)

L'algorithme proposé s'appuie essentiellement sur la supposition d'un ordre préférentiel de parcours de cette structure pour la résolution d'anaphores : le parcours se décide en fonction du rôle de l'anaphore (agent vs. autre) et se fait dans l'ordre *focus courant* > *focus préféré* > *focus par défaut.*

Pour chaque référent hypothétique retourné, une procédure de test décide si celui-ci convient ou non, en fonction de l'accord grammatical, des contraintes sémantiques et des informations inférentielles.

L'algorithme pour la résolution des définis et pronoms non agentifs est celui de la (Figure 15). Les critiques principales de cet algorithme (cf. par exemple Gaiffe, 1992) concernent trois points. D'abord, l'unification du traitement pour toutes les anaphores définies est une bonne idée, mais la façon dont elle est opérée est doublement critiquable : d'un côté, l'algorithme n'est pas unifié, car il sépare systématiquement le traitement des anaphores « agents » des autres anaphores. D'un autre côté, il ne tient plus compte de la spécificité des différents marqueurs référentiels (pronom vs. défini), dont nous avons vu pourtant au chapitre 2 qu'elle est irréductible. Le deuxième point est lié au précédent et concerne une des manifestations de la différence entre définis et pronoms lors d'une coordination comme dans « *un carré et un triangle* ». Comme C. Sidner ne gère pas la formation de tels ensembles, l'algorithme ne pourra pas prédire qu'une expression pronominale comme « *il* » ne pourra pas faire référence à un des éléments de cet ensemble. Enfin, le troisième point concerne la fusion opérée entre thème, focus et référent. En effet, le thème d'un discours n'est pas toujours un référent : il peut s'agir d'un ensemble de référents, comme pour l'exemple précédent.

L'intérêt essentiel de cet algorithme réside dans la gestion d'une structure contextuelle basée sur des facteurs autres que purement syntaxiques et dans une modélisation des référents par des représentations cognitives, ce qui permet la prise en compte de connaissances inférentielles. Par la suite, nous comparerons cet algorithme à un système plus récent qui est également supposé reposer sur des représentations mentales (Popescu-Belis, 1999).

- *« Un atelier pour le traitement de la référence » (A. Popescu-Belis, 1999)*

L'objectif des travaux de A. Popescu-Belis et de ses collègues (1998) est la mise au point d'un ensemble d'outils opérationnels, robustes et évaluables pour le traitement automatique de la référence. Les auteurs proposent un système basé sur une version simplifiée d'un modèle représentationnel, considérant que les phénomènes référentiels concernent non pas des entités textuelles, mais des représentations cognitives des référents. L'outil central du système est un résolveur de référence qui construit pour chaque référent du texte une *représentation mentale*, notion qui est empruntée au projet CERVICAL (Reboul et al., 1998)[50]. A. Popescu-Belis (1999) considère qu'il s'agit d'une structure qui contient l'ensemble des expressions référentielles se rapportant à un même référent. Une représentation mentale peut être construite pour toute chose à laquelle un locuteur peut référer : un objet concret, un personnage, ou encore un événement.

Pour une expression référentielle donnée, la tâche du résolveur consiste à déterminer si celle-ci réfère à une représentation mentale déjà introduite ou au contraire, si elle introduit une entité dont il n'avait pas encore été question et pour laquelle il faut introduire une nouvelle représentation. La décision sur ce point se prend par l'application d'un ensemble de règles. Celles-ci permettent d'éliminer les représentations auxquelles l'expression ne pourra pas référer. Pour en choisir une parmi celles qui restent, on examine leur taux d'activation selon une méthode proche de celle employée dans l'algorithme de S. Lappin et H. Leass (1994). S'il n'y a pas de représentation compatible ayant un taux d'activation suffisant, on procède à la création d'une nouvelle représentation.

[50] Nous reviendrons plus en détail sur ce projet (cf. le chapitre 7).

L'algorithme peut être résumé comme suit :

```
Pour une expression E donnée :
1.  Déterminer un ensemble de représentations qui pourraient être co-
    spécifiées E en effectuant des tests sur
        l'accord en genre et en nombre,
        la compatibilité sémantique (synonymie, hyponymie, hyperonymie)
        sachant qu'une représentation est promue compatible ssi plus de x %
        des expressions attachées à cette représentation sont compatibles avec
        l'expression à interpréter.
2.  Parmi ces représentation compatibles, alors choisir celle qui est la plus
    active
        calcul de l'activation à partir d'un taux d'activation initial de 15
        réactivation en cas de co-spécification en fonction du type de
        l'expression utilisée : nom propre = 40, nom commun = 20 , pronom = 10
        baisse d'activation avec avancement dans le temps (par mot, phrase,
        paragraphe, unité thématique)
        taille maximale de la mémoire des représentations = 15,
        taille maximale de la liste des expressions attachées à une
        représentation = 10
3.  Si une représentation a pu être identifiée, attachement de l'expression à
    celle-ci
4.  Si l'ensemble des représentations compatibles est vide ou aucune
    représentation n'est suffisamment activée :
        création d'une nouvelle représentation
5.  Mise à jour des taux d'activation
```

Figure 16 : Algorithme du résolveur de référence (Popescu-Belis, 1999)

Un des avantages de cette proposition est qu'elle a été implémentée et testée sur des textes littéraires. Les résultats obtenus sont comparables à ceux d'autres systèmes. Les auteurs font pourtant remarquer qu'avec peu de ressources, ils obtiennent de meilleurs résultats que des systèmes purement coréférentiels ou sans traitement de l'activation. L'avantage par rapport aux systèmes purement coréférentiels s'explique par le fait de calculer la compatibilité non pas par rapport à un seul antécédent, mais par rapport à l'ensemble des antécédents regroupés sous une même représentation. En ce qui concerne le traitement de l'activation, il a pour effet d'éliminer certains antécédents compatibles, mais improbables en raison d'une activation insuffisante. Selon la conclusion des auteurs, il s'agit d'un système robuste qui demande à être enrichi avec des connaissances plus élaborées : « *Pour passer d'un traitement de la référence à une compréhension de la référence, le programme doit être capable d'un ancrage perceptif de ses représentations qui porteront sur des entités de son environnement* » (A. Popescu-Belis, 1999 : 244).

Cette dernière citation montre déjà que malgré ce qu'affirme A. Popescu-Belis, nous ne pensons pas que l'implémentation proposée maintienne les caractéristiques essentielles de la théorie de représentations mentales. Nous verrons dans la suite (en particulier au chapitre 7) qu'un des objectifs de cette théorie est précisément de poser un cadre pour l'ancrage perceptif des référents. Or, la réduction de la représentation mentale à la liste des expressions référentielles utilisées est finalement plus proche d'une approche textuelle coréférentielle que d'une approche mémorielle ou cognitive de la référence. Contrairement à ce qu'envisage C. Sidner, la proposition de A. Popescu-Belis ne fait pas apparaître de quelle manière il serait possible d'intégrer des connaissances conceptuelles autres que celles liées strictement à une hiérarchie de types. Or, ce sont ces connaissances qui permettraient par exemple de traiter les anaphores associatives, jusqu'alors non prises en compte.

Un autre point nous paraît également délicat : l'algorithme proposé ne fait pas apparaître si et comment le traitement dépend du type de l'expression à interpréter. Il laisse penser qu'il ne tient pas

compte de la spécificité des différentes expressions. D'abord, nous supposons que le fonctionnement est le même pour les pronoms et les descriptions définies, ce qui pose les mêmes problèmes que pour l'algorithme de C. Sidner. Ensuite, rien n'est dit à propos des démonstratifs : en particulier, nous nous interrogeons sur la manière dont le test de compatibilité basé sur l'accord grammatical et la compatibilité sémantique peut passer le cap des reclassifications (« *un arbre – cette sentinelle* ») dont nous avons vu au chapitre 2 (section 2.4) qu'elles représentent des emplois typiques pour les démonstratifs. Enfin, nous craignons que l'absence de prise en compte de la détermination mène à un traitement coréférentiel de la suite « *un triangle – un triangle* ».

5.3 Algorithmes pour la génération automatique d'expressions référentielles

5.3.1 Génération de descriptions définies « distinctives »

Une des questions centrales en génération automatique d'expressions référentielles (Dale, 1992 ; Dale et Reiter, 1995, 1996) peut être formulée ainsi : étant donné un ensemble de référents E, comment générer une description D_e qui identifie de façon unique une entité e. Les concepts utilisés pour répondre à cette question sont les suivants :

E	l'ensemble contextuel des entités saillantes
e	le référent
$C_e = E\text{-}\{e\}$	l'ensemble des distracteurs
$P_e = \{p \mid p(e)\}$	les propriétés qui sont vraies de e
$D_c \subseteq P_e$	une description distinctive pour e

Une description qui distingue de façon unique le référent e est alors une description qui vérifie les propriétés suivantes :

$\forall p \in D_e : p(e)$	toutes les propriétés mentionnées par D_e sont vraies de e
$\forall e_i \in C_e : \exists p_j \in D_e \ \neg p_j(e_i)$	pour toutes les entités e_i de l'ensemble contextuel, la description contient une propriété qui ne soit pas vraie de e_i.

Pour donner un exemple, dans un contexte composé de trois entités :

e_1	{chien, petit, noir}
e_2	{chien, grand, blanc}
e_3	{chat, petit, noir}

les valeurs des variables sont les suivantes :

E	$\{e_1, e_2, e_3\}$
e	$\{e_1\}$
C_e	$\{e_2, e_3\}$
P_e	{chien, petit, noir}.

Une description distinctive D_e pour e doit alors contenir l'ensemble des propriétés {chien, noir} ou {chien, petit}.

A partir d'une telle formulation du problème, R. Dale et E. Reiter (Reiter et Dale, 1992 ; Dale et Reiter, 1995, 1996) ont proposé plusieurs algorithmes pour la génération de descriptions définies distinctives. Ces algorithmes sont résumés et comparés dans le Tableau 14 :

Algorithme	*« Full Brevity »*	*« Greedy Heuristics »*	*« Incremental Algorithm »*
Description	teste si une description à une seule propriété peut être générée. Si ce n'est pas possible, teste avec toutes les combinaisons de deux propriétés etc...	choisit une propriété qui élimine un nombre maximal de distracteurs et réduit de la même façon successivement l'ensemble des distracteurs résiduels	itération à travers la liste des propriétés de e et ajout à D_e de la propriété courante si celle-ci élimine au moins un des distracteurs
Complexité	NP-complexe	polynomial	polynomial
Avantages	génère une description minimale	rapidité	rapidité
Inconvénients	complexité	ne génère pas la description la plus courte	ne génère pas la description la plus courte

Tableau 14 : Comparaison des algorithmes pour la génération de descriptions définies

Ces algorithmes, tous destinés à générer des descriptions définies appropriées en fonction d'un contexte donné, soulèvent un certain nombre de questions. D'abord, ils partent tous d'une liste de propriétés à plat. Or, il semble que toutes les propriétés n'ont pas le même statut : lorsqu'il y un choix entre plusieurs propriétés, certaines seraient préférables d'un point de vue cognitif. Cette hypothèse est approfondie par R. Dale et E. Reiter (1995) sur la base d'expériences psycholinguistiques qui font apparaître qu'une description minimale n'est pas toujours adaptée. Dans certains cas, il est préférable de générer une description plus longue, mais contenant des propriétés plus facilement identifiables. Le deuxième problème est l'absence d'une prise en compte des relations entre objets (« *la bouteille sur la table* »). Les difficultés sont ici liées à la récursivité de l'emboîtement de la description (« *la bouteille sur la table sur laquelle se trouve la bouteille...* »). Elles ont été étudiées en détail par R. Dale et al. (1991). Enfin, ces algorithmes présupposent une procédure capable de décider si la génération d'une description définie est la solution la plus appropriée dans le contexte donné.

5.3.2 Choix entre description définie et pronom personnel

Le choix entre une description définie ou une pronominalisation est en effet une deuxième question centrale de la génération automatique d'expressions référentielles. Elle a été abordée entre autres par L. Danlos (1992), R. Dale (1992) et par E. Reiter (Dale et Reiter, 1996). Sur la base des maximes conversationnelles (Grice, 1979) – et en particulier celle de la quantité (« *Dites tout ce qu'il faut et seulement ce qu'il faut !* ») – les auteurs considèrent que

> une pronominalisation est appropriée lorsque le référent est focalisé (marqué par le trait « + *center* » , au sens de la théorie du Centrage) ou s'il est le dernier élément introduit dans le discours, à condition toutefois que le pronom ne soit pas ambigu ;
>
> une description définie est appropriée dans les autres cas de référence anaphorique et cette description doit être distinctive, c'est-à-dire isoler le référent visé de tous les autres éléments du contexte saillant.

K. McCoy et M. Strube (1999) ont mis en évidence encore d'autres facteurs que le seul facteur focal et la non-ambiguïté pour décider entre pronominalisation et description définie. Ils proposent en particulier de tenir compte des frontières phrastiques et de la structuration temporelle du discours. Par ailleurs, ils relâchent, suite à une étude de corpus, la condition de non-ambiguïté qu'ils considèrent comme trop contraignante. L'algorithme final a été implémenté et évalué sur un corpus du *New York Times* : il s'est avéré correct dans 84,7% des cas.

Ce qui nous intéresse particulièrement dans les approches consacrées à la génération est l'idée d'une description définie distinctive dans un contexte donné. Nous y retrouvons un point de vue sur le fonctionnement des descriptions définies qui est très proche des considérations linguistiques présentées au chapitre 2 (section 2.3). Les algorithmes pour la génération prédisent en effet qu'une description définie demande, pour être appropriée, d'isoler un élément dans un contexte plus large. Ainsi, en (93) :

(93) Dessine un triangle. Colorie [?]...

les algorithmes génèrent un pronom pour référer au triangle introduit, alors qu'ils génèrent une description définie en (94) :

(94) Dessine un triangle et un carré. Colorie [?]...

Il convient de noter ici qu'aucun des algorithmes pour l'analyse des expressions référentielles ne permet de faire ces prédictions, alors que celles-ci coïncident fortement avec l'hypothèse sur le fonctionnement linguistique des descriptions définies.

(95) a. Dessine un triangle. Colorie *le triangle.*
b. Dessine un triangle et un carré. Colorie *le triangle.*

Pour les algorithmes d'analyse, les variantes (a) et (b) de (95) présentent en effet le même taux d'acceptabilité. Cet exemple illustre notre inquiétude sur la réversibilité des algorithmes d'analyse pour la génération et soulève par conséquent une fois de plus la question de leur pertinence cognitive.

5.3.3 Dynamisation du contexte de génération

Aussi bien les algorithmes sur la génération de descriptions définies que les procédures de décision entre pronominalisation et génération d'un défini laissent de côté un problème de taille : la détermination et l'évolution dynamique de l'ensemble contextuel. R. Dale et E. Reiter précisent seulement que « *the context set [is] the set of entities that the hearer is currently assumed to be attending to ; this is similar to the set of entities in the focus spaces of the discourse focus stack in Grosz and Sidners theory of discourse structure* »[51] (1995 : 236). En pratique, il semble que le contexte est donné une fois pour toutes et qu'il n'évolue pas (Dale, 1992).

E. Krahmer et M. Theune (1999) proposent alors de structurer dynamiquement le contexte par l'intégration de la notion de saillance à l'algorithme incrémental R. Dale et E. Reiter (1995) : « *It is not feasible to dynamically reduce or enlarge the context set. Rather, it should be a structured whole, containing precisely the objects in the domain and combined with a method to mark certain objects as more prominent than others.* »[52] (Krahmer et Theune, 1999 : 5). La modification essentielle qu'ils apportent à l'algorithme incrémental est l'attribution d'un taux de saillance à chaque entité contextuelle. Après chaque itération – correspondant à la sélection d'une propriété – ils effectuent un

[51] « l'ensemble contextuel est un ensemble d'entités avec lequel l'interlocuteur est supposé être familier ; il est similaire aux espaces focaux de la pile des foci discursifs dans la théorie de la structure discursive de Grosz et Sidner. »
[52] « Il n'est pas possible de réduire ou élargir dynamiquement l'ensemble contextuel. Il devrait être plutôt un tout structuré, contenant précisément les objets du domaine et combiné avec une méthode pour marquer certains objets en tant que plus proéminents que d'autres. »

test supplémentaire pour décider si l'entité à désigner est l'entité la plus saillante possédant les propriétés de la description générée. Si cela est le cas, ils retiennent la description.

Selon les auteurs, une telle gestion dynamique de la saillance associée aux entités contextuelles a plusieurs avantages : elle permet de générer des descriptions définies plus appropriées au contexte, dans la mesure où dans des contextes réduits, certaines informations peuvent être omises. Il est par exemple possible d'utiliser moins de propriétés ou une description plus proche du niveau sémantique de base (Rosch, 1978). Par ailleurs, elle permet d'améliorer la décision sur la pronominalisation et d'intégrer des indications sur la génération d'accents contrastifs lors d'une phase éventuelle de synthèse de la parole.

La conséquence majeure de la modification est donc une reformulation des conditions appropriées pour la génération d'une description définie : contrairement aux algorithmes initiaux, une description définie n'isole plus l'objet unique, mais l'objet le plus saillant du contexte possédant les propriétés de la description.

Il est intéressant de noter que l'application du principe de saillance au fonctionnement des descriptions définies soulève quelques questions : si le besoin de domaines d'interprétation plus fins que le contexte global est incontestable, utiliser la seule saillance pour structurer ce domaine global ne semble pas non plus résoudre tous les problèmes. En particulier, il efface de nouveau les spécificités du fonctionnement des marqueurs définis et pronominaux, comme le montre l'exemple (96) :

(96) a. Le triangle bleu se trouve à gauche du triangle rouge.
b. *Il* est grand...
c. *Le triangle* est grand...

Si l'on adopte un calcul de saillance suivant les principes de la théorie du Centrage, *le triangle bleu* est clairement l'élément le plus saillant du contexte après l'interprétation de (a). Cela est d'ailleurs confirmé par l'anaphore pronominale *il* en (b) qui s'interprète sans problème. L'interprétation de l'anaphore définie *le triangle* en (c), en revanche, semble beaucoup moins aisée. Or, cette différence ne peut plus être prédite par l'approche étendant la contrainte de saillance au calcul référentiel des définis. Nous pensons donc que le fonctionnement des définis – l'isolation, par le contenu descriptif, d'un élément à l'intérieur d'un domaine – ne doit pas être remplacé par un fonctionnement expliqué en termes de saillance. Cela correspond d'ailleurs aux hypothèses issues d'expériences psycholinguistiques présentées dans le chapitre 0. Pour prendre en compte des phénomènes de quantification limitée à un sous-ensemble du contexte global, la solution ne passerait donc pas par un ordonnancement des entités contextuelles selon leur saillance, comme cela a été proposé par E. Krahmer et M. Theune (1999), mais par la formation dynamique de sous-ensembles à l'intérieur desquels le fonctionnement des descriptions définies est tout à fait régulier. C'est une modélisation allant dans ce sens que nous allons proposer dans cette thèse (chapitres 6 à 9).

5.4 Traitement de la référence dans quelques systèmes de dialogue

5.4.1 SHRDLU, Que ta voie ne meure...[53]

Le système *SHRDLU*, développé au début des années 1970 par T. Winograd (Winograd, 1972), a été un des premiers systèmes à simuler un comportement de compréhension dans un monde limité à la manipulation de blocs multiformes et multicolores. L'objectif était effectivement de reproduire une partie du comportement humain, mais sans chercher à en comprendre réellement le fonctionnement

[53] D. Hofstadter (1985). Gödel, Escher, Bach – Les Brins d'une Guirlande Éternelle. InterEditions, Paris, p. 657.

cognitif. Nous commençons notre revue du traitement de la référence dans quelques systèmes dialogiques par celui-ci et ceci pour deux raisons : premièrement, toutes les composantes du système, y compris celle qui calcule la référence, sont documentées de façon exhaustive par l'auteur (Winograd, 1972). Deuxièmement, cela nous fournit une base de comparaison pour des systèmes plus récents.

Dans le système *SHRDLU*, le traitement de la référence (comprenant l'analyse et la génération) se fait par l'application d'un ensemble d'heuristiques : en ce qui concerne les expressions anaphoriques, il traite les définis et les pronoms. Un défini est supposé référer à un objet particulier et unique. Si l'analyse d'une description laisse plusieurs solutions ouvertes, différentes heuristiques basées sur la récence de la dernière mention sont appliquées avant qu'il y ait échec. Le système traite également les quantifieurs (*all, some, every, at most, at least, three of the N*), les ordinaux et les superlatives. Quant aux pronoms personnels, la stratégie est la suivante : s'il y a un autre pronom précédent dans la même phrase ou la phrase précédente, on retourne le même référent. Si ce n'est pas le cas, on cherche un antécédent parmi les entités introduites dans la phrase précédente. En cas d'antécédents multiples, on choisit l'antécédent en fonction de la structure syntaxique, stratégie déjà proche de ce que proposent des approches plus récentes sur la compréhension du pronom. Par ailleurs, il y a la possibilité d'utiliser des pronoms pour faire référence à des événements : le dernier événement peut être désigné par *it*, l'avant-dernier par *that*. Enfin, en ce concerne les ellipses (*one-anaphora* de l'anglais), le traitement repose sur des heuristiques de contraste : la restitution de la tête nominale se fait par la recherche du dernier groupe nominal qui contraste avec la propriété exprimée dans l'anaphore :

(97) a. The green block supports the *big* pyramid but not the *little* one.
b. The *green* block supports the big pyramid but not the *red* one.

Ainsi, pour (97), l'ellipse en (a) est résolue sur *pyramid* (en raison du contraste entre *little* et *big*), alors que celle de (b) est résolue sur *block* (en raison du contraste entre *red* et *green*).

En ce qui concerne l'évaluation du système, quelques dialogues entre un utilisateur expert et le système sont transcrits dans Winograd (1972). S'ils peuvent sembler impressionnants au premier abord, une analyse approfondie fait apparaître rapidement leurs limites. Les remarques critiques concernent essentiellement deux points : d'une part l'absence d'une gestion contextuelle suffisante et d'autre part la méthodologie qui consiste à accumuler des heuristiques.

Concernant la gestion du contexte, celle-ci est limitée aux entités de la phrase précédente, ce qui rend impossible des références pronominales à des objets introduits auparavant. Par ailleurs, il n'y a pas d'opérations de groupement sur les entités contextuelles, ce qui rend impossible des références pronominales plurielles à des entités introduites séparément. Enfin, l'univers est considéré comme statique : la nature des objets ne change pas et le point de vue du locuteur sur les objets reste constant. Or, cet aspect est particulièrement important pour des systèmes de dialogue de commande. Dans le cas d'un simple logiciel de dessin, la composition progressive des figures géométriques va donner lieu à des perspectives différentes (d'un carré et d'un triangle, on passera par exemple à une maison). Ce type de comportement n'est pas prévu par *SHRDLU* : ce système travaille avec une liste de propriétés fixes et toute désignation sortant de ce cadre est considérée comme « inutile » et donc négligée.

Ce point nous amène à une critique concernant l'accumulation des heuristiques. Cette accumulation a deux inconvénients. D'un point de vue pratique, les règles sont difficilement extensibles, car il faudra à chaque fois contrôler les interactions éventuelles. D'un point de vue théorique, elles ne cherchent pas à modéliser le comportement humain ou même, plus modestement, l'usage complémentaire des différents types d'expressions référentielles. Comme la majorité des algorithmes de compréhension, le système est limité à choisir entre une description définie et un pronom, mais n'intègre pas les particularités linguistiques qui caractérisent l'emploi de ces marqueurs. Nous l'avons

déjà vu pour l'emploi très restrictif des pronoms et nous pourrions ajouter à cette liste l'absence d'un traitement approprié des anaphores associatives ou évolutives, de la référence à des objets inexistants, ou, plus important encore dans un système de dialogue, des démonstratifs.

5.4.2 DenK System (Kievit et Piwek, 2000)

L'intérêt du *DenK System,* développé par L. Kievit et P. Piwek (2000), réside dans la résolution d'expressions référentielles dans un contexte dialogique et multimodal. La modélisation de l'environnement dialogique comprend la modélisation du domaine, des connaissances de l'utilisateur et des connaissances du système. Le domaine est défini comme une partie du monde réel : il s'agit ici de la manipulation d'un microscope. Le système, implémenté en Prolog, est conçu pour comprendre une entrée multimodale – de la langue naturelle et des gestes de désignation effectués avec la souris – et répond à des questions ou effectue des actions. Il est en particulier capable de combiner des informations de différents canaux de communication pour calculer le référent d'une expression définie ambiguë ou pour résoudre un pronom sur une entité disponible dans le contexte visuel.

Le processus de résolution des références multimodales est modélisé sur la base de quatre hypothèses : premièrement, l'information pour la résolution référentielle est multimodale dans la mesure où elle vient de différentes sources (langue et gestes). Deuxièmement, si plusieurs objets sont disponibles, on choisit le plus saillant. Troisièmement, les conditions exactes de choix et de saillance dépendent du type de l'expression à interpréter. Enfin, en cas d'échec, on envisage une réaction coopérative du système.

Sur la base de ces hypothèses, les auteurs proposent un schéma algorithmique général et unifié pour le calcul référentiel. Pour chaque type d'expression, ils instancient une ou plusieurs stratégies de recherche consistant à

1. sélectionner un ensemble de référents candidats ;
2. appliquer des filtres ;
3. ordonner les référents retenus selon des critères de saillance ;
4. évaluer le résultat : échec si aucun ou plusieurs objets.

En ce qui concerne les descriptions définies, les itérations de cette stratégie portent sur différents types de contextes : la recherche commence dans un contexte « hypothétique », spécifique au traitement des conditionnelles. Ensuite, elle s'effectue dans les contextes fournis par les connaissances communes, avec des filtres sur le type des objets et contrainte par un seuil de saillance minimale pour éviter de retenir des objets non pertinents. La saillance elle-même peut venir d'une position récente dans l'historique du dialogue ou d'une disponibilité immédiate dans l'environnement visuel. Enfin, la recherche est élargie à toutes les entités du domaine, toujours contrainte par des filtres sur le type et la saillance de l'objet recherché.

L'algorithme pour les pronoms commence, lui aussi, par la recherche dans des contextes hypothétiques. Ensuite, la recherche est élargie au contexte fourni par l'historique du dialogue et plus précisément aux dix derniers objets mentionnés. Enfin, si ces recherches ne donnent pas de résultats, le référent est accommodé à partir de connaissances sur le domaine. Les filtres spécifiques pour les pronoms portent sur des contraintes grammaticales d'accord et sur le type de l'objet : cette dernière contrainte assure que les actions demandées puissent s'exécuter correctement.

Ce système est intéressant à plusieurs titres. D'abord, il tente une réelle prise en compte des particularités liées à l'interaction dialogique par la modélisation de connaissances partagées et privées. Ensuite, il tient compte du caractère multimodal du dialogue homme-machine en intégrant des facteurs de saillance perceptive, venant du contexte visuel et des gestes de pointage. Enfin, il se distingue des systèmes précédents par le souci d'unification du modèle de résolution référentielle. Même s'il semble prendre en compte seulement deux types d'expressions référentielles (le défini et le pronom), il propose, à travers le schéma algorithmique général, un cadre évolutif permettant d'intégrer d'autres mécanismes de résolution. En ce qui concerne ce schéma, il a l'avantage de s'adapter au parcours de différents contextes, construits en fonction de connaissances statiques – modélisation du domaine – et dynamiques – connaissances partagées, historique du dialogue (Luzzati, 1995)

Un point problématique pourrait néanmoins être la considération de l'historique dialogique comme un seul contexte. Nous avons déjà rencontré ce problème lors de l'examen des algorithmes de génération : la seule structuration du contexte en une pile d'entités ordonnées selon leur saillance mène, d'un point de vue théorique, à une fusion du fonctionnement des descriptions définies avec celui des pronoms et risque d'amener à des prédictions erronées lors de la phase de génération.

5.4.3 MultiDial

Nous verrons par la suite un dernier système – la plate-forme dialogique *MultiDial* – dont la particularité est précisément de proposer des structurations contextuelles plus fines : cela passe par la modélisation de zones temporelles d'une part, et par la création dynamique de sous-ensembles de référents d'autre part.

Le traitement référentiel dans *MultiDial* s'appuie sur une modélisation contextuelle qui représente les objets de l'application et leur états successifs par des schémas temporels (Romary, 1989). Ces schémas sont conçus sur la base de deux relations : l'adjacence et l'inclusion (Figure 17a). La relation d'adjacence (représentée par une flèche pointillée) lie deux états successifs d'un objet, comme par exemple la coloration d'un objet de type *fenêtre*). La relation d'inclusion (représentée par une flèche pleine) traduit un état d'un objet, comme le fait d'être *vert* à un moment donné pour un objet de type *fenêtre*. Ces schémas temporels pour les objets sont eux-mêmes regroupés dans des ensembles, créés pour les participants d'un même événement. Le contexte est alors une pile d'ensembles de schémas temporels, ordonnés selon la récence de leur activation.

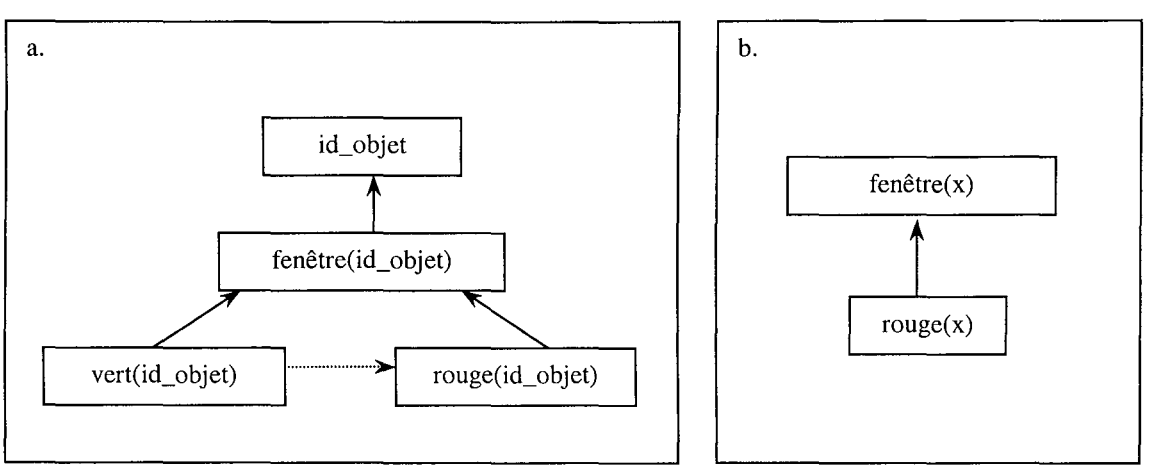


Figure 17 : Représentation des objets par des schémas temporels (Romary, 1989)

Le processus de résolution de la référence proprement dit s'inspire des travaux de B. Gaiffe (Gaiffe, 1992). Premièrement, on construit un schéma temporel hypothétique à partir de l'expression à interpréter (cf. Figure 17b pour *la fenêtre rouge*). Deuxièmement, on compare ce schéma avec la frontière droite des schémas temporels du premier élément de la pile contextuelle et l'instancie sur le premier candidat compatible.

L'originalité du système *MultiDial* par rapport aux systèmes présentés jusqu'ici consiste en la gestion d'un contexte structuré localement et ceci doublement : d'abord, la formation d'ensembles pour les participants d'un même événement est une possibilité pour créer des domaines de quantification locaux. Cela permet par exemple de maintenir le fonctionnement linguistique des descriptions définies comme extracteurs d'une entité d'un ensemble et d'éviter le recours à une fonction de saillance pour les désambiguïser (cf. la section 5.3.3).

Ensuite, la structuration des objets en zones temporelles pose une base intéressante pour la modélisation de l'évolution temporelle des référents et donc pour une prise en compte de la référence évolutive. Elle rend par exemple possible le traitement d'expressions telles que *la fenêtre qui était verte*, ou de prédire les (im-)possibilités de reprise de *les pommes* en (98) :

(98) a. Écrase *les pommes*. Mets-*les* dans un bol.
b. ? Écrase *les pommes*. Mets *les pommes* dans un bol.

D'autres travaux (Vivier et al., 1997) ont souligné, à propos d'une tâche semblable à celle du corpus Ozkan, qu'un nombre important de reformulations d'expressions référentielles (*le rectangle – la porte de la maison*) est précisément liée aux changements de perspective sur les objets manipulés. Les auteurs font remarquer que « *tout se passe comme si l'objet changeait de nature en même temps qu'il change de fonction. [...] Une fois mis en place, l'objet passe d'un statut d'élément indépendant à celui de partie d'ensemble en cours de construction, ce qui était « rectangle » devient « porte de la maison ». En fait, les sujets modifient leur manière de catégoriser en même temps qu'ils transforment l'objet.* » (Vivier et al., 1997 : 278). Cette observation, qui concerne un des aspects non pris en compte par *SHRDLU* (Winograd, 1972), nous semble très importante pour la modélisation du calcul référentiel dans un système de dialogue. Elle signifie en particulier que la modélisation du contexte devrait être capable de refléter non seulement l'évolution de la situation de communication, mais aussi les représentations qu'en ont les interlocuteurs. Nous avons déjà vu lors de la présentation de *SHRDLU* que ce point était particulièrement important pour des applications pilotées par un dialogue de commande, visant par définition à transformer l'état du monde.

L'élargissement des mécanismes de création et de gestion de sous-ensembles contextuels pourra permettre d'étendre le traitement référentiel aux extractions d'une entité d'un ensemble (*Fermer une des fenêtres.*), aux parcours d'ensembles par des expressions ordinales ou d'altérité (*l'autre fenêtre, le premier triangle*) et à un traitement des ellipses (*l'autre, le premier, le vert*) qui ne soit plus basé uniquement sur des heuristiques. Par ailleurs, l'intégration de connaissances pragmatiques sur des ensembles « virtuels » pourra mener à un traitement des anaphores associatives (*une maison – le toit*).

La modélisation que nous proposerons dans les chapitres suivants (chapitres 6 à 9) ira dans ce sens. Elle reprendra en particulier l'idée de la nécessité de sous-ensembles contextuels structurés, tout en généralisant les facteurs de structuration, jusqu'alors exclusivement temporels ou liés à la participation à un même événement.

5.5 Synthèse sur les implémentations

Le constat suivant de A. Popescu-Belis (1999 : 201) résume bien la caractéristique essentielle des algorithmes et systèmes robustes pour la traitement automatique d'expressions référentielles : « *Pratiquement tous les systèmes évalués pour la résolution de la référence dans les textes étaient fondés sur des heuristiques nettement moins élaborées que les modèles computationnels* ». En effet, par rapport aux modèles linguistiques, cognitifs et computationnels discutés dans les chapitres précédents, il se dégage nettement deux tendances :

Ou bien les algorithmes et systèmes sont robustes, implémentés et évaluables, mais essentiellement basés sur une combinaison de différentes heuristiques. Cela concerne en général les systèmes de compréhension intégrés dans des outils de recherche d'information, de compréhension de textes techniques ou d'indexation. Leur avantage est de s'adapter facilement à une grande variétés de textes et de domaines, mais ils ont comme point faible un intérêt scientifique limité, dans la mesure où ils cherchent à simuler un comportement intelligent sans chercher à comprendre réellement les processus cognitifs de la résolution référentielle. Ces systèmes ne sont généralement pas conçus sur la base d'hypothèses particulières sur les mécanismes cognitifs impliqués. Ce fait a pour conséquence de ne plus prédire la distribution effective des différents types d'expressions référentielles. Le développement d'algorithmes et de systèmes spécifiquement consacrés à la résolution du pronom personnel de troisième personne est symptomatique à cet égard. Or, ne pas prédire la distribution des marqueurs référentiel signifie que ces systèmes sont le plus souvent inadaptés à la génération d'expressions référentielles. Leur implémentation dans un système de dialogue demanderait donc de gérer le module de génération à part, ce qui est coûteux d'un point de vue pratique et vraisemblablement inadapté d'un point de vue cognitif ;

Ou bien les algorithmes et systèmes essaient d'intégrer des apports théoriques issus d'une ou de plusieurs modélisations, mais ne sont alors plus adaptés à une couverture large, parce qu'ils demandent en général une modélisation coûteuse de connaissances « de haut niveau », c'est-à-dire de connaissances sémantiques et/ou pragmatiques. C'est dans la majorité des cas la solution adoptée pour les systèmes de dialogues : cela s'explique par le fait que des systèmes de dialogue exigent de toute façon, sous une forme ou une autre, la modélisation du domaine d'application, car l'exécution d'actions est impossible sans une réelle identification des objets sur lesquels porte l'action. Or, la modélisation du domaine d'application est déjà une forme de modélisation contextuelle, indispensable à la gestion de la résolution référentielle basée sur un modèle plus élaboré qu'un ensemble d'heuristiques. Par ailleurs, la modélisation du contexte de l'application est aussi une condition nécessaire pour la phase de génération de réponses en langue naturelle, indispensable à tout système méritant réellement le qualificatif « *dialogique* ». Les algorithmes de génération montrent effectivement que la génération d'expressions référentielles ne se fait pas sans la prise en compte d'un ensemble d'entités contextuelles duquel le référent intentionné doit être isolé.

Enfin, malgré la difficulté que pose la modélisation de connaissances « de haut niveau », retenons que certains systèmes récents (DenKSystem, MultiDial) ont pu être implémentés dans des applications réelles. Notons aussi que ces systèmes prennent tous en compte, d'une manière ou d'une autre, la gestion d'ensembles contextuels, formés sur la base de différentes informations. La modélisation que nous proposons dans la suite s'inscrit dans cette perspective : nous défendrons l'hypothèse qu'il est possible de concevoir un modèle de la référence implémentable tout en respectant le fonctionnement linguistique des marqueurs référentiels, à condition d'étendre et de généraliser la gestion d'ensembles contextuels que nous appellerons « *domaines de référence* ».

6 Les domaines de référence : hypothèse et arguments

6.1 Introduction

L'objectif que nous avons défini pour ce travail est l'élaboration d'un modèle de la référence, qui soit unifié pour tous les types d'expressions référentielles, qui tienne compte des informations discursives, perceptives et gestuelles et qui soit, si possible, prédictif aussi bien pour l'analyse que pour la génération d'expressions référentielles. Nous avons vu à travers les chapitres précédents que les approches linguistiques sont celles qui répondent le mieux aux critères que nous nous sommes fixé. Elles permettent de faire les prédictions les plus fines sur différents emplois des expressions référentielles, tout en veillant à dériver ces variations d'usage de principes de fonctionnement plus généraux. Ces approches fournissent donc une base intéressante pour une modélisation de la référence qui prenne en compte tout type de marqueur référentiel et qui soit prédictive en analyse, mais aussi en génération d'expressions référentielles. En revanche, elles ne renseignent pas directement sur une modélisation appropriée du contexte et ne sont pas immédiatement opérationnelles.

Le présent chapitre est plus particulièrement consacré au premier point : étant donné les travaux sur le fonctionnement linguistique des différentes expressions référentielles (présentés au chapitre 2), nous nous posons comme objectif d'en dériver les contraintes sur une modélisation contextuelle compatible et appropriée. Sur la base d'un principe supposé général – toute acte de référence consiste à isoler un référent à l'intérieur d'un cadre contextuel donné –, nous défendrons l'hypothèse d'une modélisation du contexte en termes de **domaines de référence** : il s'agit d'ensembles contextuels locaux structurés de façon à prédire la distribution des différents marqueurs référentiels.

Après une présentation de cette hypothèse, nous donnerons des arguments qui la renforcent. Ces arguments reviennent, pour certains, sur des points problématiques mis en évidence dans le chapitre précédent. D'autres arguments viennent de travaux que nous n'avons pas inclus dans le chapitre précédent, en raison de leur caractère « marginal » par rapport aux modélisations « standards » de la référence. Ils permettent néanmoins de donner une ouverture à la problématique référentielle, par exemple sur la grammaire cognitive (Langacker, 1991) ou sur des problèmes de la référence spatiale (Schang, 1997) et fournissent ainsi une assise plus large à la notion de domaine de référence. A partir des points abordés lors de cette argumentation, nous terminerons ce chapitre par la formulation d'un ensemble de spécifications pour la modélisation des domaines de référence.

6.2 L'acte référentiel est relatif à un cadre contextuel : le domaine de référence

Le problème central de notre travail est le suivant : Comment construire et mettre à jour un modèle du contexte qui soit au maximum compatible avec le fonctionnement linguistique des différents marqueurs référentiels tel que nous l'avons présenté dans au chapitre 2 ? En réponse, nous proposons une modélisation du contexte sous forme de domaines de référence. Il s'agit d'ensembles contextuels qui regroupent et structurent les référents selon des critères permettant de prédire l'emploi des différentes expressions référentielles.

Notre proposition repose sur l'hypothèse selon laquelle tout acte de référence passe par l'identification d'un domaine à l'intérieur duquel le « bon » référent est isolable sur des critères distinctifs. Cette hypothèse peut être reformulée en trois points, dont chacun sera commenté de façon plus détaillée par la suite.

1. L'identification référentielle consiste en une opération de prélèvement d'un référent dans un ensemble. Ce principe est valable pour tout type d'expression référentielle.

2. Par conséquent, tout acte de référence présuppose un ensemble contextuel : son domaine de référence. Le modèle du contexte doit donc mettre à disposition de tels ensembles.

3. Le rôle d'une expression référentielle est de donner toutes les informations nécessaires pour isoler son référent des autres entités appartenant au même domaine de référence. Cela signifie en particulier que l'identification du référent implique aussi l'identification du complément de l'ensemble contextuel ou des « alternatives exclues ».

6.2.1 L'acte de référence comme opération de prélèvement

Voir le processus d'identification d'un référent par une expression référentielle non pas comme l'établissement d'un lien direct entre une expression et une entité désignée, mais plutôt comme une opération de prélèvement sur un ensemble contextuel aboutit à une vue unifiée sur les mécanismes d'identification référentielle. Il s'agit en quelque sorte du « plus petit commun dénominateur » pour tout type de marqueur référentiel. Nous développerons cet argument dans une des sections suivantes (cf. 6.3.2), où nous montrerons en particulier comment cette vision permet d'intégrer dans un modèle unifié de la référence des emplois parfois considérés comme marginaux, comme les anaphores associatives, les références mentionnelles et les groupes nominaux sans nom.

Par ailleurs, l'hypothèse de l'acte référentiel en tant qu'opération de prélèvement est tout à fait transposable au fonctionnement des indéfinis, comme le constate très clairement M. Galmiche (1986 : 52) : « *L'opération de prélèvement est une propriété très générale des déterminants indéfinis et définis* ». Selon lui, tout comme pour F. Corblin (1987), les définis *le* ou *les N(s)* sélectionnent un ou des éléments de type *N* au sein d'un ensemble de départ hétérogène, appelé « ensemble relationnel » et correspondant à notre domaine de référence. L'usage d'une description indéfinie présuppose, en revanche, un ensemble constitué d'éléments homogènes (pouvant faire lui-même partie d'un ensemble hétérogène) dont la description prélève la quantité d'éléments indiqués par le déterminant. Nous défendrons, lors de notre modélisation, le point de vue selon lequel l'acte de référenciation comme opération de prélèvement est également compatible avec le fonctionnement des pronoms et démonstratifs, à condition de disposer d'un ensemble de départ structuré selon des critères à définir.

Comme le fait déjà remarquer M. Galmiche (1986), cette hypothèse exige effectivement que la connaissance, la maîtrise et l'identification de ces ensembles soient partagées par les interlocuteurs. Selon lui, les ensembles peuvent être donnés par la situation d'énonciation, par des connaissances partagées ou par le contexte linguistique lui-même. Essayons alors de voir comment les modélisations contextuelles présentées au chapitre précédent peuvent fournir de tels ensembles.

6.2.2 La structuration du contexte en domaines de référence

Comme nous l'avons mentionné dans l'introduction de ce chapitre, les approches linguistiques ne se préoccupent pas de façon prioritaire de la modélisation du contexte, mais du fonctionnement des expressions référentielles. Ce fonctionnement permet néanmoins de conclure à la nécessité d'une structuration contextuelle sous forme de domaines de référence.

Parmi les approches que nous avons examinées aux chapitres précédents, les approches cognitives (chapitre 0) essaient de faire un lien entre une expression référentielle et le contexte : plus précisément,

elles prédisent l'usage d'une expression en fonction de la saillance du référent dans un modèle contextuel. Il s'agit donc d'une structuration unidimensionnelle du contexte qui repose essentiellement sur le degré de familiarité. Or, comme nous l'avons constaté en 3.3, cette caractéristique limite le pouvoir de prédiction de ces approches. Elle amène même certains des auteurs d'expérimentations psycholinguistiques à remettre en cause le rôle universel attribué au facteur de saillance : « *Manifestement, une expression référentielle plus explicite, comme le nom répété, n'est pas sous la tutelle du focus de discours.* » (Fossard, 1999 : 9).

Bien que le facteur de saillance doive certainement être pris en compte, en particulier pour le traitement des expressions pronominales, la seule structuration des entités contextuelles selon leur degré de saillance ne suffit donc pas à créer des sous-ensembles contextuels pouvant fonctionner comme domaines de référence. Concernant la création de tels sous-ensembles, nous avons vu émerger au chapitre 4, consacré aux modélisations contextuelles, certaines hypothèses : les espaces focaux dans les travaux de B. Grosz et C. Sidner (Grosz, 1981; Grosz et Sidner, 1986), les ensemble de centres dans la théorie du Centrage (Grosz et al., 1995) ou les espaces d'interprétation basés sur les relations discursives (Asher, 1993).

Lors de la synthèse de ces travaux (cf. la section 4.4), nous avons déjà souligné quelques points problématiques communs à ces approches. D'abord, elles ne tiennent pas compte du contexte perceptif pour introduire ou structurer des entités contextuelles. Pourtant, dans des dialogues homme-machine, le contexte perceptif se substitue ou se superpose facilement aux structures discursives : des facteurs perceptifs empruntés à la théorie de la Gestalt (Wertheimer, 1923) – comme la similitude et la proximité – introduisent en effet des ensembles pouvant servir de domaines d'interprétation. Ensuite, ces approches, essentiellement préoccupées par la modélisation du contexte, ont tendance à réduire les spécificités d'usage des différents marqueurs référentiels à des opérations d'introduction ou de ré-identification d'entités contextuelles.

6.2.3 Le rôle de l'expression référentielle : identifier le référent et les alternatives exclues

Toutes les propositions de modélisation contextuelle définissent les espaces d'interprétation (espaces focaux, ensemble de centres, S-DRS, ...) indépendamment de l'expression à interpréter et négligent ainsi le fait que l'expression elle-même peut, en fonction de son type de détermination et de sa sémantique, imposer des contraintes sur l'extension et la structuration de son domaine. Or, cette idée ressort directement de la description du fonctionnement linguistique des différents marqueurs référentiels. Reprenons comme exemple les principes de fonctionnement des expressions définies et démonstratives, tels que présentés au chapitre 2 : un défini comme *le triangle rouge* cherche à isoler, dans un ensemble d'objets de type *triangle*, un objet distingué par sa couleur alors qu'un démonstratif comme *ce triangle* cherche à isoler, dans un ensemble d'objets, un objet saillant pouvant être reclassifié en tant que « triangle ».

Lors de notre argumentation pour le concept de domaine de référence (cf. 6.3.1, ci-dessous), nous nous attacherons à montrer que l'idée selon laquelle une expression référentielle spécifie un domaine tout autant qu'un référent est loin d'être récente, en particulier dans la tradition des grammaires et sémantiques cognitives : elle est, par exemple, déjà présente dans les travaux de D. Olson (1970), qui défend une théorie de la référence sous forme de sémantique cognitive, affirmant que « *it is impossible to specify the meaning of a word or sentence unambiguously unless one knows the context and hence the set of alternative referents being entertained by the listener* »[54] (Olson, 1970 : 260). La conséquence de ce point de vue est que l'interprétation d'une expression est plus que l'identification

[54] « Il est impossible de spécifier le sens d'un mot ou d'une phrase de façon non ambiguë sans connaître le contexte et donc l'ensemble de référents alternatifs dont dispose l'auditeur. »

d'un référent : c'est aussi l'identification d'un ensemble d'alternatives exclues. Il est particulièrement intéressant de constater que ce point de vue sur la nature de l'opération référentielle a été intégré depuis longtemps dans les travaux de génération automatique d'expressions référentielles, en particulier par R. Dale (1992), en collaboration avec E. Reiter (1995, 1996). Dans les systèmes de compréhension, en revanche, l'approche la plus courante pour identifier le référent d'une expression référentielle (définie) consiste toujours à filtrer successivement les entités de l'application jusqu'à n'en retenir qu'une seule, compatible avec la description (cf. par exemple Kievit et Piwek, 2000). En principe, ce processus est réitéré pour chaque expression, mis à part éventuellement des heuristiques pour traiter les expressions d'altérité (Vivier et al, 1997), les ellipses et les *one*-anaphora (Winograd, 1972), où il est nécessaire de revenir sur des alternatives précédentes.

Contrairement à ces approches, nous défendrons l'hypothèse que l'identification systématique des « alternatives exclues » peut être mise à profit d'une modélisation des processus de compréhension qui soit à la fois plus efficace d'un point de vue informatique et plus proche du fonctionnement cognitif : ces « alternatives exclues » ou compléments d'ensemble semblent en effet fournir l'espace d'interprétation préféré pour les expressions elliptiques, les expressions ayant trait à l'altérité et les expressions ambiguës, comme nous le montrerons ci-dessous (6.3.2). Nous pensons que cela traduit un principe d'interprétation plus fondamental : ces compléments d'ensemble, activés par le « prélèvement » d'un premier référent, forment l'espace d'ancrage préférentiel pour toutes les expressions à interpréter. Ce principe peut s'appuyer sur la théorie de la pertinence (Sperber et Wilson, 1989) qui prédit que l'interprétation optimale est celle qui demande un minimum d'efforts cognitifs pour un maximum d'effets. Si l'on considère – étant donné un contexte activé C_i – que l'effet consiste en l'identification d'un référent pour une expression référentielle ER_{i+1}, l'effort peut alors être mesuré par la distance entre C_i et le contexte d'interprétation effectif C_{i+1}. Cet effort est minimal si $C_i = C_{i+1}$, c'est-à-dire si le domaine d'interprétation reste le même.

6.2.4 Synthèse : le processus d'interprétation d'une expression référentielle

Nous considérons qu'une expression référentielle isole un référent dans un ensemble de référents qui comprend l'objet à identifier et les « alternatives » – objets desquels le référent se distingue par la valeur d'une propriété distinctive. Nous appellerons cet ensemble le **domaine de référence** (DR) et la propriété distinctive le **critère de différenciation** (CD). L'expression à interpréter impose elle-même, par sa détermination et sa sémantique, des contraintes sur la structuration de son domaine de référence. Ces contraintes, que nous modéliserons sous forme d'un **domaine de référence sous-spécifié**, portent sur le type des objets du domaine et les propriétés permettant de partitionner le domaine afin d'isoler le référent. Pour un défini *le N P*, par exemple, le type des objets du domaine doit être *N* et le critère de différenciation doit être l'attribut associé à la valeur *P*. Sur la base de ces contraintes, **un sous-ensemble contextuel approprié est choisi** comme domaine de référence. Cet ensemble n'est pas obligatoirement constitué par la totalité des référents de type *N* disponibles dans le contexte de l'application : il peut s'agir d'un sous-ensemble. Nous verrons au chapitre 7 que de tels ensembles sont créés dynamiquement lors de l'introduction d'un référent pluriel, d'une coordination ou d'une énumération discursive et sur la base d'indices perceptifs, comme la proximité ou la similarité. C'est ici que les propositions concernant les structures discursives et les espaces focaux pourraient aider à délimiter d'autres ensembles d'objets, en particulier ceux liés à la structuration de la tâche en sous-tâches. Une fois un domaine de référence identifié, le référent en est extrait par une **opération de restructuration**. Cette opération restructure le domaine retenu en opposant le référent aux alternatives et en focalisant le référent. C'est sur la base de ce nouveau contexte que l'expression suivante sera interprétée.

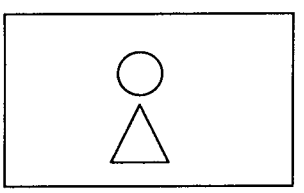
La différence essentielle d'une modélisation en terme de domaines de référence par rapport aux modèles existants consiste donc en la supposition d'un cadre référentiel impliqué dans tout acte de référence. Alors que les modèles traditionnels – comme par exemple celui de C. Sidner (1979) ou la DRT (Kamp et Reyle, 1993), avec l'extension qu'en proposent J. Bos et al. (1995) pour traiter les descriptions définies – ramènent le problème de la référence à une relation dont le prototype est le « liage » entre un référent accessible et une expression référentielle, nous pensons que **l'accès au « bon » référent se fait systématiquement via l'activation d'un domaine de référence**. Dans la section suivante, nous présenterons des arguments supplémentaires – d'ordre cognitif, linguistique, multimodal, psycholinguistique et discursif – qui renforcent cette hypothèse

6.3 Arguments pour les domaines de référence

6.3.1 Arguments cognitifs : référence et points de vue sur le monde

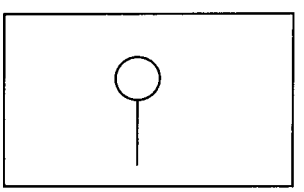
Dans notre optique, le concept de domaine de référence n'est pas (uniquement) un outil pour l'attribution d'un référent à une expressions référentielle, mais d'abord une hypothèse sur la nature d'un acte de référenciation. Comme nous l'avons déjà introduit dans la section précédente, nous faisons l'hypothèse qu'un acte de référence ne consiste pas à coller une étiquette sur un objet, mais à exprimer un point de vue du sur le monde. L'expression de ce point de vue permet à l'interlocuteur de se construire un modèle du monde en fonction de ce que conçoit le locuteur : celui-ci peut imposer, par exemple, un cadre référentiel particulier ou un critère de différenciation spécifique.

(99) I_1 donc tu vas prendre *le petit rond* pour faire la tête
I_2 voilà
[...]
I_3 ensuite tu vas prendre le petit triangle,
I_4 que tu vas mettre sous la *tête* (C8Eglise)



L'exemple (99) illustre cette possibilité : en I_1, le locuteur désigne le rond à manipuler par l'expression *le petit rond*. Le domaine de référence de cette expression ne peut être que la palette des figures géométriques, dont un élément est identifié comme référent et sera posé sur l'espace du dessin. Or, en I_4, le locuteur réfère à la même figure géométrique par l'expression *la tête*. Ce changement peut s'expliquer par une évolution du point de vue sur cette figure : d'un élément à isoler dans la palette des figures disponibles, elle devient un élément figuratif, constitutif du dessin cible qu'est la fillette. Les deux désignations se caractérisent donc par une différence du classement de l'objet perçu dans le contexte de la tâche.

(100) I_1 ça fait donc *un ballon*
avec une grande euh *une grande barre* (C8Ballon)



Il est par ailleurs intéressant d'observer qu'un changement de cadre référentiel sans motivation apparente peut avoir des effets « secondaires » : l'exemple (100) semble traduire une certaine maladresse, car le locuteur mélange en effet le domaine figuratif (*un ballon*) avec le domaine géométrique (*une grande barre*) à l'intérieur d'une même description.

Notre hypothèse ainsi que ces observations peuvent trouver une assise théorique dans la grammaire cognitive développée par R. Langacker (1991). Contrairement aux théories grammaticales et

sémantiques « classiques », la grammaire cognitive refuse l'autonomie de la syntaxe et l'approche vériconditionnelle du sens. Elle n'est pas conçue comme un ensemble de règles permettant de générer des phrases correctes, mais comme un inventaire d'unités symboliques qui résultent d'une association entre une structure phonologique et une structure sémantique. La décision sur la « bonne formation » d'une construction grammaticale se fait par une opération de comparaison qui évalue la distance entre la structure de cette unité et des schémas abstraits, obtenus par une généralisation des structures spécifiques.

La compréhension linguistique est considérée comme une opération cognitive de conceptualisation à partir des structures sémantiques. Chaque structure sémantique (appelée « *predication* » par R. Langacker) se caractérise relativement à un ou plusieurs domaines. Ces domaines, comparables à nos domaines de référence, sont de nature cognitive et peuvent être fournis par des expériences perceptives, des connaissances conceptuelles ou des connaissances encyclopédiques, sans que cette liste soit exhaustive : « *Any cognitive structure can function as the domain for a predication* »[55] (Langacker, 1991 : 61). Comprendre le sens d'une expression linguistique demande alors d'identifier correctement son domaine cognitif. Ainsi, parler d'une *hypoténuse* n'a de sens que par rapport à un *triangle rectangle*, une *pointe* se définit pour un *objet longitudinal* et un *oncle* est identifié dans un *réseau de parenté.*

Mais l'identification du domaine n'est pas suffisante. Encore faut-il tenir compte du fait qu'une structure sémantique véhicule une « imagerie conventionnelle ». Par là, R. Langacker entend « *our manifest capacity to structure or to construe the content of a domain in alternate ways* »[56] (ibid : 61). Cette caractéristique distingue clairement la grammaire cognitive des approches sémantiques qui considèrent que le sens d'un énoncé est exprimé par les conditions de vérité. Comme le montre l'exemple adapté de B. Partee (1984), deux phrases peuvent être vraies dans les mêmes mondes, sans pour autant imposer les mêmes contraintes sur la suite du discours :

(101) J'ai perdu dix billes et je n'en ai retrouvé que neuf. *Elle est ...

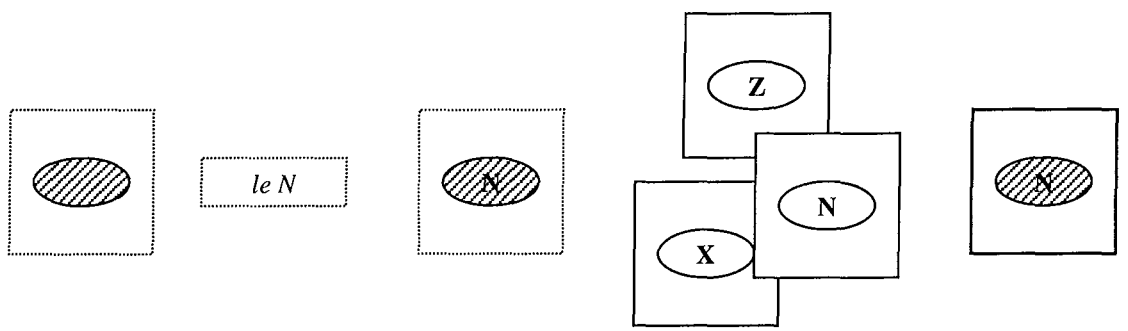
(102) J'ai perdu dix billes et je les ai toutes retrouvées, sauf une. Elle est ...

Analysées par les concepts de la grammaire cognitive, les phrases (101) et (102) n'ont pas les mêmes structures sémantiques. Plus exactement, celles-ci ne focalisent pas les mêmes éléments du domaine. En (102), l'accent est mis sur la bille perdue, alors que (101) focalise les billes retrouvées. La structure sémantique reflète donc la perception de la réalité par le locuteur et le rôle de la grammaire consiste précisément à structurer des scènes complexes en mettant en valeur certains éléments selon des découpages variés : selon la catégorie des éléments, selon leur positionnement, selon leur fonction ou encore selon leur saillance. C'est ainsi que l'on obtient différentes images d'un même domaine, car celui-ci est présenté sous différents points de vue. L'emploi du passif est un autre exemple, grammatical cette fois-ci : le même événement peut être présenté sous différentes perspectives, soulignant soit le rôle de l'agent, soit celui du patient. Encore une autre application de ce principe est l'expression des relations spatiales : Selon que la lampe se trouve *sur* la table ou la table *sous* la lampe, les éléments du domaine ne sont pas mis en valeur de la même façon.

Dans la grammaire cognitive, le rôle d'une construction grammaticale est donc de « profiler » une partie du domaine cognitif par rapport à une « base » : « *an expression is said to impose a particular image on its domain* »[57] (ibid : 61). Notre hypothèse, selon laquelle l'acte référentiel consiste en l'opposition du référent aux « alternatives exclues » d'un domaine de référence, pourrait être vue

[55] « N'importe quelle structure cognitive peut fonctionner comme domaine pour une prédication ». (c'est-à-dire comme domaine pour la conceptualisation d'une structure sémantique).
[56] « notre capacité manifeste de structurer ou de construire le contenu d'un domaine de plusieurs façons »
[57] « une expression est supposée imposer une image particulière sur son domaine »

comme l'application des principes de la grammaire cognitive au calcul référentiel. Si le processus d'analyse consiste, pour la grammaire cognitive, à « *identifier des structures schématiques ou prototypiques pertinentes et à les instancier avec les constituants reconnus pour constituer l'image mentale correspondant à la scène décrite* » (Sabah, 1995 : 5), l'analyse d'une expression référentielle pourrait se décrire telle que dans la Figure 18 : les structures schématiques ou prototypiques peuvent être dérivées des schémas abstraits proposés pour les expressions nominales : un nom, par exemple, est vu comme délimitant une région dans un domaine (a). Une expression référentielle sous forme de groupe (pro-)nominal (b) sur-spécifie donc le schéma abstrait pour les noms en l'appariant avec un (ou plusieurs) domaine cognitif correspondant à la structure sémantique de l'expression nominale (c). Ensuite, ce schéma est lui-même utilisé comme motif de recherche pour trouver un domaine compatible fourni par le contexte d'interprétation (d). L'identification d'un tel domaine permet effectivement d'instancier le schéma et d'identifier le référent qui sera profilé (e).



a. b. c. d. e.

Figure 18 : Application des principes de la grammaire cognitive au calcul référentiel

La modélisation du calcul référentiel que nous présenterons aux chapitres suivants est en effet comparable à ce scénario. Si notre modélisation peut être vue comme une tentative de dépasser l'absence notoire de formalisation dans les travaux de R. Langacker, la possibilité de son rapprochement de la grammaire cognitive n'est pas surprenante : notre objectif consiste à interpréter des expressions référentielles et nous avons vu que cela dépasse une modélisation purement linguistique, dans la mesure où l'identification des référents doit s'appuyer sur un modélisation dynamique du contexte, construit à partir de critères linguistiques et non linguistiques. La préoccupation centrale de R. Langacker est précisément de représenter le sens d'un énoncé par des structures mentales créées dynamiquement, à partir d'expériences verbales et perceptives. La prise en compte du contexte verbal et perceptif est donc fondamentale dans la grammaire cognitive, car contrairement à des théories linguistiques, elle ne fait pas de distinction entre l'analyse d'un énoncé et la prise en compte du contexte : « *Les deux sont simultanés. Pour construire le sens d'un nouvel énoncé, on essaie d'apparier globalement la paire* énoncé– contexte d'énonciation *à une ou plusieurs structures cognitives antérieurement acquises.* » (Bordeaux, 1993 : 71).

6.3.2 Arguments linguistiques : réhabilitation des emplois « marginaux »

Cette section est consacrée au traitement de quelques expressions référentielles qui nous sont apparues, à travers l'examen des modélisations proposées, comme « marginales », au sens où elles ne semblent pas s'intégrer facilement dans les mécanismes proposés par ailleurs dans la littérature. Cela concerne certains emplois des descriptions définies, les expressions ordinales et d'altérité, les ellipses ou groupes nominaux sans nom et enfin les références mentionnelles. Nous allons montrer ici

comment la notion de domaine de référence pourrait permettre de « réintégrer » ces emplois dans un modèle unifié de la référence.

Un de nos objectifs est de proposer une modélisation unifiée qui soit capable de prendre en compte les différents usages des marqueurs référentiels. Lors de la présentation d'un certain nombre de modélisations et algorithmes (cf. les chapitres 4 et 5), nous avons vu que ce point est souvent délicat en ce qui concerne les descriptions définies : d'une part, beaucoup d'approches ont du mal à prédire correctement la distribution entre l'emploi d'une description définie et d'un pronom. D'autre part, elles n'arrivent pas à intégrer les différents usages d'une description définie (anaphorique vs. associatif) dans un même principe de fonctionnement et ont alors tendance à considérer les emplois associatifs comme moins « typiques ». Considérer que le principe général pour l'interprétation des descriptions définies est l'extraction du référent d'un domaine de référence (Corblin, 1987 ; Gaiffe et al., 1997) permet à la fois de séparer l'emploi du pronom de celui du défini et de restaurer l'unité des différents usages du défini :

(103) I_1 on a *une lampe et une fleur* à réaliser
M_1 d'accord
I_2 bon on va commencer par *la lampe / *elle*[58] (C8Lampe)

(104) I_1 on a *une lampe* à réaliser
M_1 d'accord
I_2 *elle / ? la lampe* se trouve à droite (C8Lampe, modifié)

(105) I_1 on a *une salle de réception* à réaliser
M_1 d'accord
I_2 bon on va commencer par *la lampe* (C8Lampe, modifié)

La comparaison des exemples (103) et (104) illustre le fait que l'usage d'une description définie, contrairement à celui du pronom, s'inscrit beaucoup mieux dans un contexte ou domaine mettant à disposition un ensemble de référents à l'intérieur duquel il est possible d'en isoler un sur des critères distinctifs. Or, les algorithmes classiques pour l'analyse des expressions anaphoriques (Sidner, 1979) ou la sémantique formelle (Kamp et Reyle, 1993) ne permettent pas de faire cette prédiction, contrairement d'ailleurs aux algorithmes de génération automatique d'expressions référentielles (Dale, 1992). Une comparaison de (103) et (105) montre aussi que l'usage associatif des définis n'est pas un cas particulier ou « moins prototypique », comme le pensent J. Bos et al. (1995) : à condition de disposer d'un domaine de référence approprié – que celui-ci soit donné par un ensemble discursif comme en (103) ou par des connaissances sur la composition des entités contextuelles comme en (105) –, le défini fonctionne comme extracteur d'un élément du domaine.

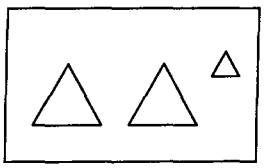
Un autre argument linguistique pour une modélisation du contexte en termes de domaines de référence vient de l'interprétation des expressions ordinales et d'altérité. Ces expressions, souvent traitées à part (*SHRDLU*, Winograd, 1972 ; *Compérobot*, Vivier et al., 1997) et dont une partie – les références mentionnelles – semblent occuper un statut particulier (Corblin, 1999 ; Laborde, 1999), peuvent effectivement être intégrées dans une modélisation du calcul référentiel sur la base de domaines contextuels.

La justification du traitement à part des expressions ordinales et d'altérité dans le système *Compérobot*, par exemple, est précisément que leur interprétation implique la prise en compte d'un ensemble contextuel. Ce constat correspond effectivement à des hypothèses sur la sémantique de ces expressions (Schnedecker, 1998, Berrendonner et Reichler-Béguelin, 1996). Les ordinaux y sont considérés comme adjectifs « paradigmatisant », car leur emploi présuppose l'existence d'un ensemble

[58] Afin de montrer que l'inacceptabilité du pronom personnel n'est pas due à l'ambiguïté en genre de « lampe » et « fleur » en I_1, on aurait pu remplacer « fleur » par « ballon ».

d'éléments ordonnés semblables à celui spécifié par l'ordinal. L'ordonnancement peut provenir de l'ordre des mentions textuelles, de la chronologie événementielle exprimée linguistiquement ou de la disposition spatiale perçue par le locuteur et/ou par l'auditeur.

(106) I1 ensuite il va falloir tracer l'horizon.
I2 donc tu vas commencer par prendre une petite barre
que tu vas mettre à gauche de la pointe *du premier triangle*
(C8Egypte)



En (106), l'ensemble présupposé par l'expression *le premier triangle* est un ensemble perceptif de triangles, ordonnés selon leur disposition horizontale sur l'écran. En tenant compte du sens de lecture habituel, cet ensemble permet en effet d'interpréter l'expression en question comme référant au triangle le plus à gauche.

L'interprétation de *autre* s'appuie également sur un domaine de référence, avec la particularité qu'il doit y exister un objet-repère déjà identifié. Contrairement aux ordinaux, ce n'est pas un ordre « extérieur » qui distingue ici les éléments du domaine : *autre* installe un ordre épistémique qui s'appuie sur le seul fait que les objets du domaine sont ontologiquement discernables. Mais comme pour les relations d'ordre, cette distinction peut s'appuyer sur des critères discursifs ou perceptifs : en (107), elle repose sur une mention de l'objet-repère, alors qu'elle se fait par rapport à un élément manipulé, mais non mentionné explicitement en (108) :

(107) I_1 maintenant il faut prendre un grand triangle et le mettre à gauche de l'écran
[...]
M_{I1} d'accord
I_8 maintenant prendre *un autre grand triangle* (C11Egypte)

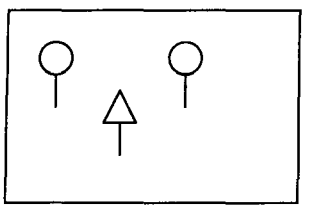
(108) I_1 oui la route alors c'est deux traits verticaux, deux grands traits verticaux
I_2 alors j'y vais alors la route elle monte
[manipule un grand trait...]
I_4 oui, voilà je vais la mettre là
I_5 on va chercher *l'autre morceau* (C7Route)

Un troisième argument linguistique pour la défense de la notion de domaine de référence est la possibilité de traiter certaines ellipses nominales, présentées sous l'appellation de « groupes nominaux sans nom » au chapitre 2 (section 2.6). Les mécanismes de calcul référentiel pour ces constructions sont soit séparés des autres mécanismes (*SHRDLU,* Winograd, 1972), soit peu abordés. En revanche, un certain nombre de travaux sont explicitement consacrés au traitement automatique de l'ellipse (par exemple Carberry, 1989, pour l'anglais ; Sauvage, 1993, pour le français), ce qui contribue à séparer cette problématique de la problématique référentielle. Il existe pourtant un lien entre ces deux problématiques et ce lien apparaît à travers la notion de domaine de référence.

L'examen linguistique de ces constructions a mis en évidence que la particularité de l'interprétation référentielle des groupes nominaux sans nom consiste en une séparation de deux opérations distinctes : d'une part, la récupération de la tête du syntagme et d'autre part le calcul du référent de la totalité du groupe. Or, ce fonctionnement s'intègre fort bien dans notre hypothèse sur le fonctionnement général des expression référentielles : ils spécifient un cadre référentiel et isolent un des éléments de ce cadre. Pour les ellipses nominales, la première de ces opérations consiste à récupérer un domaine de référence donné par le contexte. En raison de l'absence de tête nominale, le type des éléments de ce domaine reste sous-spécifié. La deuxième opération, l'isolement du référent dans ce domaine, correspond au calcul référentiel pour la totalité du groupe. Comme l'a déjà intégré T. Winograd (1972) dans la formulation d'une heuristique pour le traitement des ellipses (cf. la section 5.4.1), le choix du

domaine, s'il n'est pas *a priori* contraint par le type des éléments, peut dépendre de la structuration de ses éléments selon une propriété indiquée par le groupe elliptique : dans l'exemple (109), le groupe nominal sans nom *le deuxième* demande effectivement à s'interpréter dans un domaine dont on ne connaît pas le type des éléments, mais dont on sait que les éléments sont structurés selon un ordre strict. Cette contrainte est remplie par le domaine des deux arbres ronds de la scène, ayant été explicitement introduit en I_2, mais aussi par l'ensemble des trois arbres du dessin, ce qui provoque d'ailleurs l'hésitation lors de la formulation en I_7 et lors de la compréhension en M_1.

(109) I_1 si tu veux prendre un petit triangle
I_2 et tu le mets entre les deux arbres que je viens de dessiner en fait [...]
I_3 voilà là
I_4 et tu vas prendre une petite tige verticale et tu vas la mettre dessous
I_5 voilà
I_6 et tu fais la même chose là tu prends un triangle
I_7 tu mets euh vers *le deuxième* ... tout à droite ... de l'arbre rond
M_1 à droite de l'arbre euh du deuxième à droite
I_8 voilà (C5Forêt)



Enfin, un dernier argument en faveur des domaines de référence est la possibilité de traiter les références mentionnelles, sans les considérer comme une classe à part. M.-C. Laborde (1999) propose en effet un traitement spécifique de ces expressions dans le cadre de la DRT. Ce traitement passe par l'ajout d'informations grammaticales et structurelles à la DRS. Cette procédure permet effectivement de calculer les antécédents des références mentionnelles, mais elle masque ce qu'il y de commun entre de tels fonctionnements et les autres emplois des ordinaux ou expressions d'altérité.

Pour notre part, nous proposons de ne pas séparer ces emplois mentionnels du principe d'interprétation des groupes nominaux sans nom en général. Il nous semble que l'interprétation mentionnelle est un cas particulier de certaines interprétations de *autre* et *premier/second*, mais elle reste compatible avec le principe général d'interprétation des groupes nominaux sans nom. Ces expressions nécessitent un domaine de référence saillant donné par le contexte, duquel il est possible d'extraire un élément en raison d'une propriété particulière. Cette propriété peut être, entre autre, l'ordonnancement discursif des éléments (pertinent pour les ordinaux, mais pas pour *l'un / l'autre*), mais c'est seulement une possibilité parmi beaucoup d'autres. Les critères invoqués pour justifier l'autonomie de la catégorie des références mentionnelles sont en fait une conséquence directe d'une interprétation dans laquelle on s'appuie, en l'absence d'autres possibilités, sur la matérialité du texte pour structurer les référents en domaines. Mais dans la mesure où cette particularité n'affecte pas le principe d'interprétation tel que nous l'avons formulé, la mise à part de cette catégorie ne nous semble pas indispensable.

6.3.3 Arguments multimodaux : domaines perceptifs et interprétation spatiale

Si nous sommes à la recherche d'un modèle de la référence qui soit unifié pour différents types d'expressions référentielles et différents usages d'un même type d'expression, il nous semble également indispensable de tenir compte des différentes sources d'information pour l'interprétation référentielle. Nous avons déjà remarqué à plusieurs reprises que les seules informations discursives ne suffisent pas à structurer un modèle contextuel qui soit adapté au dialogue homme-machine : si nous considérons un utilisateur qui se trouve face à une machine, les informations perceptives et en particulier la disposition des objets sur l'écran vont influer sur sa manière de référer. De plus, si l'utilisateur a la possibilité de se servir d'un dispositif de désignation (souris, gant de désignation,...),

il choisira le mode de référenciation qui lui semble le plus efficace. Cela signifie que le modèle contextuel ne peut pas se passer d'une prise en compte du contexte perceptif et des gestes.

Or, plutôt que d'envisager un traitement séparé des informations provenant de différents canaux de communication, il peut être intéressant de définir des modèles de représentation du contexte qui intègrent les contraintes liées à l'usage du mode gestuel sans pour autant oublier les connaissances que nous avons déjà sur le fonctionnement des mécanismes linguistiques de la référence. Nous allons donc montrer qu'une structuration du contexte en domaines de référence correspond à des besoins exprimés par des études consacrées à l'interprétation spatiale. Ces études ont en effet en commun de s'appuyer sur un contexte sous forme d'un cadre spatial – construit à partir de données perceptives – à l'intérieur duquel le référent est distingué par rapport à des alternatives. Le rôle du geste y est considéré comme délimitant un espace en fonction de la granularité du cadre.

Un premier argument est fourni par la redéfinition du principe d'interprétation de l'adverbe de lieu *ici* (Romary, 1993). A partir d'une observation des seuls énoncés de positionnement dans des corpus d'interaction sur des tâches finalisées (110), on pourrait aboutir à une conception « naïve » du fonctionnement de cet adverbe en l'associant directement à un élément de l'univers de la tâche, désigné par un clic de souris.

(110) I_1 et tu le mets entre les deux arbres que je viens de dessiner en fait
M_1 [+ geste de désignation avec la souris] : *ici*
I_2 un peu plus en bas
I_3 voilà là (C5Forêt)

Or, concevoir l'interprétation de *ici* par une association avec les coordonnées du point désigné par la souris est une vue simplificatrice qui n'est valide que dans des tâches particulières. Dès que la tâche change (111) ou si l'espace uniforme de l'écran est abandonné pour un environnement plus complexe (112), il faut envisager une représentation contextuelle plus élaborée, donnant une extension et une structure au lieu.

(111) Mets de la moquette *ici.* (Mignot et al., 1993, repris par Romary, 1993)

(112) Éteindre la cigarette *ici.* (Romary, 1993)

Le fonctionnement de *ici* est alors revisité comme réalisant « *un filtrage sur une structure de pavage de l'espace pour isoler l'un des éléments de cette structure qui se trouve directement mise en évidence par le locuteur, soit par la simple situation d'énonciation, soit par un geste lorsque l'unité de pavage est relativement fine* » (Romary, 1993 : 7). L'isolation d'un élément dans cette structure établit un contraste spatial par rapport à un ensemble d'alternatives. Cela signifie que l'interprétation de *ici* demande l'abandon d'une structure à base de points au profit d'une représentation par l'intermédiaire d'un pavage de l'espace en unités élémentaires. Or, cette structure contextuelle peut être modélisée par un domaine de référence, regroupant des entités pertinentes sur un critère spatial dont la granularité dépend de la nature de l'objet à positionner et de la nature du prédicat.

Cette proposition revient en fait à étendre l'idée de D. Schang (1997) qui modélise le contexte pour la référence spatiale par des cadres de référence. Ceux-ci sont définis comme « *entités qui restreignent l'univers à une portion d'espace ou à une connexion de portions d'espaces nécessaires et suffisantes, considérées d'un point de vue particulier, autorisant l'achèvement du raisonnement spatial* » (Schang, 1997 : 131). Le principe d'interprétation d'une expression spatiale comme *le carré à droite du triangle* consiste alors à construire une image mentale du cadre de référence associé à la préposition *à*

droite de et à en instancier le site[59] *le triangle* (Figure 19a). Cette image sert ensuite de motif de recherche sur l'environnement perceptif disponible pour d'identifier la cible[60] *le carré* (Figure 19b).

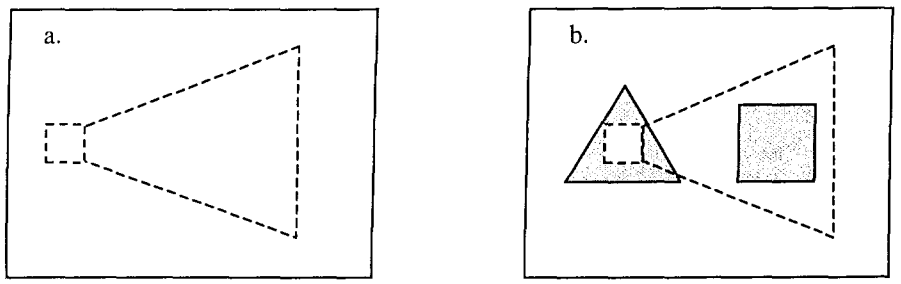
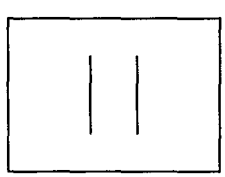


Figure 19 : Interprétation de « le carré à droite du triangle » dans un cadre spatial (Schang, 1993)

Un des intérêts de ce travail par rapport à ce que nous proposons est donc de souligner le caractère plus général des domaines de référence sous-spécifiés : en ce qui concerne la référence aux objets, nous avons effectivement noté que l'expression à interpréter impose elle-même, par sa détermination et sa sémantique, des contraintes sur la structuration de son domaine de référence. Nous avons proposé de modéliser ces contraintes par des domaines sous-spécifiés. Pour la référence spatiale, D. Schang part du même principe en guidant la recherche par des images mentales dont l'extension dépend des prépositions et de la nature des sites. Ces images mentales remplissant la même fonction que nos domaines sous-spécifiés.

Un autre intérêt de ce travail est la possibilité d'établir un parallèle entre la nature des cadres de référence et nos domaines de référence : Comme nos domaines de référence, les cadres de référence sont des entités qui évoluent dynamiquement au cours du dialogue, qui focalisent un objet en l'opposant aux autres objets du domaine et qui sont capables d'exprimer plusieurs points de vue sur une même structure contextuelle. Par ailleurs, les cadres de référence sont présentés comme relevant d'un principe d'économie cognitive, dans la mesure où l'interprétation à effort minimal est celle qui vise la stabilité d'un cadre établi. Cette hypothèse, que nous reprenons à notre compte pour ce qui est des domaines de référence en général (cf. la section suivante), permet d'ailleurs d'expliquer certains malentendus portant sur les domaines spatiaux du corpus observé : en (113), I_5 ne spécifie pas le site par rapport auquel l'expression *au milieu* doit être interprétée. Selon le principe d'économie cognitive, ce site est donc repris d'un cadre antérieur, mais il y a divergence sur la nature de ce cadre. Pour I, il s'agit de la dernière barre posée, alors que M se positionne dans le cadre fourni par l'ensemble des deux barres posées sur l'écran.

(113) I_1 alors tu prends une grande barre verticale
I_2 voilà
I_3 tu en mets une deuxième à coté [...]
I_4 c'est bon tu prends un petit carré
I_5 et tu le mets e *au milieu...*
M_1 [place le carré entre les deux barres]
I_6 non pas au milieu des deux barres
I_7 sur le coté d'une barre au milieu (C12Route)



Les problèmes spécifiques à l'interprétation spatiale semblent donc rejoindre la problématique générale de la référence : la nécessité d'un contexte d'interprétation structuré en ensembles locaux à l'intérieur desquels s'opère une opposition entre référent et alternatives. La particularité des références spatiales consiste à s'appuyer non pas sur un contexte discursif, mais sur un contexte perceptif. C'est

[59] Le site est l'objet servant de référence spatiale (Vandeloise, 1986).
[60] La cible est l'élément à positionner (Vandeloise, 1986).

aussi le cas pour l'interprétation de certaines expressions « indexicales » (déictiques et démonstratifs) accompagnés ou non de gestes, dont le fonctionnement référentiel a été modélisé par J. Moulton et al. (1994). Cette modélisation est également plus proche du modèle des domaines de référence que d'un modèle « standard » considérant que le calcul référentiel consiste en un filtrage successif des entités contextuelles sur la base de propriétés dérivées de la dénotation de l'objet. La proposition de ces auteurs est de replacer un tel modèle par un «*figure-ground-model* »[61], reposant sur les hypothèses suivantes : l'emploi d'une expression « indexicale » détermine un contexte qui contient le référent. Le rôle d'un geste est d'attirer l'attention sur une partie de ce contexte. Le contenu descriptif de l'expression instaure un contraste entre l'arrière-plan et le référent. La proximité avec la notion de domaine de référence est assez claire : les auteurs définissent le rôle majeur de la référence comme « *giving contrast, rather than true description* »[62]. Par ailleurs, leur modèle a l'avantage de projeter un point de vue unifié sur le fonctionnement des expressions référentielles et des gestes : « *Both gestures and descriptions function in the figure-ground-model in virtue of contrasts to backgrounds. Gestures contrast the location of the referent to other parts of surroundings. Descriptions need not be true of the referent, as long as they suffice for distinguishing it from other possible referents.* »[63] (Moulton et Roberts, 1994 : 3).

6.3.4 Arguments psycholinguistiques : l'hypothèse de la stabilité du domaine référentiel

Un argument supplémentaire pour la réalité cognitive des domaines de référence pourrait venir d'une confirmation de l'hypothèse sur la stabilité de ces domaines lors de l'interprétation d'expressions ambiguës. Au vu du principe détaillé au début de ce chapitre (6.2.4), nous faisons effectivement l'hypothèse qu'une interprétation est d'autant plus pertinente qu'elle est moins coûteuse. Or, le coût d'interprétation est lié à la distance entre des domaines activés successivement. A partir de là, nous pensons effectivement qu'il est raisonnable de supposer qu'en cas d'ambiguïté entre deux interprétations, l'interlocuteur choisira celle qui lui demande le moins d'efforts en termes de changement domanial. Avant de donner les résultats de deux expériences, voici un exemple qui montre que notre hypothèse se trouve confirmée par les données du corpus Ozkan :

(114)	I_1 (S1)	il faut prendre une grande horizontale	
	I_2	et la placer à la pointe *des deux grands triangles* [...]	
	M_1	(+geste)	
	I_3 (S2)	voilà comme ça ... et tu en prends une deuxième	
	I_4	une petite [...]	
	I_5	tu la places à gauche de *la pyramide de gauche* [...]	
	M_2	(+geste)	
	I_6 (S3)	voilà comme ça [...] et t'en prends une autre petite [...]	
	I_7	et tu la places à droite de *la pyramide de droite* [...]	
	M_4	(+geste)	
	I_8 (S4)	voilà comme ça... et tu prends une autre petite verticale	
	M_5	une autre petite verticale	
	I_9	e horizontale pardon	
	I_{10}	et puis tu la places dans la même lignée à droite de la petite pyramide	(C5Egypte)

[61] éminemment Gestaltiste... (Wertheimer et al., 1923)
[62] « donner du contraste, plutôt qu'une description vraie »
[63] « Dans le modèle *figure-ground*, gestes et descriptions fonctionnent toutes les deux en vertu de contrastes avec des fonds. Les gestes contrastent la localisation d'un référent par rapport à d'autres parties de l'environnement. Les descriptions n'ont pas besoin d'être vraies du référent, tant qu'elles suffisent pour le distinguer d'autres référents possibles. »

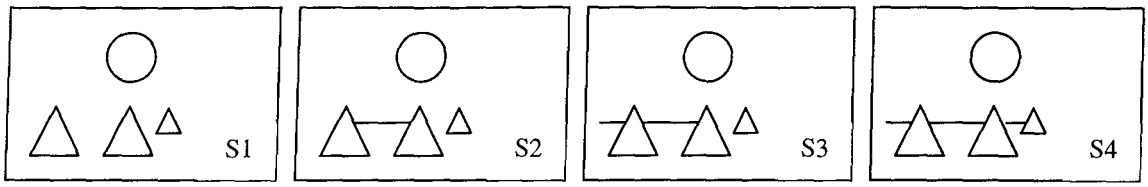


Figure 20 : Scènes successives pour l'exemple (114)

L'extrait proposé contient un exemple d'ambiguïté référentielle : hors contexte, *la pyramide de droite* en I_7 réfère à la pyramide la plus à droite de la scène courante (S3), à savoir la petite pyramide. Or, ce n'est pas ce qu'envisage I ni ce que comprend M. La raison pour cela est un domaine actif comprenant les deux grandes pyramides. Ce domaine a été introduit en I_2 par l'expression *les deux grands triangles* et est maintenu comme domaine actif en I_5 par l'extraction de *la pyramide de gauche.* L'hypothèse sur la stabilité des domaines de référence prédit donc que l'interprétation la moins coûteuse de I_7 est une référence à la deuxième pyramide de ce domaine (la plus à droite des deux grandes) et c'est effectivement celle-ci qui est identifiée.

Des expérimentations psycholinguistiques entreprises par K. Kessler et al. (1996) confirment également cette hypothèse. Nous avons déjà mentionné ces travaux en fin du chapitre 4 pour montrer l'insuffisance des modèles contextuels purement discursifs. Mais ces expériences apportent en plus des données expérimentales intéressantes par rapport à notre hypothèse sur la stabilité du domaine de référence. Une des hypothèses des auteurs est effectivement de considérer « l'historique des éléments focaux » (*focus history*) comme facteur pertinent pour la résolution référentielle dans des situations ambiguës.

A une scène comportant divers objets (Figure 21) ont été associées des séries d'instructions selon le schéma présenté dans l'exemple (115). Les instructions (a) et (b) servent à introduire le focus primaire (a) et secondaire (b) à travers les expressions *à gauche* et *à droite*, qui attirent l'attention sur un des groupes perceptifs de la scène. L'instruction (c) comporte un expression référentielle ambiguë par rapport à la totalité des objets de la scène. Selon l'hypothèse des auteurs, dans une suite instructionnelle de type (a) – (c) (omission de la phase (b)), l'identification de l'objet devait se faire en fonction de l'activation du groupe perceptif sous (a). Les auteurs constatent effectivement que c'est le cas : ils montrent par des analyses statistiques que la distribution du focus primaire (gauche vs. droite) est le seul facteur indispensable à la modélisation probabiliste de l'identification du référent sous (c). Ce résultat confirme clairement notre hypothèse sur la continuité du domaine de référence : le focus primaire active un des domaines perceptifs – gauche vs. droite – et c'est ce domaine qui fournit majoritairement le cadre d'interprétation de l'expression référentielle suivante.

(115) a. Marquez un cercle sur *la droite / la gauche* !
b. Marquez un cercle sur *la droite / la gauche* !
c. Marquez *le cube* !

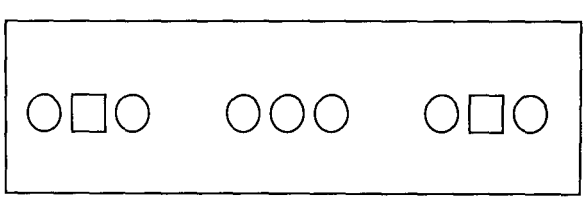


Figure 21: Scène d'ambiguité référentielle (d'après Kessler et al.,1996)

Cependant, l'expérience apporte une donnée supplémentaire intéressante. En intégrant l'instruction (b) dans l'expérimentation, les auteurs ont voulu tester deux hypothèses concurrentes :

« *primacy effect* »[64] : l'identification de l'objet en (c) guidée par le focus primaire (a)) *versus*

« *recency effect* »[65] : l'identification de l'objet en (c) guidée par le focus secondaire (b)

Les résultats sont différenciés (Tableau 15) : dans le cas d'une continuité entre focus primaire et secondaire (combinaison 1 et 4), les résultats précédents ont été confirmés. Cela signifie pour nous que l'activation, puis la stabilité d'un domaine sous forme de groupe perceptif permet effectivement de faire des prédictions sur l'interprétation référentielle d'une expression ambiguë En ce qui concerne les cas de disjonction entre focus primaire et secondaire (combinaison 2 et 3), l'identification s'est faite majoritairement à gauche. Ce résultat a été attribué à l'effet de balayage de gauche à droite, influencé par le sens de lecture habituel. L'expérience a alors été répétée en remplaçant le facteur positionnel par un focus induit par des couleurs. Dans cette seconde expérience, le « *primacy effet* » l'a emporté surle « *recency effect* ». Cela signifie d'abord qu'il faut garder plus d'un domaine dans l'historique. Par ailleurs, dans certaines tâches, le domaine le plus stable ne serait pas le dernier activé, mais le premier instauré, ce qui s'explique éventuellement par le caractère plus fort d'une structuration initiale de l'espace de la tâche.

Instruction	***Combinaison 1***	***Combinaison 2***	***Combinaison 3***	***Combinaison 4***
focus primaire (a)	gauche	gauche	droite	droite
focus secondaire (b)	gauche	droite	gauche	droite
identification majoritaire	gauche	gauche	gauche	droite
hypothèse confirmée	continuité du domaine activé	« *primacy effect* »	« *recency effect* »	continuité du domaine activé

Tableau 15 : Résultats de l'expérimentation de Kessler et al. (1996)

Enfin, la recherche d'une stabilité des représentations contextuelles peut être vue comme relevant d'un principe cognitif plus général. Une expérience menée par I. Rock et S. Palmer (1991)[66] montre les difficultés qu'éprouvent les sujets lors d'une tâche qui consiste à relier, par quatre traits et sans lever le stylo, les neuf points de la Figure 22. Cela n'est en effet possible que si les sujets dépassent le cadre instauré par l'espace limité par les points, ce qui semble leur poser des difficultés. Si nous transposons cette expérience à la problématique référentielle, elle plaide effectivement en faveur de l'hypothèse sur la stabilité des domaines de référence.

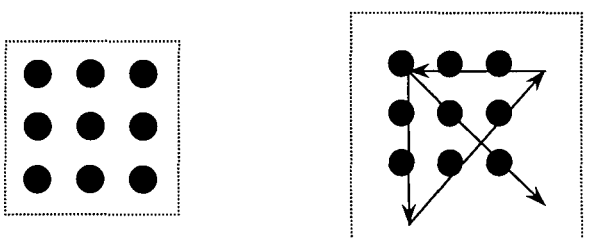


Figure 22 : Sortir d'un cadre donné ? (Rock et Palmer, 1991)

[64] « effet de primauté »
[65] « effet de récence »
[66] Nous décrivons cette expérience d'après Schang (1997).

6.3.5 *Arguments discursifs : des domaines de référence aux structures du discours ?*

Un dernier argument en faveur des domaines de référence pourrait être le fait qu'ils arrivent à intégrer certains aspects liés à la structuration du discours. Nous avons effectivement vu au chapitre 4 que la structuration discursive d'un texte – que ce soit sous forme de relations rhétoriques, temporelles, événementielles ou intentionnelles – est liée à l'organisation référentielle et détermine en partie l'accessibilité des référents. L'hypothèse qui réunit un certain nombre de travaux (dont Grosz et Sidner, 1986 ; Asher, 1993 ; Cristea et al., 1998) est la suivante : l'interprétation d'une expression référentielle ne se fait pas par la recherche d'un antécédent dans la globalité du texte, mais seulement dans des segments du texte qui ont en commun d'être liés par des relations particulières.

Il apparaît alors que le concept de domaine de référence pourrait couvrir – au moins en partie – ces aspects, à condition de l'élargir à d'autres types d'entités. Si nous l'avons introduit dans un premier temps pour regrouper et structurer des objets, nous aurions ici besoin de l'appliquer à des entités de type événementiel. Regrouper et structurer des événements en domaines de référence permettrait en effet de tenir compte d'un certain nombre d'observations ayant trait à des problèmes de référence : Premièrement, un événement complexe peut fournir le cadre pour l'extraction ou la particularisation de ses sous-événements (Asher, 1993 ; Danlos, 1999). Deuxièmement, le regroupement d'événements peut contraindre l'accessibilité référentielle de ses participants (Asher, 1993). Enfin, l'interprétation des « déictiques discursifs » (*this, that*) ou d'autres marqueurs linguistiques peut demander la prise en compte d'un contexte regroupant plusieurs événements (Webber, 1991 ; Asher, 1993). L'exemple (116) montre en effet que l'expression *la même chose* en I_3 réfère à un ensemble d'actions, regroupant les événements correspondant à *prendre une tige* et *la coller dessous* :

(116) I_1 tu prends une tige, une petite tige verticale et tu mets dessous
M_1 je la colle ?
I_2 voilà
I_3 tu fais *la même chose* e à gauche de l'écran

Élargir la notion de domaine de référence aux événements demande bien sûr de réfléchir aux conditions permettant de grouper des événements en domaines. Au chapitre 4, nous avons déjà mentionné quelques problèmes lié à cette opération. Un des problèmes importants est la représentation des connaissances nécessaires au calcul des relations entre événements. A. Lascarides et N. Asher (1993) proposent par exemple une formalisation du calcul des relations entre propositions – comme la narration ou l'élaboration – sur la base de règles logiques par défaut. Ces règles expriment des connaissances encyclopédiques intervenant dans l'interprétation du discours. L'application de ces règles à l'exemple (117) aboutit à la structuration des événements de la Figure 23 :

(117) a. Guy experienced a lovely evening last night.
b. He had a fantastic meal.
c. He ate salmon.
d. He devoured lots of cheese.
e. He won a dancing competition.

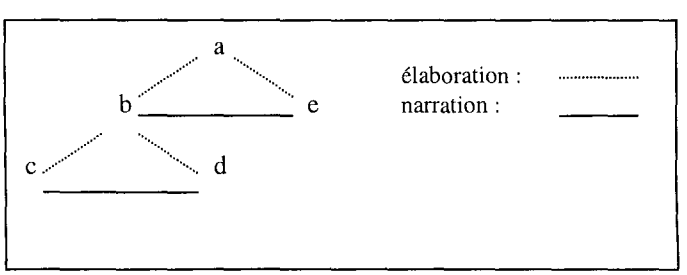


Figure 23 : Structure discursive de l'exemple (117) selon Lascarides et Asher (1993)

Le problème général de ces approches consiste alors à ne donner que des indications imprécises sur la manière de segmenter le texte. Si l'on adopte par exemple les propositions de A. Lascarides et N. Asher (1993), il faudrait beaucoup d'axiomes pour calculer la structure discursive de textes réels. Un deuxième écueil à éviter est le problème de circularité, observé à propos des travaux de B. Grosz et de C. Sidner (1986) : d'une part, elles se servent des expressions référentielles pour calculer la structure discursive et d'autre part, elles proposent de considérer cette structure comme prédictive quant à l'interprétation de ces mêmes expressions.

Une solution pourrait alors consister à considérer, dans un premier temps, l'organisation référentielle indépendamment de la structuration discursive. Ceci serait d'ailleurs une piste pour étudier la motivation d'une structuration discursive à travers l'évolution des domaines de référence. Ainsi, dans l'exemple (117), nous pensons qu'il est au moins aussi plausible de supposer qu'il y a élaboration entre (b) et (c) parce qu'il y a extraction de *salmon* d'un domaine de référence introduit par *fantastic meal* que de supposer l'existence d'un axiome postulant que manger du saumon peut être une partie d'un repas fantastique. Parallèlement à cela, on pourrait faire l'hypothèse que le parcours d'un même domaine correspond à une relation de narration : *lots of cheese* en (d) s'interprète effectivement dans le même domaine de référence que *salmon.*

6.4 Bilan : caractéristiques des domaines de référence

A partir du fonctionnement linguistique des marqueurs référentiels (chapitre 2), nous avons fait l'hypothèse d'une modélisation du contexte par des domaines de référence : jusqu'ici nous avons considéré que ces domaines sont des ensembles contextuels locaux qui regroupent et structurent des entités contextuelles de façon à prédire la distribution des différentes expressions référentielles. Les arguments que nous avons rassemblés au cours des sections précédentes ont montré que cette idée est compatible avec des hypothèses plus générales sur le fonctionnement cognitif (arguments cognitifs et psycholinguistiques), qu'elle permet d'intégrer des emplois considérés quelquefois comme problématiques (arguments linguistiques) et qu'elle pourra s'adapter à la prise en compte d'autres types de références, comme la référence spatiale ou la référence événementielle. Mais à travers notre argumentation, nous avons également eu l'occasion de mettre en lumière un certain nombre de caractéristiques qu'un domaine de référence doit posséder pour pouvoir répondre à tous les besoins du calcul référentiel.

6.4.1 Un ensemble partitionnable...

En premier lieu, nous sommes partie du principe qu'un domaine de référence regroupe un ensemble d'entités contextuelles, à l'intérieur duquel une expression référentielle sélectionne, sur un critère de différenciation, un référent parmi plusieurs alternatives. Cela signifie qu'il s'agit d'une structure ensembliste pouvant être partitionnée et ce en fonction de différents critères : nous avons vu qu'une expression référentielle peut sélectionner son référent en vertu d'informations prédicatives (indéfinis), de sa description (définis) ou d'une saillance externe du référent (démonstratifs et pronoms). En ce qui concerne le pouvoir distinctif d'une description, elle peut s'appuyer sur le type des objets du domaine, mais aussi sur des propriétés intrinsèques, comme la taille ou la couleur d'un objet, ou sur des propriétés relationnelles, comme la position par rapport à d'autres objets. Un domaine de référence regroupe donc des entités qui sont susceptibles de se distinguer par au moins une propriété, leur critère de différenciation.

Les partitions d'un domaine peuvent être pré-existantes à une opération référentielle, par exemple lorsqu'elles sont le résultat d'interprétations précédentes ou lorsqu'elles reflètent des caractéristiques perceptives des entités visibles à l'écran. Mais une partition peut aussi être créée au cours de

l'interprétation, à condition que des connaissances conceptuelles ou encyclopédiques le permettent. L'interprétation des « anaphores associatives » s'appuie, par exemple, sur des connaissances sur une éventuelle décomposition – c'est-à-dire partition – d'un domaine donné.

La caractéristique fondamentale des domaines de référence est donc d'organiser les entités contextuelles dans des partitions. Celles-ci permettent de distinguer individuellement les entités selon une propriété particulière, le critère de différenciation. Le critère de différenciation n'est pas obligatoirement donné par le discours : des entités d'un même domaine peuvent se distinguer selon des critères visuels ou selon des connaissances encyclopédiques sur leur rôle ou leur localisation.

6.4.2 ... selon différents points de vue

Un domaine de référence n'est pas limité à une seule partition de ses éléments. Nous avons déjà souligné à plusieurs reprises la nécessité de pouvoir tenir compte de différentes perspectives d'un locuteur sur son environnement de communication. Nous considérons alors que chaque partition modélise un point de vue possible sur les éléments du domaine de référence. Ces différentes vues sont reliées à différents aspects compositionnels (spatiaux, temporels, fonctionnels,...) du domaine. Le fait important lié au calcul référentiel est que chaque point de vue particulier sur un domaine et donc chaque partition, fournit des accès référentiels spécifiques aux éléments.

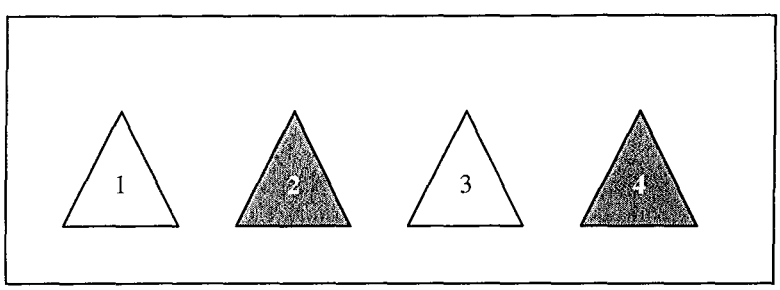


Figure 24 : Différentes possibilités de partition d'un même domaine

Pour donner un exemple de plusieurs partitions perceptives, prenons l'ensemble des objets de la Figure 24 : ce domaine, correspondant à la totalité des objets visibles, peut être vu sous deux angles qui se traduisent par deux possibilités de partition. La première est une partition en deux éléments sur la base des couleurs : {{1,3}, {2,4}}. La seconde est une partition en quatre éléments sur la base de la position horizontale {{1},{2},{3},{4}}. Ce qui est important, c'est que les expressions référentielles disponibles pour extraire des éléments du domaine ne sont pas les mêmes selon la partition active. Un accès référentiel à un élément vu à travers la première partition passera par l'emploi d'une expression liée à l'axiologie des couleurs : *les noirs*, *les blancs*, *les foncés* etc. Un accès référentiel à un élément vu à travers la deuxième partition passera par l'emploi d'une expression liée à l'axiologie de l'alignement horizontal : *les deux premiers*, *celui de gauche*, *le dernier* etc.

Cependant, l'observation suivante montre que le maintien de plusieurs partitions, s'il permet effectivement une représentation « multi-dimensionnelle » d'un domaine, a aussi un coût cognitif qui doit être justifié :

(118) Dans une classe de 20 élèves, *certains* font de la natation et *certains* font du vélo.

Selon F. Corblin (1987 : 37), en (118), « *chacun des indéfinis successivement considérés opère sur la classe entière* ». Autrement dit, il n'y a pas de raison de supposer que les cyclistes sont des non-nageurs, puisque faire de la natation et faire du vélo ne sont pas des propriétés mutuellement exclusives. Le domaine de référence, correspondant ici aux *20 élèves*, contiendra alors deux partitions,

une séparant les élèves cyclistes des élèves non-cyclistes et une autre séparant les élèves nageurs des élèves non-nageurs. Ensuite, une reprise par *les nageurs* activera la partition « natation » alors qu'une reprise par *les cyclistes* activera la partition « cyclisme ». Le même domaine permet donc, à travers ses différentes partitions, l'accès à des sous-ensembles différents, mais non mutuellement exclusifs. Ceci étant, F. Corblin (ibid : 38) constate tout de même une lecture préférentiellement disjonctive. Il l'attribue à une norme discursive consistant à considérer comme différent tout ce qui n'est pas linguistiquement marqué comme identique. Ce constat s'explique dans notre cadre par une économie de représentation : continuer à extraire d'une partition existante coûterait moins cher en calcul et en espace de mémoire que d'en créer une nouvelle.

6.4.3 ... permettant de profiler certains éléments par une focalisation

Pour que l'extraction référentielle puisse fonctionner correctement pour les expressions démonstratives et pronominales, il doit y avoir la possibilité de pouvoir mettre en premier plan (« focaliser ») un ou plusieurs éléments d'une partition. Nous avons effectivement retenu que l'interprétation de telles expressions nécessite un domaine contenant au moins un élément saillant. Nous définissons la saillance comme le résultat d'une structuration domaniale précédente, en particulier par une opération de focalisation. La focalisation a lieu lors d'une extraction d'un élément par une mention discursive, mais aussi lors d'une caractérisation perceptive exceptionnelle (par la taille, une couleur signalétique, un clignotement,...), ou encore par une désignation gestuelle. Nous reprenons donc à notre compte la distinction entre *base* et *profile* de la grammaire cognitive et le point de vue de J. Moulton et al. (1994) qui considèrent que le rôle d'une mention ou d'un geste consiste à imposer une *figure* face à un *arrière-plan*, fourni ici par le domaine de référence.

7 Le modèle du contexte

7.1 Introduction

Au chapitre précédent, nous avons présenté les domaines de référence en tant qu'entités de base d'un modèle contextuel adapté à la résolution de la référence. Le bilan de la fin de ce chapitre nous a amenée à définir leurs caractéristiques essentielles : ensembles partitionnables selon différents points de vue et permettant de focaliser certains éléments. En cela, nous avons formulé un début de réponse aux interrogations soulevées lors de la synthèse des modélisations contextuelles existantes, présentée dans à la fin du chapitre 4. Pour mémoire, ces interrogations portaient premièrement sur la nature des entités composant le modèle contextuel, deuxièmement sur la structuration dynamique du contexte en domaines d'accessibilité et troisièmement sur le rôle accordé aux expressions référentielles elles-mêmes.

Le présent chapitre a pour but de répondre de façon plus détaillée aux deux premières questions en abordant les domaines de référence, introduits au chapitre précédent, sous l'angle de leur modélisation. La réponse à la première question s'inspirera de la théorie des représentations mentales (Reboul et al., 1998), développée initialement pour répondre à des problèmes spécifiques de la référence dans le dialogue homme-machine, mais suffisamment ouverte pour permettre une modélisation plus générale de phénomènes référentiels. Nous verrons en particulier que la structure des représentations mentales est compatible avec les contraintes structurelles définies sur les domaines de référence (section 7.2). La deuxième question, celle de la structuration dynamique du contexte, sera abordée à travers l'opération de groupement, permettant de créer de nouveaux domaines de référence au cours du dialogue et ceci en fonction de critères discursifs et perceptifs (7.3).

Enfin, le chapitre suivant (chapitre 8) abordera la troisième question : il sera consacré au rôle des expressions référentielles elles-mêmes lors du processus d'interprétation référentielle. Nous verrons que ce rôle consiste en l'imposition de contraintes sur la structure d'un domaine d'interprétation compatible, suivie d'une opération de restructuration de celui-ci. L'ensemble des deux chapitres est non seulement organisé autour des trois questions soulevées lors de la synthèse des modélisations existantes (cf. la section 4.4), mais mènera aussi à la construction d'un tableau qui donnera une vision synthétique de l'interprétation de différents types d'expressions référentielles dans différentes structures contextuelles (cf. le Tableau 16). Le présent chapitre permettra d'en instancier les lignes, en spécifiant différentes structures contextuelles possibles. Le chapitre suivant permettra d'en instancier les colonnes, en spécifiant le rôle accordé aux expressions référentielles. Le corps du tableau – les prédictions possibles – fera l'objet d'un chapitre à part (chapitre 8).

Structures contextuelles	*Expression indéfinie*	*Expression définie*	*Expression pronominale*	*Expression démonstrative*
Structure 1	Prédictions sur l'emploi des expressions indéfinies	Prédictions sur l'emploi des expressions définies	Prédictions sur l'emploi des expressions pronominales	Prédictions sur l'emploi des expressions démonstratives
Structure 2				
Structure 3				

Tableau 16 : Structure du tableau synthétique sur la compatibilité entre structures contextuelles et expressions référentielles

7.2 Entités du contexte : les représentations mentales comme domaines de référence

7.2.1 La théorie des représentations mentales

Depuis les années 1970, l'intérêt croissant pour des modèles de compréhension, que ce soit dans une optique psycholinguistique (Kintsch, 1974; Sanford et Garrod, 1982; Johnson-Laird, 1983, 1988) ou computationnelle (Karttunen, 1976 ; Webber, 1978 ; Grosz, 1981) a soulevé la question de la représentation mentale des informations véhiculées par un discours. P. Johnson-Laird (1988), qui situe cette interrogation par rapport aux travaux de sémantique, logique et psycholinguistique de l'époque, formule la problématique ainsi : « *Les réseaux [sémantiques] sont circulaires, en supposant que la signification consiste simplement à mettre un ensemble de symboles en relation avec un autre.* » (p. 61). « *Les logiciens n'ont fait que relier le langage à des modèles sous diverses formes et les psychologues ne l'ont relié qu'à lui-même. Or, ce dont il s'agit réellement, c'est de montrer comment le langage se rapporte au monde par l'intermédiaire de l'esprit.* » (p. 66). En partant du postulat qu'une théorie qui relie les mots au monde permet aussi de relier les mots entre eux, il défend alors l'hypothèse qu'une assertion réfère à un état de choses, ou plus précisément, dès lors qu'un auditeur est capable d'imaginer cet état de choses et d'en faire évoluer l'image mentale, à la représentation mentale que se fait l'auditeur de cet état de choses. Pour P. Johnson-Laird, la dénotation d'une phrase est donc un modèle mental représentant l'état de choses particulier auquel la phrase réfère.

D'autres travaux, plus particulièrement consacrés à des problèmes (co)référentiels, font, sous différentes formes, des propositions comparables : L. Karttunen (1976), dans le but de concevoir une machine capable de comprendre un discours et d'en garder le contenu en mémoire, introduit l'idée des fichiers qui sauvegarderaient des informations sur des référents discursifs. Cette idée est également développée par I. Heim (1982), à travers la sémantique du changement de fichiers. Aussi bien L. Karttunen que I. Heim considèrent qu'une description indéfinie crée un nouveau fichier pour son référent, alors qu'une description définie ré-identifie son référent dans un des fichiers existants. Enfin, la DRT (Kamp, 1981; Kamp et Reyle, 1993) peut être considérée comme une approche représentationnelle, dans la mesure où elle construit des représentations du discours qui vont au-delà d'une simple représentation des conditions de vérité. Un des avantages de la DRT est effectivement de pouvoir prédire, pour deux énoncés partageant les mêmes conditions de vérité, certaines différences dans les possibilités des reprises et ceci sur la base des représentations construites pour ces énoncés.

Le point commun de ces propositions consiste donc à considérer les processus de compréhension comme des opérations de construction et de mise à jour de « représentations mentales » pour un état de choses. Mais ce qui leur est également commun, c'est le fait de limiter ces processus à l'établissement de relations entre informations discursives. La question de l'ancrage de ces représentations sur des entités réelles est seulement abordée à propos des problèmes particuliers que pose l'interprétation du nom propre. Or, la finalité d'un module de calcul référentiel pour le dialogue homme-machine est précisément d'aboutir à l'identification des objets particuliers dont parle le locuteur. Comme nous l'avons déjà signalé lors de la synthèse des modèles contextuels (section 4.4), la seule mise à disposition d'entités discursives est insuffisante par rapport à nos objectifs et ceci pour deux raisons : d'une part, le référent d'une expression peut ne pas être mentionné auparavant et d'autre part, les informations nécessaires à la détermination d'un référent ne sont pas exclusivement de nature discursive. Elles sont aussi de nature perceptive et encyclopédique.

En réponse à notre interrogation sur les entités composant le contexte, la théorie des représentations mentales (Reboul et al., 1997 ; Reboul et Moeschler, 1998) – issue d'un projet de recherche commun

entre le LORIA et le LIMSI (CERVICAL[67]) – propose une modélisation des référents qui répond à ces critères. Le calcul référentiel n'y est pas considéré comme un processus de recherche du « bon antécédent discursif », mais comme un processus d'assignation d'une représentation mentale identifiante à une expression référentielle. Dans cette perspective, une représentation mentale est conçue comme une charnière cognitive entre une expression référentielle et son référent. Il s'agit d'une représentation identifiante pour une entité contextuelle particulière, capable d'intégrer toutes les informations hétérogènes (linguistiques, encyclopédiques et perceptives) susceptibles de contribuer à sa (re-)identification lors d'un acte de référence.

Nous considérons que la différence entre cette représentation cognitive et l'objet de l'application qu'elle représente réside dans le fait que la représentation mentale représente une vue partielle, particulière et évolutive de l'objet, vue qui correspond idéalement à la perception qu'a un locuteur à un moment donné sur l'objet représenté. De ce fait, la représentation mentale est supposée être prédictive quant à des accès référentiels privilégiés. La théorie des représentations mentales présente, de ce point de vue, des similitudes avec la conception représentationnelle et constructiviste de la référence défendue par D. Apothéloz et M.-J. Reichler-Béguelin (1995) : ces auteurs renoncent à une conception des référents en tant que « choses », au profit d'une conception en termes d'« objets-du-discours ». Ceux-ci sont modélisables sous la forme d'un ensemble évolutif d'informations, incluses dans le savoir partagé des interlocuteurs et conditionnant contextuellement les désignations. Cependant, la différence entre ces « objets-du-discours » et notre conception des représentations mentales réside dans le fait que les représentations mentales ne proviennent pas seulement de l'activité langagière, mais aussi d'autres activités cognitives (la perception, en particulier).

Nous montrerons dans la suite de cette section que les caractéristiques des représentations mentales sont compatibles avec les contraintes que nous avons définies, à la fin du chapitre précédent (section 6.4), sur la structure des domaines de référence.

7.2.2 Des représentations mentales aux domaines de référence

La section précédente a répondu à notre première question – concernant les entités du contexte – en proposant, sur la base de la théorie des représentations mentales, que le contexte doit être composé de représentations mentales. Mais cette réponse n'est que partielle, dans la mesure où nous avons fourni, au chapitre 6, des arguments en faveur de l'hypothèse selon laquelle tout acte de référence passe par l'identification d'un référent dans un domaine de référence, qui peut être lui-même un sous-ensemble de la totalité des entités contextuelles. Cela signifie alors que le contexte doit (aussi ?) contenir des domaines de référence.

La solution consiste à considérer qu'une représentation mentale introduite pour représenter une entité contextuelle peut, de façon récursive, fonctionner comme domaine de référence dans la suite du discours. D'abord, d'un point de vue linguistique, cette solution correspond à des observations telles qu'en (119) ou (120) : le référent de l'expression *la première* est en effet extrait de la représentation mentale introduite auparavant par *des pyramides* et celui de *un toit* de celle introduite par *la maison* :

(119) I_1 bon, alors il faut faire des pyramides, donc je vais essayer...
M_1 ouais
I_2 voilà *la première* (C5Egypte)

(120) I_1 maintenant je vais faire la maison
I_2 faut mettre *un toit* dessus, c'est un petit triangle (C6Route)

[67] « Communication Et Référence : Vers une Informatique Collaborant Avec la Linguistique », financé par le GIS Sciences Cognitives

Ensuite, cette solution semble justifiée d'un point de vue cognitif par le fait qu'il paraît effectivement difficile d'envisager la gestion parallèle de deux structures différentes (représentations mentales pour les référents ; domaines de référence pour les contextes locaux). Enfin, d'un point de vue structurel, les représentations mentales présentent toutes les caractéristiques nécessaires aux domaines de référence, caractéristiques que nous avons mises en évidence à la fin du chapitre précédent. Nous allons voir ci-dessous qu'elles sont effectivement capables de représenter des ensembles d'objets, partitionnables selon différents points de vue et permettant de focaliser certains éléments.

La structure de base d'une représentation mentale (RM) est celle d'un schéma attribut-valeur (Figure 25). Ce schéma, qui s'inspire des propositions de CERVICAL sans les reprendre entièrement, comporte au minimum un identificateur unique (introduit par @) et un pointeur sur le type de l'entité représentée (attribut type). Par ailleurs, une représentation mentale peut comporter d'autres champs : entre autres, une entrée indiquant la cardinalité du référent[68], une entrée pour les propriétés spécifiques au référent, comme ses caractéristiques physiques et son implication dans des événements et une entrée visuelle, pour son aspect perceptif et son orientation par défaut. Un point faible des propositions de CERVICAL est que la gestion de ces champs n'a pas été suffisamment détaillée. Nous montrerons par la suite qu'une partie de ces informations – pourtant indispensables au calcul référentiel – peut être gérée dans des partitions. Les partitions, qui donnent des informations compositionnelles sur les entités représentées, sont une autre composante d'une représentation mentale. Dans notre modélisation, nous leur attribuons un rôle essentiel car c'est à travers elles que nous allons modéliser différents points de vue permettant de prédire des accès référentiels privilégiés aux entités contextuelles.

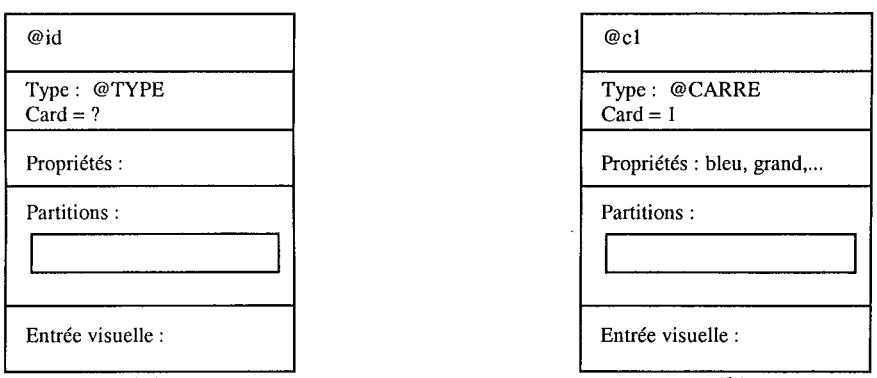


Figure 25 : Canevas général d'une RM (a) et instanciation pour « un grand carré bleu » (b)

La ou les partitions d'une représentation mentale modélisent des décompositions possibles de l'objet représenté. Un des principes méréologiques stipule que les parties d'un objet seront toujours des objets (Guarino, 1998). Par conséquent, une partition correspondra à un ensemble de pointeurs sur d'autres représentations mentales qui, conformément aux caractéristiques ensemblistes d'une partition, doivent être des entités disjointes.

La gestion de l'accès aux éléments d'une partition nécessite l'existence d'un critère de différenciation. Il s'agit d'un attribut commun à tous les éléments d'une partition, mais pour lequel ces éléments prennent des valeurs différentes. Les critères de différenciation peuvent être d'origine perceptive, discursive ou encyclopédique (connaissances sur la composition d'un objet). L'existence d'au moins un critère de différenciation justifie non seulement l'existence de la partition (pour que l'on puisse partitionner un ensemble, les parties doivent être différenciables), mais permet aussi de

[68] La cardinalité d'une RM peut-être supérieur à 1, lorsque celle-ci représente un ensemble (cf. la section 7.3).

prédire certaines caractéristiques d'une expression qui puisse servir à désigner un élément de la partition. Ainsi, une représentation mentale pour un ensemble de deux billes (une bille rouge et une bille bleue) peut être partitionnée selon le critère de différenciation *COULEUR*. Cela permet de supposer que l'accès référentiel à une de ces billes se fait préférentiellement sur cette propriété. Comme une entité peut se décomposer généralement selon plusieurs points de vue (la même représentation pour les deux billes peut comporter une deuxième partition selon le critère de différenciation *POSITION_HORIZONTALE* en une bille à gauche et une bille à droite), une représentation mentale peut comporter plusieurs partitions.

Par ailleurs, les éléments d'une partition peuvent, pour certains critères de différenciation, être ordonnés. C'est par exemple le cas pour une partition d'un groupe d'éléments alignés horizontalement. Le critère de différenciation *HORIZONTAL_POSITION* non seulement distingue les éléments alignés par leur coordonnée X, mais impose aussi un ordre entre ces éléments, ordre dont on peut supposer qu'il correspond à l'ordre de lecture par défaut, c'est-à-dire de la gauche vers la droite. Cette caractéristique des partitions est particulièrement importante pour l'interprétation d'expressions s'appuyant sur un ordre temporel, spatial ou mentionnel (cf. Corblin, 1999 ; Laborde, 1999).

Enfin, à l'intérieur d'une partition, un élément au plus peut être focalisé. Il s'agit généralement de la représentation pointant sur le dernier référent créé ou identifié à l'intérieur du domaine représenté par la représentation mentale en question. Comme cette focalisation dépend des opérations de restructuration contextuelle (groupement et extraction), nous reviendrons sur le calcul focal lors de la présentation de ces opérations.

7.2.3 Les représentations mentales génériques

La première question liée à la gestion d'un modèle contextuel sous forme de représentations mentales est celle de la création de nouvelles représentations. Cette opération, déclenchée par l'introduction d'un nouveau référent dans l'univers discursif ou perceptif, correspond à une instanciation de représentations mentales génériques. Nous faisons donc la différence entre représentations spécifiques et représentations génériques : les premières sont des instances des secondes.

Cela signifie que parallèlement aux représentations spécifiques pour les entités effectivement introduites dans le contexte (que cela soit par une mention discursive ou par la perception), nous supposons l'existence d'un réseau conceptuel modélisé sous forme de représentations mentales génériques. Ce réseau correspond à la composante « encyclopédique » dont un système de dialogue ne peut se passer et dont la fonction est de représenter et de structurer l'univers de référence dans lequel le système doit évoluer (Luzzati, 1995). Nous considérons cette composante comme indépendante de la langue, dans la mesure où elle modélise des concepts. En revanche, les liens entre lexique et concepts sont de nature linguistique, dans la mesure où ils sont régis par la sémantique de chaque langue (Figure 26).

La raison d'être des représentations mentales génériques est la nécessité de représenter des connaissances encyclopédiques intervenant dans la résolution référentielle. La disponibilité de ces connaissances a d'ailleurs été présupposée par les concepteurs des représentations mentales, en particulier à travers l'invocation d'un mécanisme d'héritage entre informations conceptuelles et représentations mentales spécifiques : « *La base de l'entrée encyclopédique est l'ensemble des informations conceptuelles héritées, par défaut, de la catégorie à laquelle le référent correspond.* » (Reboul et al, 1998, chap. II/4).

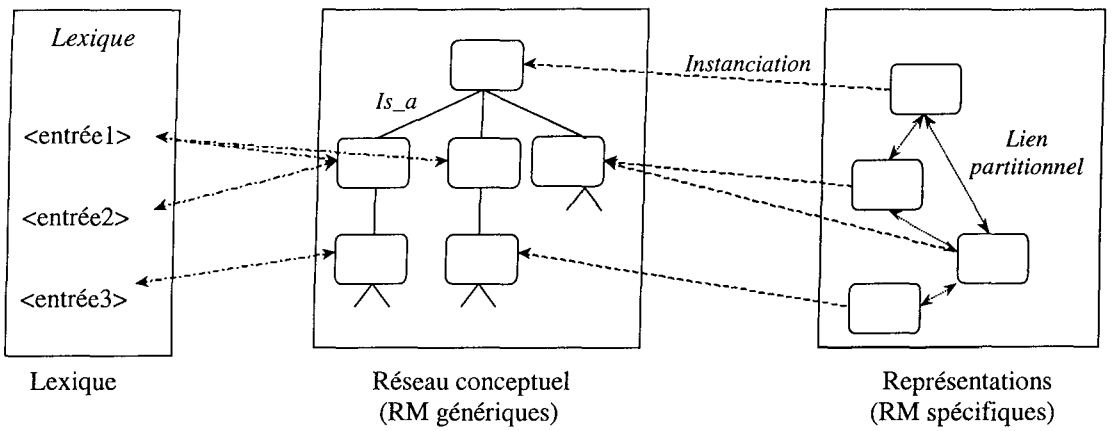


Figure 26 : Relation entre lexique, RM génériques et RM spécifiques

Nous considérons alors que la structuration interne des représentations mentales génériques n'est pas fondamentalement différente de celle des représentations mentales spécifiques et ceci pour deux raisons : d'une part, cela nous permet de maintenir un format unique de représentation pour les connaissances encyclopédiques et les connaissances contextuelles particulières, solution qui semble d'ailleurs justifiée par le fait que les connaissances mises en jeu par les processus de résolution référentielle ne sont pas seulement de nature discursive, mais aussi de nature encyclopédique. Par ailleurs, maintenir le canevas général des représentations mentales représente évidemment un avantage pour la mise en œuvre informatique, dans la mesure où l'instanciation d'un objet à partir de la définition d'une classe correspond à une opération de base dans la programmation objet. Étant donné notre hypothèse selon laquelle les mécanismes référentiels sont essentiellement liés à la gestion des partitions, nous attirons l'attention sur le rôle important que jouent les partitions des représentations mentales génériques : ce sont elles qui donnent les informations méréologiques sur la composition habituelle des objets et c'est sur la base de ces partitions que pourront être interprétées, en particulier, les anaphores associatives.

Suite à la spécification de la structure d'une représentation mentale générique, s'impose une réflexion sur l'organisation de ces représentations entre elles. Les liens entre représentations génériques sont, eux aussi, subordonnés à notre objectif qui est la modélisation des connaissances encyclopédiques nécessaires au calcul référentiel. Plus précisément, ce réseau doit constituer le support pour des calculs inférentiels impliqués dans la (ré-)identification d'objets. Avant tout, il convient de s'interroger sur le type des connaissances intervenant dans ces calculs. Les exemples suivants peuvent servir de support à la réflexion :

(121) Ingrédients : 500 g de congre coupé en tranches, 4 carottes, 3 oignons, 2 poireaux [...]. Épluchez et lavez *les légumes.* (http://www.cuisine.free)

(122) Laissez mijoter pendant 25 minutes puis égouttez les tranches de congre avec une écumoire. Retirez-en *la peau* et *les arêtes.* (http://www.cuisine.free)

(123) John Lennon a été assassiné hier soir à New York. *Le meurtrier*, un détraqué nommé M. Chapman, a tiré sur le chanteur alors qu'il rentrait chez lui. (Manuélian, 1999)

L'exemple (121) illustre la ré-identification d'un ensemble d'objets à travers un type subsumant, linguistiquement réalisée par l'emploi d'un hyperonyme. Ce raisonnement nécessite comme support une hiérarchie de types. L'exemple (122) montre l'importance de la prise en compte de relations méréologiques, réalisées linguistiquement par des anaphores associatives. L'introduction d'une entité peut en effet donner accès à ses parties, comme en (122). Enfin, l'introduction d'un événement permet

de référer à ses participants (123). La modélisation de ces calculs demande alors que ces connaissances soient représentées sous une forme ou une autre.

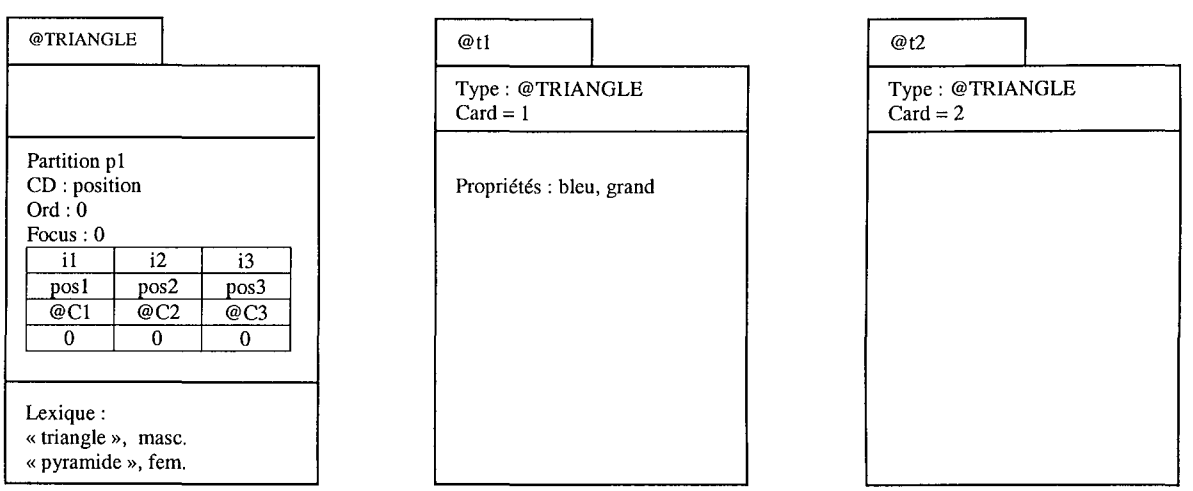
Nous proposons donc de construire le réseau des représentations mentales génériques autour d'un squelette qui est une hiérarchie de types. En plus des liens *is_a* entre types et sous-types, nous proposons d'introduire d'autres liens entre représentations mentales génériques. Un lien *tout–partie* permettra, à travers les partitions, l'accès d'un objet à ses parties. En théorie, il est intéressant de noter que certains de ces liens donnent lieu à des mécanismes d'héritage très particuliers : alors qu'il est impossible de considérer l'héritage comme une opération générale le long des liens *tout–partie*, on observe que certaines propriétés peuvent se transmettre par ces liens : il en est ainsi, par exemple, pour la couleur d'une fleur qui se transmet aux pétales. En pratique, nous ne tenons pas encore compte de ces spécificités.

7.2.4 Représentations mentales : bilan, exemples et représentations graphiques

En (124), la mention *un grand triangle bleu* mène à l'introduction d'une nouvelle représentation mentale pour son référent :

(124) il faut prendre *un grand triangle bleu*

Cette représentation (Figure 27b) est elle-même une instanciation de la représentation mentale générique pour le type *TRIANGLE* (Figure 27a). Elle comporte un identificateur unique @t1 et son entrée *TYPE* pointe sur la représentation générique @TRIANGLE. Comme l'expression comporte des indications sur des propriétés spécifiques du référent, ces informations entrent dans le champ des propriétés. La représentation hérite de son type une partition qui donne accès aux composantes de l'entité : un triangle, par exemple, est composé de trois côtés. Le critère de différenciation de cette partition est la position spatiale des ces éléments, les éléments de la partition ne sont pas ordonnés et aucun de ces éléments n'est focalisé. Il convient de préciser que cette partition, bien qu'héritée, n'est pas instanciée dans la représentation mentale spécifique @t1 tant qu'il n'y a pas d'accès référentiel à un de ses éléments et ceci pour des raisons d'économie de représentation.



(a) RM générique TRIANGLE *(b) RM « un grand triangle bleu »* *(c) RM « deux triangles »*

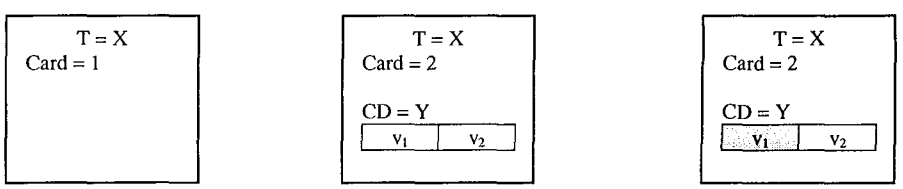
Figure 27 : Exemples de représentation graphiques de RM

L'exemple (125) illustre une autre instanciation :

(125) Prends deux triangles.

L'expression *deux triangles* mène à la création d'une représentation pour un ensemble d'éléments de type @TRIANGLE dont la cardinalité est *deux* et dont aucune propriété spécifique n'est connue (Figure 27c). Il est à noter que bien qu'il s'agisse d'un ensemble, cette représentation n'est pas partitionnée : étant donné le point de vue sous lequel cet ensemble a été présenté, il est en effet impossible de distinguer individuellement les deux référents.

Les représentations graphiques de la Figure 27 sont des représentations exhaustives, correspondant à la totalité de la structure des données définie pour les représentations mentales. La modélisation informatique de cette structure est présentée de façon détaillée dans le chapitre 10 et sa connaissance est nécessaire à la compréhension des algorithmes spécifiant les opérations définies sur les représentations mentales. Il est néanmoins possible de suivre les grandes lignes de notre modélisation indépendamment des spécifications informatiques. Nous utiliserons dans les chapitres suivants des représentations graphiques simplifiées, telles que présentées dans la Figure 28. Elles contiennent seulement les informations essentielles à la modélisation, c'est-à-dire le type des éléments représentés (T), la cardinalité et des informations sur la partition : critère de différenciation (CD), valeur du critère de différenciation pour les éléments de la partition (v) et la focalisation (partie grise de la partition). La Figure 28a donne l'exemple d'une RM sans partition, avec information sur le type. La Figure 28b représente une RM avec partition et critère de différenciation. La Figure 28c est une RM avec une partition contenant un élément focalisé.



(a) (b) (c)

Figure 28 : Exemples de représentations graphiques simplifiées

7.3 Structurations locales du contexte : l'opération de groupement

7.3.1 Motivation de l'opération de groupement

La section précédente a répondu à la question des entités de base composant le contexte (représentations mentales des domaines de référence), mais elle n'a pas abordé la question de la mise à jour dynamique du contexte autrement qu'à travers l'opération de création d'une nouvelle représentation mentale, lors de la mention ou perception d'une nouvelle entité. Or, l'introduction de nouvelles représentations mentales ne suffit pas à elle seule à créer des structures contextuelles prédictives pour l'interprétation des expressions à venir : cette solution mènerait en effet à une structure contextuelle « plate », proche des structures obtenues par la DRT (en faisant abstraction des structures particulières pour des conditionnelles, négations et modalisations). Comme nous l'avons déjà souligné à plusieurs reprises, nous avons besoin de créer dynamiquement des domaines de référence locaux : cette création se réalise par une opération de groupement.

Les motivations principales pour regrouper des représentations mentales individuelles en représentations plus complexes sont les suivantes :

Le groupement crée des domaines fonctionnant eux-mêmes en tant que référents : en (126), l'expression *les deux* réfère effectivement à l'ensemble des référents introduits par *une (à ce niveau)* et *une (là)* :

(126) I_1 alors là c'est censé être des pyramides dans le désert
I_2 euh ... *une à ce niveau* et *une là*
I_3 euh il faut mettre une barre entre *les deux* (C6Pyramides)

Le groupement crée des domaines, dont la prise en compte est nécessaire pour opérer des extractions (*le triangle rouge*) et pour interpréter des expressions telles que *autre* (127) :

(127) Prends *un triangle rouge* et *un triangle vert*. Mets *le triangle rouge* sur la droite. Supprime *l'autre*.

Le groupement crée des domaines, dont les entités sont, sous certaines conditions, inaccessibles individuellement pour des reprises pronominales (128) :

(128) ? Prends *un triangle rouge* et *une barre verte*. Mets *le* sur la droite.

La question suivante porte, par conséquent, sur les facteurs susceptibles de déclencher ces opérations de groupement.

7.3.2 Type de déclencheurs et niveaux de groupement

Lors de la synthèse sur les modélisations contextuelles existantes (en particulier dans la section 4.4.2), nous avons relevé les principaux critères considérés comme des indices pour un groupement des entités du modèle contextuel : il s'agit de facteurs linguistiques (essentiellement syntaxiques), discursifs (structures rhétoriques), intentionnels (structure de la tâche) et cognitifs (récence et saillance). Cette section les rappelle brièvement avant d'en proposer une vue synthétique et de justifier nos propres choix de modélisation.

Pour ce qui est des facteurs linguistiques, l'énumération et la coordination par la conjonction *et* sont les critères les plus importants. D'un point de vue linguistique, on peut en effet souligner, avec G. Kleiber (1986 : 77), que « *ce ne sont pas deux nouveaux référents qui sont en fait introduits, mais bien un seul référent, en l'occurrence l'ensemble des référents constitué par la coordination* ». De fait, certaines modélisations proposent des mécanismes de groupement adéquats : la DRT (Kamp et Reyle, 1993), par exemple, prévoit une opération de sommation des référents du discours introduits par un groupe nominal coordonné. Cela permet de reprendre, dans la suite, ce complexe par un pronom au pluriel. En revanche nous verrons que le problème n'est pas entièrement résolu, car l'accès individuel aux référents du complexe n'est pas pour autant bloqué. Or, comme le suggère G. Kleiber en parlant d'« *un seul référent* », une reprise pronominale individuelle dans un exemple tel que (128) semble très difficile.

En plus de la coordination, d'autres facteurs linguistiques contribuent à la création d'ensembles référentiels. Parmi ceux-ci, on trouve la structure argumentale d'un prédicat (Fillmore, 1968 ; Saint-Dizier, 1998). En effet, beaucoup de modélisations – dont l'algorithme de C. Sidner (1979), la DRT ou la théorie du Centrage (Grosz et al., 1995) – prévoient, d'une façon ou d'une autre, la possibilité de regrouper les participants d'une même éventualité. La théorie du Centrage en fait même son principe de structuration contextuelle principal. Parmi les travaux moins formalisés, on pourra mentionner la grammaire cognitive (Langacker, 1986, 1991). On y rencontre en effet cette même idée dans la modélisation des processus de compréhension : ceux-ci consistent en une élaboration de positions sous-spécifiées (une sorte de « remplissage ») dans des schémas abstraits représentant des éventualités (« relations « ou « processus ») par des schémas plus concrets (« choses »). La représentation finale est

un schéma composite qui profile un ou plusieurs participants d'une éventualité. Enfin, il est intéressant de constater qu'un certain nombre d'expériences psycholinguistiques sur la saillance relative de référents du discours reposent entièrement sur l'hypothèse d'un groupement « cognitif » des participants d'un événement (Schnedecker et Bianco, 1995). Nous reviendrons sur la grammaire cognitive et les expériences dans la section suivante, portant sur la focalisation.

A un niveau supérieur au segment phrastique élémentaire, on trouve des propositions de regroupement sur critères syntaxiques et discursifs. Les critères syntaxiques sont formulés de façon élégante par la DRT : en fonction de certaines constructions (conditionnelles, temporelles et aspectuelles), des segments forment des complexes qui restreignent l'accessibilité de leurs référents discursifs. L'avantage de ces conditions de groupement est leur calculabilité. L'inconvénient par rapport à notre objectif, le dialogue homme-machine, est leur taux d'occurrence très faible dans notre corpus de dialogues ainsi que leur dépendance d'une grammaticalité normative, pas toujours assurée dans la langue orale. Parmi les travaux les plus représentatifs des groupements discursifs, nous avons présenté ceux de B. Webber (1991) et de N. Asher (1993) : leur objectif est de construire une représentation hiérarchique du discours, reflétant des relations de coordination ou de subordination entre des segments. Idéalement, cette représentation serait prédictive quant à la reprise des entités discursives, mais dans la section 4.3.6, nous avons formulé un certain nombre de réserves à cet égard. Nous considérons en effet que les principaux points problématiques sont la calculabilité de cette structure (rappelons que B. Webber suppose, pour résoudre le problème, rien de moins qu'un oracle) et son pouvoir prédictif effectif qui nous semble limité.

Enfin, les travaux de B. Grosz (1977) ont fait apparaître que la structure de la tâche, c'est-à-dire une hiérarchie intentionnelle, permet de créer des espaces attentionnels, limitant l'ensemble de recherche pour les antécédents d'une expression donnée. Or, comme l'a noté B. Gaiffe (1992), l'application de cette proposition suppose une tâche fortement hiérarchique, bien définie, connue et exécutée dans un ordre strict par les sujets. Cela est loin d'être le cas pour des tâches tout-venant, mais ce fait ne doit pas empêcher de faire jouer ce facteur, en combinaison avec d'autres, lorsque cela est possible. Pour le corpus Ozkan, il mériterait en effet d'être exploité, car l'ordre de construction des éléments des dessins a été prédéfini (bien que pas toujours respecté) et de plus, la clôture des sous-tâches – pose d'un élément au bon endroit, fin de la construction d'un élément figuratif – est très souvent linguistiquement marquée (*voilà, puis, et ensuite, ben ça y est,...*), comme a pu le constater N. Colineau (1997). La prise en compte de tels facteurs permettrait par exemple de résoudre l'ambiguïté sur l'interprétation du pronom *le* en I_3 de l'exemple (129). Ce pronom peut en effet être considéré comme renvoyant à l'instance désignée auparavant par *le gros triangle*, mais aussi au type *GRAND_TRIANGLE*. Étant donné la marque de clôture (*voilà*), le pronom peut être considéré ici comme ayant moins de probabilité de référer à l'instance qui vient d'être manipulée qu'au type dont relève cette instance[69].

(129) I_1 ensuite on a les pyramides dans le désert
alors il faut que tu prennes les deux euh enfin le triangle
I_2 le gros triangle...
I_3 voilà et tu *le* reprends une deuxième fois un peu décalé par rapport au premier (C8Egypte)

Un résumé synthétique de ces différents facteurs de groupement est donné dans la Figure 29. Nous concevons en effet qu'ils sont complémentaires, dans la mesure où aucun parmi eux n'en exclut d'autres et que les meilleurs systèmes de résolution référentielle seront ceux qui tenteront de les combiner. Cette combinatoire pourra suivre l'ordonnancement hiérarchique suivant : les représentations pour les entités individuelles entreront dans des groupements sur critères syntaxiques

[69] Au chapitre 2, section 2.5, nous avons rapproché cet exemple des « *pay-check-sentences* », caractérisées par la présence d'un pronom divergent.

« locaux », tels que la coordination et l'énumération. Au niveau sémantique, cela concerne des entités jouant un même rôle thématique dans une structure argumentale (a). Ces ensembles ainsi que des entités individuelles se regroupent en tant que participants à une même éventualité (b). Les éventualités se combinent selon des relations discursives (c) et des ensembles d'éventualités peuvent être regroupés pour former des espaces attentionnels (d).

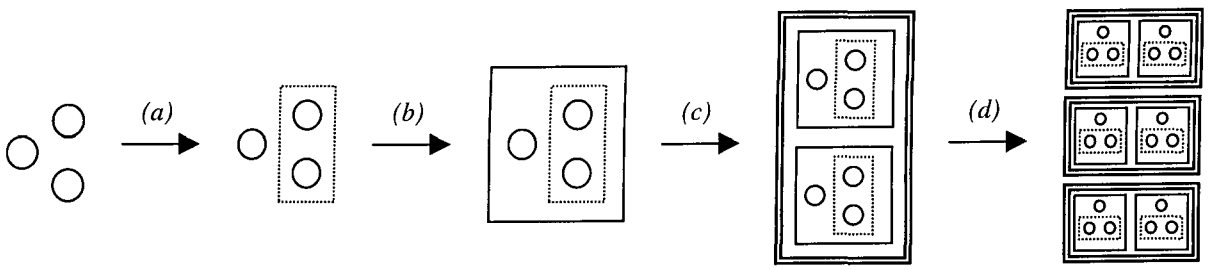


Figure 29 : Synthèse des niveaux de groupement

7.3.3 Solution retenue pour notre modélisation

Si notre synthèse a pu contribuer à donner une vision plus globale des différents facteurs participant à la création de domaines d'interprétation, elle est certainement non exhaustive, comme le montre l'exemple des prépositions spatiales sur lesquelles nous reviendrons ci-dessous. D'autre part, nous avons déjà mentionné à plusieurs reprises les difficultés que pose le calcul automatique de certains de ces facteurs, en particulier à partir du niveau (c) de la Figure 29. Pour ces deux raisons, notre premier souci est de proposer une opération de groupement qui reste paramétrable, à la fois en ce qui concerne les conditions de déclenchement et le calcul de la structure des domaines résultants. Comme le détaillent les algorithmes présentés au chapitre 10, une opération de groupement prend en argument au moins deux représentations mentales et en crée une nouvelle, ce qui implique en particulier le calcul d'un type commun des entités regroupées ainsi que la création d'une partition à l'intérieur de cette nouvelle représentation.

Pour l'instant, nous avons décidé de ne tenir compte, pour notre modélisation, que des niveaux (a) et (b) de la Figure 29. Ce choix nous semble justifié à plusieurs égards : premièrement, dans l'état des connaissances actuelles, il n'est pas possible de calculer automatiquement des regroupements de niveau (c). Ensuite, si cela avait été possible, il aurait fallu gérer des phénomènes dépassant le cadre de ce travail : un groupement entre éventualités implique en effet de réfléchir aux types des entités regroupées, ce qui demande d'abord de disposer d'une représentation pour les éventualités ainsi que d'une hiérarchie de type adaptée. Ensuite, nous n'aurions pas pu tester des prédictions sur ces domaines, car nous nous sommes limitée volontairement à la formulation de prédictions sur l'extraction référentielle à partir d'ensembles d'objets. Par rapport aux données dont nous disposons, il nous semble en effet que la prise en compte des deux premiers niveaux de groupement est prioritaire pour modéliser le calcul référentiel. Il sera par ailleurs intéressant d'explorer jusqu'à quel point un contexte modélisé exclusivement par des ensembles (structurés) d'objets peut rendre compte des processus d'interprétation référentielle. Nous reviendrons sur ce point dans la conclusion du chapitre consacré aux prédictions de notre modèle (section 9.6).

En ce qui concerne les facteurs syntaxiques locaux (niveau (a) de la Figure 29), nous retenons donc les groupes nominaux coordonnés ou énumérés comme déclencheurs d'une opération de groupement. Les données du corpus Ozkan montrent effectivement que les référents de ces syntagmes forment un domaine de référence, pouvant être repris (130) ou servant à des extractions ultérieures (131) :

(130) I_1 alors là c'est censé être des pyramides dans le désert
I_2 euh ... *une à ce niveau* et *une là*
I_3 euh il faut mettre une barre entre *les deux* (C6Pyramides)

(131) I_1 maintenant il faut faire *des sapins et des arbres*
I_2 alors il y a *trois sapins* à peu près à ce niveau-là (C6Forêt)

En ce qui concerne le niveau (b) – le groupement des participants d'une même éventualité – l'opération est déjà plus délicate. Étant donné la nature des actes de langage du corpus Ozkan, qui sont essentiellement des ordres directs ou indirects (Colineau, 1997), le problème posé est celui du rôle du sujet.

(132) I_1 *tu la* mets entre *les deux pyramides* (C7Egypte)

Dans un énoncé tel que (132), un groupement des participants impliquerait de grouper la représentation pour l'interlocuteur avec celle des objets manipulés. Or, étant donné que nous nous sommes limitée à la référence aux objets, en excluant pour l'instant les références déictiques, il nous a semblé qu'un tel groupement n'était pas souhaitable. Force est alors de constater l'existence de groupements « intermédiaires » entre la totalité des arguments d'un prédicat et des groupes coordonnés. Une question qui se pose ici est celle du rôle des prépositions. Le corpus Ozkan, composé essentiellement d'énoncés de positionnement, contient en effet un grand nombre de prépositions spatiales. Il est alors permis de s'interroger sur la façon dont ces propositions contribuent à la création d'ensembles contextuels. Un exemple comme (133) montre que ces prépositions (*sur, à côté*) mettent en relation des référents (*personne, chaise, table*) formant un domaine de référence pour des extractions futures (*la table, la chaise*) :

(133) I_1 donc l'autre dessin c'est *une personne*
I_2 assise donc sur *une chaise* et juste à côté *d'une table*
I_3 donc *la table* est représentée par un grand cercle
I_4 et *la chaise* c'est deux traits donc verticaux (C9Homme)

Là encore, la grammaire cognitive peut fournir une piste : comme nous l'avons déjà mentionné, l'interprétation d'expressions complexes se fait par l'élaboration successive de schémas abstraits. Or, les schémas des prépositions sont précisément conçus comme « mettant en relation » des « choses ». Ainsi, comme le montre la Figure 30a, l'expression *à gauche de* est considérée comme instaurant une relation sur l'axe horizontal entre deux objets. Les schémas des objets en question viennent alors élaborer le schéma de la préposition, ce qui a pour conséquence leur groupement à l'intérieur d'un schéma plus complexe. Parallèlement à cela, nous proposons donc de grouper les représentations mentales des référents reliés par une préposition spatiale, comme dans la Figure 30b.

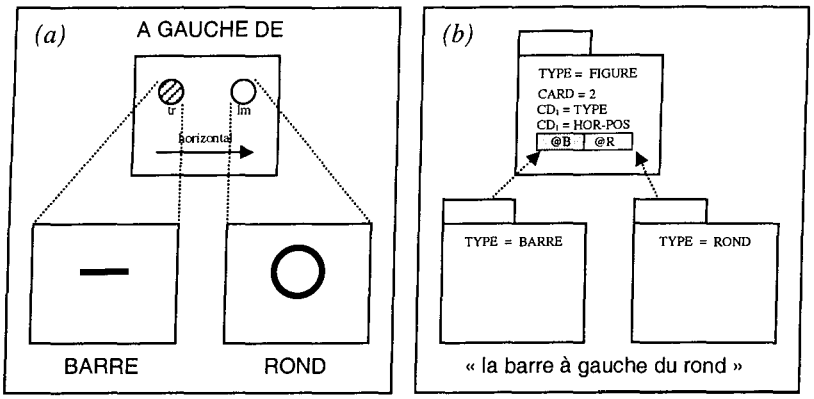


Figure 30 : Interprétation d'une préposition spatiale dans la grammaire cognitive (a) et modélisation par groupement (b)

7.3.4 Le calcul focal

Le calcul d'une représentation résultant d'un groupement implique, comme nous l'avons mentionné ci-dessus, le calcul d'un type commun des entités regroupées ainsi que le calcul d'une partition. Parmi les informations liées à la partition, la structure focale mérite une attention particulière, dans la mesure où elle est un des éléments-clé de la suite de notre modélisation.

Par structure focale, nous entendons le marquage éventuel d'un des éléments d'une partition en tant qu'élément le plus saillant et donc le plus accessible. Cette opération – la focalisation – est évidemment dépendante de la dynamique discursive. Nous avons déjà mentionné qu'elle est effectuée à chaque extraction référentielle, mais aussi lors d'une opération de groupement. Étant donné que nous avons décidé d'inclure certains groupements dans la modélisation, nous sommes obligée de préciser les conditions qui influent sur la structure focale d'un domaine créé par un groupement. Il s'agit donc de décider quel élément appartenant à un domaine issu d'un groupement sera l'élément focalisé.

Cette question a déjà fait l'objet d'un certain nombre de travaux, aussi bien en intelligence artificielle (Hobbs, 1977 ; Sidner, 1979 ; Grosz et al., 1995) qu'en psycholinguistique (Gernsbacher et Hargeaves, 1988; Clifton et Ferreira, 1987; Greene et al., 1992) ou en linguistique cognitive (van Hoek, 1995). Nous présenterons d'abord deux hypothèses psycholinguistiques « extrêmes » et contradictoires, avant de les nuancer et présenter les choix que nous avons faits pour notre modélisation.

Comme le relèvent C. Schnedecker et M. Bianco (1995), des expériences psycholinguistiques ont fait apparaître une forte homologie entre les deux déclencheurs de groupement qui nous intéressent ici : au vu des résultats obtenus, il semblerait en effet que le mode de conjonction – coordination ou prédicat – n'influe pas sur la structure de la représentation mentale résultante. Cependant, il se dégage deux tendances contradictoires, suivant la nature même de cette représentation résultante. Les travaux de M.A. Gernsbacher et al. (1988) suggèrent qu'il y a dissymétrie, dans la mesure où le référent de la première mention garde dans tous les cas le bénéfice de la saillance. Transposé à notre modélisation, cela signifie qu'un domaine issu d'un groupement, qu'il soit déclenché par une coordination ou par le prédicat, comporte toujours un élément focalisé correspondant au référent de la première mention. Contrairement à cette hypothèse, les travaux de C. Clifton et F. Ferreira (1987) et de S. Greene (1992) laissent penser que les référents individuels des entités sont fusionnés en une représentation complexe et ceci indépendamment du mode de conjonction. Cela signifierait que le domaine résultant de notre modélisation ne comporte pas d'élément focalisé. Or, comme le suggère l'étude de C. Schnedecker et

M. Bianco (1995), les deux hypothèses méritent d'être nuancées en fonction du déclencheur du groupement.

En ce qui concerne le déclenchement par la coordination *et*, l'hypothèse qui favorise systématiquement la première mention est en contradiction avec les observations de G. Kleiber (1986). Elle ne parvient effectivement pas à expliquer ni à prédire la difficulté de la reprise pronominale, ni même démonstrative, d'un des éléments du groupement, fût-ce le premier mentionné :

(134) ? Prends *un triangle rouge* et *un triangle vert*. Mets-*le* sur la droite.

(135) ? Prends *un triangle rouge* et *un triangle vert*. Mets *ce triangle* sur la droite.

Il semble donc bien que les groupes coordonnés bloquent l'accès individuel à leur constituants, pour ce qui est des reprises pronominales ou démonstratives. Mentionnons ici que certaines modélisations contextuelles, bien que prévoyant des groupements déclenchés par une coordination, n'arrivent pas non plus à rendre compte de ce blocage, comme c'est par exemple le cas de la DRT.

Pour affiner encore un peu cette vision du phénomène de la coordination, il est intéressant de constater, avec C. Schnedecker et M. Bianco (1995), que le degré de cohésion des éléments d'un ensemble coordonné n'est pas constant. Et de façon encore plus surprenante, ce sont les verbes réciproques – dénotant des actions qui impliquent de façon réciproque et complémentaire les deux participants (*discuter*, *se regarder* etc.) – qui, au vu des résultats expérimentaux[70], désolidarisent leurs participants. Cet effet est encore renforcé par l'anaphore réciproque *l'un/l'autre* (Milner, 1982). Ainsi, dans la série (136), la cohésion des référents d'un verbe commitatif[71] (a) est plus forte qu'en (b), elle-même supérieure à (c).

(136) a. Tina et Lisa ont marché dans la montagne.
b. Tina et Lisa ont discuté la proposition du chef.
c. Tina et Lisa se sont regardées l'une l'autre. (Schnedecker et Bianco, 1995)

L'explication proposée par les auteurs repose sur l'hypothèse d'un effet de singularisation, lié au fait que les participants à des actions réciproques n'occupent pas de façon stable les rôles instanciés. Ils tiennent plutôt en alternance des places interchangeables. Cet effet de singularisation serait encore renforcé par l'anaphore réciproque, dissociant explicitement le couple initial.

En conclusion du premier cas de groupement – déclenché par la coordination – nous proposons alors la modélisation suivante : aucun des éléments du domaine ne reçoit une focalisation particulière. Cette décision tient compte des observations linguistiques sur les (im)possibilités de reprise individuelle par un pronom atonique ou un démonstratif. En revanche, elle ne permet pas de rendre compte des observations supplémentaires sur l'influence de la sémantique des verbes, mais dans la mesure où les cas exposés en (136) ne diffèrent pas dans leur comportement vis-à-vis d'une telle reprise (elles nous paraissent difficiles pour toutes les variantes), cela n'est pas nécessaire pour l'instant.

En ce qui concerne le deuxième cas de groupement – déclenché par le prédicat regroupant ses arguments – aucune des deux hypothèses initialement proposées ne convient : d'une part, la thèse de la saillance du premier élément mentionné peut être contrebalancée par la thèse de la saillance sur critères de récence, focalisant plutôt le dernier élément mentionné. D'autre part, la thèse « fusionniste » est en contradiction avec la majorité des modélisations jusqu'alors proposée : nous pouvons rappeler ici l'algorithme de C. Sidner (1979) qui distingue le focus des agents des autres foci,

[70] Il s'agissait d'une tâche de vérification de présence d'un mot dans une phrase lue préalablement. La variable dépendante est le temps nécessaire pour effectuer cette vérification. Pour plus de détails, cf. Schnedecker et Bianco (1995).
[71] absence de réciprocité

eux-mêmes ordonnés selon leur rôle thématique, ou encore la modélisation de la théorie du Centrage (Grosz et al., 1995) qui ordonne les centres selon la hiérarchie *sujet* > *objet direct* > *objet indirect*.

Par ailleurs, l'idée d'une hiérarchisation des rôles thématiques a aussi influencé des travaux dans le cadre de la grammaire cognitive : l'élaboration de schémas complexes s'accompagne d'une distinction entre le « profil » et la « base » à l'intérieur des schémas composés. Le profil correspond à des entités plus proéminentes, définies en fonction du schéma élaboré. Ainsi, le schéma de *à gauche de* (Figure 30a) permet non seulement de regrouper des entités, mais encore d'en profiler une : dans le cas de cette préposition spatiale, elle correspond à la cible, c'est-à-dire à l'élément à positionner, plutôt qu'au site, c'est-à-dire l'objet servant de référence spatiale[72]. De la même façon que pour l'élaboration de schémas prépositionnels, des schémas verbaux instaurent une hiérarchie entre les schémas d'objets qui viennent les élaborer. Si cette hiérarchisation correspond, *grosso modo*, à la hiérarchie des rôles grammaticaux, le cadre théorique de la grammaire cognitive permet d'aller plus loin, en proposant des explications unifiées de différents problèmes liés à l'interprétation anaphorique de pronoms : K. van Hoek (1995) a ainsi montré que des phénomènes relevant à la fois de la syntaxe (*c-commande* : Reinhart, 1976) et de la pragmatique pouvaient s'expliquer de façon uniforme sur la base de cette hiérarchisation, car elle est capable de dépasser le cadre phrastique.

Si la hiérarchisation des arguments d'un prédicat sur la base des rôles thématiques fait l'objet d'un consensus entre un grand nombre d'auteurs de différentes orientations (linguistique, psychologie, intelligence artificielle), force est de constater qu'elle n'est pas pour autant inconditionnelle. Il y a en effet d'autres facteurs qui influent sur la saillance respective des référents impliqués dans une même éventualité : les travaux de M. Charolles et L. Sprenger-Charolles (1989) ont par exemple souligné que la causalité verbale est un autre facteur déterminant de la focalisation des arguments. Ils ont montré que dans des énoncés tels que (137) et (138), les sujets produisent plus facilement des suites référant au *tigre* dans (137) et à l'*actrice* dans (138) :

(137) Le tigre effraie le chasseur parce qu'il ...

(138) L'habilleuse envie l'actrice parce qu'elle... (Charolles et Sprenger-Charolles, 1989)

L'explication serait la suivante : ce qui est focalisé par un verbe causal est la cause de l'événement. Or, celle-ci peut être portée par le sujet ou par l'objet du verbe, ce qui se superpose effectivement à une échelle de focalisation basée uniquement sur les rôles thématiques. Comme le montrent C. Schnedecker et M. Bianco (1995) par une liste importante de travaux psycholinguistiques, le statut cognitif privilégié de l'élément causal est un phénomène très robuste. Cependant, des interactions avec d'autres facteurs encore existent : parmi ceux-ci nous mentionnerons, pour terminer, les connecteurs. Il se trouve que les recherches centrées sur la causalité verbale ont systématiquement utilisé des énoncés dont la proposition principale était suivie d'une subordonnée introduite par *parce que*. Or, une telle subordonnée induit en effet une concentration sur la cause de l'événement, en ce qu'elle ouvre sur une explication de l'événement. D'autres psycholinguistes (Ehrlich, 1980 ; Stevenson et al., 1994) ont donc confirmé l'influence des connecteurs sur l'interprétation d'un pronom ambigu. Ainsi, pour des énoncés tels qu'en (139), K. Ehrlich a pu observer que les choix d'interprétation étaient majoritairement compatibles avec l'orientation causale en présence du connecteur *because*, alors que c'était le contraire pour *but*. Les interprétations en présence de *and* ont été indépendantes du biais causal.

(139) Steve blamed Frank (because/but/and) he spilt the coffee.[73]

[72] La terminologie « site » et « cible » est empruntée à C.Vandeloise (1986).

[73] Steve a fait des reproches à Frank (parce que / mais / et) il a renversé du café.

Pour résumer les travaux sur la focalisation des référents impliqués dans un même événement, on peut constater un consensus large sur l'importance accordée à une hiérarchisation en fonction du rôle thématique. Cela signifie pour notre modélisation que ce type de regroupement, contrairement au regroupement déclenché par la coordination, implique la création d'un domaine focalisé. Le choix de l'élément focalisé dépend de la structure argumentale du prédicat et suit l'échelle des rôles thématiques. Le problème particulier posé par les regroupements des éléments reliés par des prépositions spatiales peut également être résolu dans ce cadre, en suivant les intuitions de la grammaire cognitive : un groupement d'objets déclenché par une préposition spatiale instaure une hiérarchisation entre ses éléments et focalisera celui qui fonctionne en tant que cible par rapport à celui fonctionnant comme site. Les biais supplémentaires (verbes causaux, connecteurs) sont pour l'instant laissés de côté, mais le cadre général de notre modélisation permet d'intégrer ce type de contraintes supplémentaires, si cela s'avère nécessaire.

7.3.5 Les groupements perceptifs

Jusqu'ici, nous nous sommes concentrée sur des groupements de référents introduits par mention discursive et déclenchés par des critères discursifs (coordination ou implication dans une même éventualité dénotée par le prédicat). Nous avons pourtant mentionné, lors de la synthèse des modélisations contextuelles existantes (section 4.4), que cela est insuffisant : dans un modèle contextuel pour la résolution de la référence dans le dialogue homme-machine, il faut aussi tenir compte des facteurs perceptifs, importants dès lors que les interlocuteurs partagent un même espace visuel mettant à leur disposition des référents potentiels.

Les expériences menées par K. Kessler et al. (1996)[74] ainsi que bon nombre d'exemples du corpus Ozkan montrent clairement que des groupements perceptifs sont également nécessaires : ainsi, en M_2 de l'exemple (140), *en* ne peut pas s'interpréter sans la création d'une représentation pour l'ensemble des trois triangles, jamais introduite discursivement, mais visible sur la scène en construction :

(140) I_1 pyramides dans le désert... alors une grosse pyramide
I_2 [geste: I prend un premier triangle]
I_3 alors tu mets un grand triangle sur la droite de l'autre et un petit peu au-dessus
M_1 [geste: M prend un deuxième triangle]
I_4 ouais
I_5 [geste: I prend un troisième triangle]
M_2 tu veux pas que j'*en* enlève un (C12Egypte)

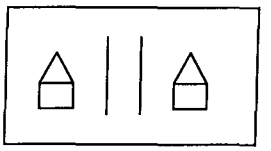
Afin de créer de tels groupes, il est par exemple possible de tenir compte des principes mis en évidence par la théorie de la Gestalt (Wertheimer, 1923). Ce courant de la psychologie s'est développé en même temps que le béhaviorisme, par opposition aux tendances introspectionnistes dominantes à la fin du XIX^{ième} siècle. Il considère que le donné psychique est constitué d'emblée d'unités et de groupes qui sont des formes. Les conditions d'apparition de ces formes résident à la fois dans les dispositions et attentes du sujet (*Einstellung*) et dans l'excitant lui-même (*Reizlage*). Ces conditions ont été illustrées par les Gestaltistes eux-mêmes, en empruntant des exemples au domaine de la perception visuelle.

En ce qui concerne les dispositions du sujet lui-même, les Gestaltistes ont montré que selon l'attention, on est capable d'inverser la perception du fond et de la figure d'une scène ambiguë. Ils défendent aussi l'hypothèse selon laquelle on perçoit non seulement en fonction de ce que l'on attend, mais aussi en fonction de sa mémoire, c'est-à-dire de ce que l'on connaît. Le corpus Ozkan permet de confirmer cette hypothèse : certains instructeurs ne donnent en effet aucune information figurative sur

[74] cf. la présentation que nous en avons faite dans la section 4.4

la scène à construire. Or, cela n'empêche pas le manipulateur de se rendre compte, à partir d'une certaine complexité de la scène, de la nature des figures à construire et ce d'autant plus qu'il s'attend à composer un dessin :

(141) I_1 voilà c'est tout, le deuxième dessin est fini
M_1 d'accord, c'est une route au milieu ?
I_2 une quoi ?
M_2 une route ?
I_3 une route ouais (C11Route)



En (141), le manipulateur perçoit donc en effet un ensemble de deux barres verticales qu'il groupe en un objet complexe dont il suppose qu'il s'agit d'une route. Il fait cela en raison de l'attente induite par la tâche (construction de dessin) et en raison de la perception de routes (figuratives) qu'il a pu avoir antérieurement.

Concernant les stimuli visuels eux-mêmes, la théorie de la Gestalt propose des critères permettant de décider si un groupe d'éléments perceptifs suit la loi de la « bonne forme » : celle-ci traduit le principe selon lequel toute forme tend à la plus grande simplicité. Cette loi permet alors de prédire quels ensembles d'éléments ont des chances d'être perçus de façon holistique, c'est-à-dire en tant que forme. La loi de la « bonne forme » se décline selon plusieurs lois de groupement, dont les plus connues reposent sur des principes de proximité, de similarité et de bonne continuité.

Si les principes dégagés par la théorie de la Gestalt semblent mettre à jour des mécanismes essentiels de la façon dont les sujet structurent leur environnement, ils ne permettent pas, comme le remarque F. Wolff (1999), de répondre au problème de la structuration hiérarchique des groupements perceptifs. Ainsi, la palette du corpus Ozkan (Figure 31) peut donner lieu à une multitude de groupements perceptifs : les critères de proximité permettent de regrouper les barres deux par deux ; les critères de similarité permettent de former cinq groupes, selon le type des figures géométriques ; les critères de bonne continuité permettent de diviser la palette en deux groupes de formes (formes du haut vs. formes du bas).

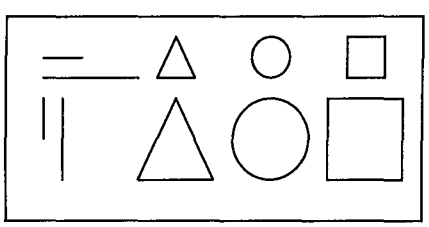


Figure 31 : La palette des formes disponibles du corpus Ozkan

Ce problème a été abordé, entre autres, dans des travaux sur la perception appliquée à l'interaction homme-machine. X. Briffault (1993) propose un algorithme permettant d'effectuer des groupements à différents niveaux de granularité, mais seulement sur le critère de proximité. N. Bellalem (1995) définit trois relations élémentaires entre objets, permettant de construire des groupes entre objets et de définir, de manière récursive, des relations entre groupes. Enfin, les règles la théorie de la Gestalt ont été utilisées de façon plus directe et systématique par un algorithme de simulation de la perception, développé par K. Thorisson (1994). Cet algorithme calcule des scores de proximité et de similarité entre deux percepts et utilise ceux-ci pour détecter des groupes à saillance prédominante par un algorithme itératif détectant des contrastes. L'avantage de cet algorithme est de fournir des groupements à granularité variable, triés en fonction de leur saillance.

En supposant que ce type d'algorithme nous fournit donc des domaines de référence perceptifs, il nous reste à calculer, comme pour les groupements discursifs, les caractéristiques de la partition de ces domaines. Cela signifie que nous devons spécifier le calcul du critère de différenciation et de la focalisation. En reprenant l'exemple des groupes perceptifs formés par les deux lignées horizontales de figures dans la palette (haut vs. bas), le critère de différenciation pourrait en effet être le type des éléments (*CARRÉ vs. CERCLE* etc.), mais il est aussi possible d'accéder aux éléments de ces groupes par leur position horizontale (*PREMIER, DEUXIÈME* etc.). Les choix à faire peuvent alors s'inspirer d'un article de R. Dale et E. Reiter (1995) consacré à la génération d'expressions référentielles : ils y rapportent des expériences ayant montré que des locuteurs ont tendance à utiliser de préférence certaines caractéristiques et à en délaisser d'autres pour référer à des objets d'un ensemble contextuel donné. *Grosso modo*, l'échelle de préférence est la suivante : *type > propriété intrinsèque > propriété relationnelle*. On pourrait alors adopter cette échelle pour décider entre plusieurs critères de différenciation possibles. Le choix d'un critère de différenciation particulier ne réduirait pas les autres possibilités d'accès référentiel, mais il serait prédictif dans la mesure où il correspondrait à celui qu'un interlocuteur semble préférer dans la situation donnée.

Enfin, la question de la focalisation d'un élément à l'intérieur d'un groupe perceptif a fait l'objet d'une étude entreprise par F. Landragin (1998). L'auteur propose une classification des types de saillance visuelle (par catégorie, par caractéristiques physiques, par fonctionnalités, par localisation dans la scène, par incongruité, par dynamique, par mise en évidence par l'application), avant de centrer son étude sur la saillance spatiale dans un environnement virtuel 3D. Il propose en particulier un algorithme permettant de calculer la saillance spatiale d'un objet en tenant compte des regroupements et de la distance de l'objet par rapport au locuteur. L'intérêt de ce travail par rapport à notre préoccupation est de montrer qu'il est possible de calculer certains types de saillance perceptive et qu'il faut en conséquence prévoir une opération de focalisation lors de la création de domaines de référence perceptifs. Mais l'étude fait aussi apparaître la nécessité d'investigations futures sur le poids respectif de chacun des facteurs mentionnés ainsi que sur leurs interactions.

7.3.6 Groupement : bilan, exemples et représentations graphiques

Pour résumer cette section consacrée à l'opération de groupement, retenons que la création de domaines de référence locaux est indispensable à la gestion d'un modèle du contexte prédictif pour l'interprétation référentielle. L'opération de groupement participe ainsi à la création et la mise à jour d'un historique du discours, composé de domaines de référence, qui correspondent soit à des objets individuels, soit à des ensembles. Cette opération est cependant tributaire de critères de natures très différentes : certains sont relativement faciles à calculer, d'autres ne semblent pas encore entièrement formalisables. Par ailleurs, il faut être capable d'intégrer des informations discursives et perceptives dans un même cadre théorique. Pour cette raison, il nous a d'abord semblé nécessaire de proposer une opération générique, instanciable selon les conditions particulières d'une application spécifique : les caractéristiques communes à toute opération de groupement sont le calcul, à partir d'au moins deux représentations mentales fournies en arguments, d'une nouvelle représentation, ce qui implique en particulier d'en calculer un type et la structure partitionnelle (critère de différenciation et focalisation). Le chapitre 10 présentera les algorithmes détaillés pour les trois cas que nous avons décidé d'intégrer à notre modélisation : groupement déclenché par une coordination *et*, groupement déclenché une préposition spatiale, groupement déclenché par des facteurs perceptifs. Dans la suite, nous illustrons ces opérations par trois exemples :

(142) Prends *un triangle et un carré.*

En (142), l'opération de groupement est déclenchée par la coordination de deux expressions référentielles. Cela a pour conséquence la création d'une représentation mentale complexe à partir des représentations individuelles des référents. L'application des principes de groupement aux représentations individuelles des référents de l'exemple (142) mène à la construction d'une nouvelle représentation telle qu'illustrée dans la Figure 32a : le type du nouveau domaine, calculé à partir de la hiérarchie de types, est *FIGURE*. Une partition distingue les éléments regroupés en fonction de leur type : le critère de différenciation est donc *type*. Enfin, conformément au principe appliqué à des groupements déclenchés par la coordination *et*, la partition n'est pas focalisée.

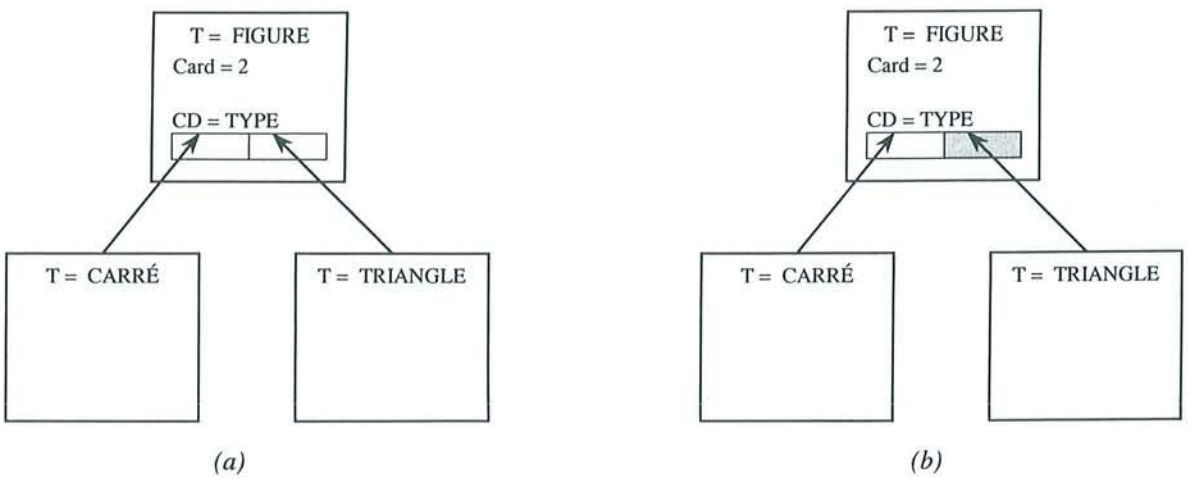


Figure 32 : Exemples de groupement

(143) Mets un triangle sur un carré.

Un exemple tel que (143) mène à un groupement des référents individuels correspondant aux arguments de l'action *METTRE_SUR*. Comme en (142), il s'agit d'un référent de type *TRIANGLE* et d'un référent de type *CARRÉ*. En revanche, le groupement implique ici la création d'une partition focalisée. Pour l'instant, nous avons décidé d'appliquer des principes de focalisation proches de la théorie du Centrage (Grosz et al., 1995), postulant que l'ordre focal des éléments se calcule en fonction du rôle thématique, selon le schéma *agent* > *patient* > *autre*. Étant donné que nous n'intégrons pas les représentations des locuteurs dans notre modélisation, l'agent de l'action n'est pas disponible. En tant que patient de l'action ainsi qu'en tant que cible du positionnement, le référent correspondant à *un triangle* sera donc focalisé (Figure 32b).

Enfin, la scène visuelle de l'exemple (144) donne lieu à un groupement des représentations individuelles des deux triangles. Les facteurs perceptifs qui déclenchent ce groupement sont la proximité et la similitude, relativement à la troisième figure de la scène. En appliquant les algorithmes de groupement perceptif, on aboutit à une représentation pour un ensemble de type *TRIANGLE*, partitionné selon la position horizontale de ses éléments (Figure 33).

(144)

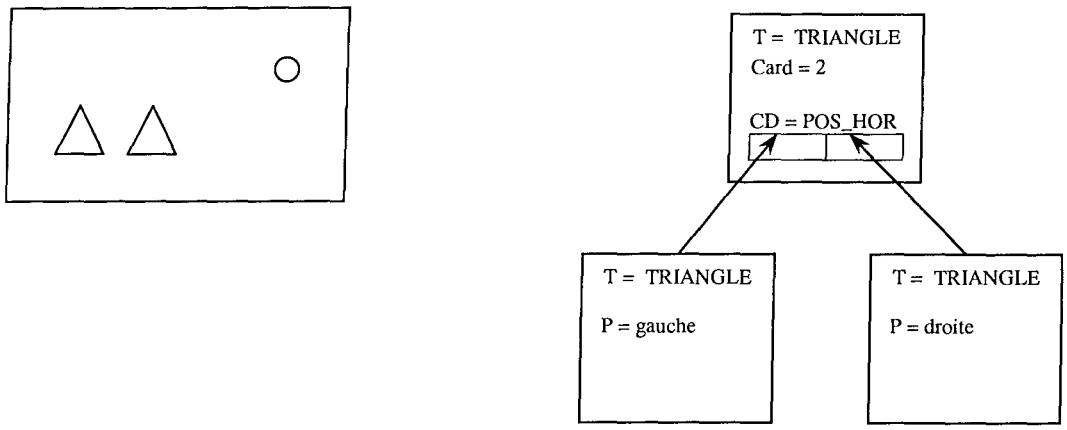


Figure 33 : Groupement de l'exemple (144)

7.4 Synthèse : les différentes structures contextuelles possibles

En réponse aux deux premières questions à l'origine des deux dernières sections de ce chapitre, nous retenons que les unités fondamentales du modèle contextuel sont des domaines de référence, modélisés par des représentations mentales. Les domaines sont soit introduits par la mention ou la perception d'une nouvelle entité contextuelle, soit issus d'une opération de groupement. Leurs caractéristiques les plus importantes pour la suite sont les informations sur le type et la décomposition des entités représentées. Les décompositions possibles sont modélisées dans une ou plusieurs partitions, organisées autour d'un critère de différenciation qui permet un accès individuel à chacune des entités de la partition. Enfin, un des éléments d'une partition peut être focalisé. Par ailleurs, les domaines de référence de notre modèle contextuel sont ordonnés selon leur récence d'activation, sachant qu'un domaine est activé lorsqu'il vient d'être créé ou lorsqu'il vient de servir à l'extraction d'un référent[75].

Cela nous amène donc à trois structures contextuelles fondamentales, retenues pour la construction des lignes du tableau synthétique (Tableau 17) :

les domaines non partitionnés : cf. la mention de *deux triangles* de l'exemple (125) ;

les domaines partitionnés sans élément focalisé : cf. le groupement de *un triangle et un carré* dans l'exemple (142) ;

les domaines partitionnés contenant un élément focalisé : cf. le groupement de *un triangle* et *un carré* en (143).

Par la suite, nous allons élaborer les colonnes et le corps de ce tableau. L'en-tête des colonnes fait l'objet du chapitre suivant : il détaillera les opérations nécessaires à l'interprétations de différents types d'expressions (indéfinies, définies, démonstratives, pronominales). Le corps, commenté au chapitre 8 donnera une vue synthétique de la compatibilité de ces opérations avec les différentes structures contextuelles introduites dans ce chapitre.

[75] L'opération d'extraction est abordée au chapitre suivant.

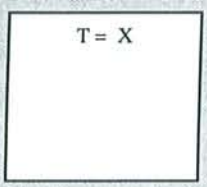
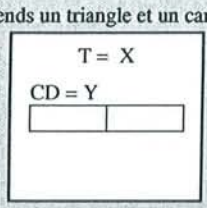
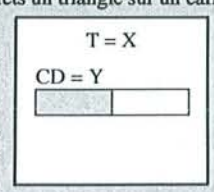
Structures contextuelles possibles	*Expression indéfinie*	*Expression définie*	*Expression démonstrative*	*Expression pronominale*
Prends deux triangles. 	*Prédiction sur l'emploi des différents types de marqueurs face à différentes structures contextuelles*			
Prends un triangle et un carré. 				
Mets un triangle sur un carré. 				

Tableau 17 : Construction du tableau synthétique : Élaboration des structures contextuelles possibles

8 Le modèle d'interprétation des expressions référentielles

8.1 Introduction

Ce chapitre est consacré à la modélisation du processus d'interprétation d'une expression référentielle. Il répondra ainsi à la troisième question soulevée lors de la synthèse sur les modélisations existantes (cf. la section 4.4) et restée en suspens jusqu'ici : celle du rôle des expressions elles-mêmes lors du recrutement de leur référent.

Nous verrons que ce rôle est double : nous suivons en cela le schéma synthétique conçu à la fin du chapitre consacré aux travaux linguistiques (Figure 4 : Synthèse des conditions contextuelles et effets interprétatifs associés aux différents marqueurs référentiels, page 45). Ce schéma classe les expressions référentielles en fonction des structures contextuelles présupposées et en fonction de leurs effets interprétatifs. Premièrement, nous considérons qu'une expression impose des contraintes sur la structure de son contexte d'interprétation, c'est-à-dire sur la structure d'un domaine de référence compatible parmi tous ceux du modèle contextuel. Le deuxième rôle des expressions référentielles consiste à extraire de ce domaine leur référent, processus qui passe par une opération de restructuration du domaine sélectionné et qui mène à une mise à jour du modèle contextuel.

La spécification de ces deux caractéristiques se fera en fonction des différents types d'expressions : indéfini, défini, pronom et démonstratif. Elle nous permettra ainsi de compléter l'en-tête des colonnes du tableau synthétique : nous préciserons pour chaque type d'expression les contraintes sur les domaines compatibles et les instructions de restructuration du domaine retenu (Tableau 18). Le corps du tableau donnera, pour chaque type d'expression, le résultat de l'interprétation en fonction de la structure initiale du domaine sélectionné.

Type d'expression	*Indéfini*	*Défini*	*Démonstratif*	*Pronom*
Contraintes sur la structure des domaines compatibles				
Opération de restructuration				
Structure 1	Prédictions sur l'emploi des expressions indéfinies	Prédictions sur l'emploi des expressions définies	Prédictions sur l'emploi des expressions pronominales	Prédictions sur l'emploi des expressions démonstratives
Structure 2				
Structure 3				

Tableau 18 : Ajout des informations spécifiques aux différentes expressions référentielles

8.2 Principe d'interprétation commun à toutes les expressions référentielles

Nous considérons que le principe d'interprétation commun à toutes les expressions référentielles consiste en deux étapes :

(a) La première étape est la formulation, sous forme d'un **domaine de référence sous-spécifié**, de contraintes sur la structure d'un domaine de référence compatible :

Différents types d'expressions (indéfini, défini, pronom, démonstratif) imposent en effet différentes contraintes sur la sélection d'un domaine de référence parmi ceux du contexte. Cette considération s'appuie sur une série d'études linguistiques fines consacrées au fonctionnement propre des différents marqueurs référentiels. Il nous est apparu à travers les travaux linguistiques (présentés au chapitre 2) que le point commun de toutes les expressions référentielles est de s'interpréter dans un domaine contextuel. Mais au-delà de ce point commun, les différentes formes d'expressions référentielles sont à considérer comme autant d'instructions différentes pour le recrutement du référent dans ce domaine. Nous essayerons donc de concrétiser cette idée dans le cadre de notre modélisation : nous montrerons en effet qu'il est possible de modéliser ces instructions de traitement par des domaines de référence sous-spécifiés. Un domaine de référence sous-spécifié dépend à la fois du type et de la sémantique de l'expression à interpréter. Il exprime des conditions structurelles sur la représentation mentale d'un domaine contextuel qui soit approprié à son interprétation. A titre d'exemple, considérons (145) :

(145) alors tu vas prendre *le gros rond* (C8Homme)

L'expression à interpréter est la description définie *le gros rond.* Sur la base du type et de la sémantique de celle-ci, le domaine de référence sous-spécifié devra exprimer que cette expression s'interprète dans un domaine réunissant des éléments de type *ROND*, à l'intérieur duquel un élément particulier puisse être isolé sur la base d'une propriété qui est sa taille. Ce domaine sous-spécifié a donc la forme du domaine $@DR_{ER}$ de la Figure 34. Le domaine de référence sous-spécifié sert ensuite de motif de recherche pour la sélection d'un domaine compatible parmi ceux du contexte.

(b) La deuxième étape est l'extraction, par une **restructuration** du domaine sélectionné, du référent de l'expression à interpréter :

Une fois un domaine compatible trouvé, il s'agira de le mettre à jour par une opération de restructuration. L'identification du référent de l'expression à interpréter est en quelque sorte un « effet de bord » de cette restructuration.

Par rapport aux autres modèles computationnels et algorithmes pour le calcul référentiel (présentés aux chapitres 4 et 5), l'aspect original de notre modélisation vient de ce que nous ne considérons pas les expressions référentielles uniquement comme des instructions de recherche d'un référent convenable. Nous défendons plutôt l'hypothèse qu'une expression référentielle remplit deux fonctions à la fois : elle désigne effectivement une entité de l'univers du discours, mais elle le désigne d'une certaine façon qui n'est pas sans importance pour la mise à jour du modèle de l'univers discursif. C'est ce point que souligne G. Kleiber lorsqu'il constate que « *un des défauts les plus courants des analyses référentielles est d'oublier qu'avec l'identification du référent l'affaire n'est pas finie pour autant et que le mode de donation du référent est tout aussi important que le référent lui-même. Préoccupés essentiellement de l'assignement d'un référent à l'expression référentielle, beaucoup de modèles de la référence s'arrêtent une fois le référent identifié et négligent ainsi la manière même dont le référent est donné. Or, si l'on veut comprendre d'un peu plus près comment se passent les processus référentiels, on ne peut faire l'impasse sur le mode de donation référentielle.* » (Kleiber 1994 :18).

Nous essayerons alors d'intégrer ce point de vue dans notre modélisation en considérant que les différents types d'expressions référentielles sont autant d'instructions sur la manière de restructurer leur domaine de référence. A travers les représentations mentales, nous disposons

en effet d'un moyen efficace pour modéliser les effets que l'interprétation d'une expression référentielle peut avoir sur la représentation du contexte. En reprenant l'exemple du défini *le gros rond*, nous considérerons par exemple que l'opération de restructuration consiste en un changement de la structure focale de la partition du domaine d'interprétation ($@DR_{Rs}$ de la Figure 34).

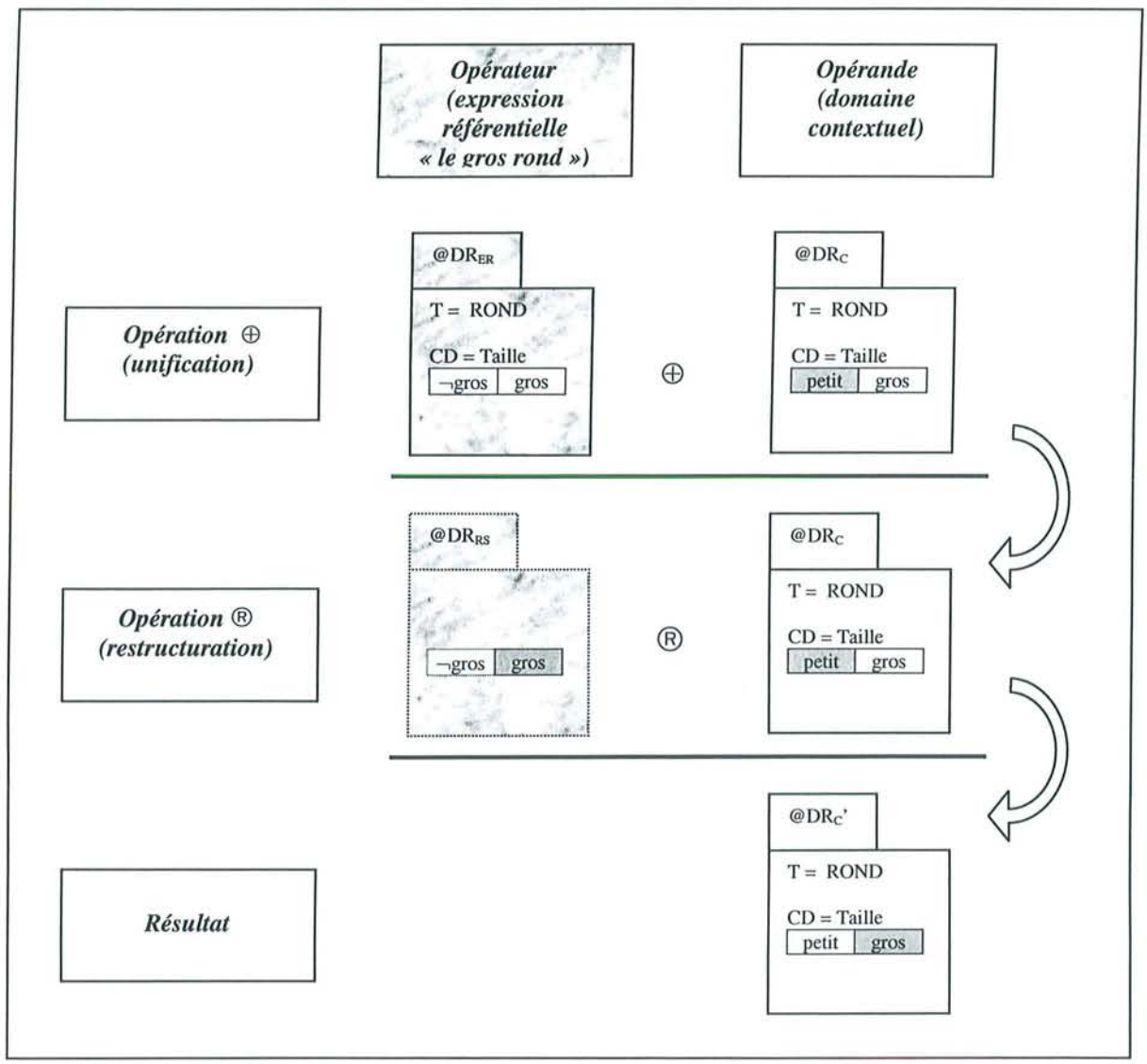


Figure 34 : Interprétation d'une expression référentielle – unification et restructuration

La Figure 34 donne une illustration graphique de ce principe : une expression référentielle à interpréter y est considérée comme un opérateur agissant sur une opérande qui est le contexte d'interprétation. Cet opérateur agit doublement : le processus d'interprétation mène d'abord à la construction d'un domaine de référence sous-spécifié ($@DR_{ER}$) qui exprime des contraintes sur un ou des domaines contextuels disponibles. Ce domaine sous-spécifié sert ensuite de motif de recherche pour un domaine compatible du modèle contextuel. Lorsqu'un tel domaine est identifié, le domaine sous-spécifié est unifié, par l'opération « ⊕ », avec ce dernier. Le résultat de cette première étape est un domaine ($@DR_C$) à l'intérieur duquel le référent de l'expression reste à identifier. Cette identification se fera à travers une deuxième opération – l'opération de restructuration ® – également contrainte par le type de l'expression à interpréter ($@DR_{Rs}$). Dans tous les cas, le résultat de

l'opération de restructuration est un domaine contextuel restructuré ($@DR_{C'}$) contenant une partition avec un élément focalisé qui correspond au référent.

Dans les sections suivantes, nous proposons donc d'instancier ce principe général, en examinant de façon plus détaillée :

a) la construction des domaines de référence sous-spécifiés et

b) les instructions de restructuration,

en fonction des différents types d'expressions à interpréter : expressions indéfinies (8.3.1), expressions définies (8.3.2), expressions démonstratives (8.3.3) et expressions pronominales (8.3.4). Ensuite, nous serons capable de présenter et de commenter le tableau synthétique complet (8.4). Nous terminerons par une section qui illustre comment les principes dégagés permettent de rendre compte de l'interprétation des groupes nominaux sans nom, sans pour autant leur attribuer un statut particulier (8.5).

8.3 Instanciation du principe d'interprétation

8.3.1 Les expressions indéfinies

- *Le domaine de référence sous-spécifié*

Les travaux linguistiques sur le fonctionnement des expressions indéfinies ont montré que la caractéristique interprétative essentielle de ces marqueurs consiste en une opération de dénombrement sur la classe nominale de type *N*. Cette opération se fait relativement à une énonciation et elle est indépendante du contexte et des opérations d'extractions précédentes (cf. notre présentation au chapitre 2, section 2.2). Pour la modélisation du volet (a) de notre principe général – les contraintes interprétatives en termes de domaine de référence sous-spécifié – nous retenons donc le principe suivant : étant donné qu'une description indéfinie de type *n N* (*une/deux/... maisons*) extrait *n* éléments d'un domaine composé d'éléments de type *N*, elle présuppose donc, pour s'interpréter, un **domaine de référence composé d'éléments de type N** (Figure 35).

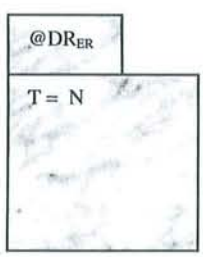


Figure 35 : Structure du DR_{ER} des indéfinis – « un N »

Conformément à ce principe, l'indéfini *une maison* de l'exemple (146) demande à s'interpréter dans un domaine d'éléments de type *MAISON*. Le domaine de référence sous-spécifié qui exprime cette contrainte est le domaine $@DR_{ER}$ de la Figure 37 (page 151).

(146) donc au bord de cette route, il y a deux maisons, donc *une maison* qui se trouve à gauche de cette route et l'autre à droite. (C9Forêt, modifié)

Une fois le domaine sous-spécifié construit, l'étape suivante consiste à parcourir les domaines disponibles du modèle contextuel, afin d'en identifier un qui puisse être unifié avec ce domaine sous-

spécifié. En (146), le domaine sous-spécifié pour *une maison* s'unifie par exemple avec celui des deux maisons, introduit auparavant dans le modèle contextuel (@DR_c de la Figure 37). Au vu de cette recherche dans le modèle contextuel, on pourrait objecter que le principe de l'indépendance contextuelle totale des indéfinis doit être nuancé. Mais cette objection ne remet pas en cause les mécanismes fondamentaux de l'interprétation d'un indéfini : celle-ci n'est en effet aucunement subordonnée à l'existence d'un domaine contextuel spécifique, comme celui des deux maisons en (146). Lorsqu'aucun domaine compatible n'est trouvé parmi ceux du contexte, l'interprétation se fera à l'intérieur de la classe conceptuelle des *N*, fournie par la représentation mentale générique. Ainsi, pour un énoncé tel que :

(147) euh maintenant tu vas prendre un petit carré et le mettre en bas à gauche pour faire *une maison*

le domaine de référence sous-spécifié pour *une maison* s'unifiera avec le domaine générique *@MAISON*, car aucun ensemble spécifique d'éléments de type *MAISON* n'est disponible dans le modèle contextuel.

- *L'opération de restructuration*

Le principe d'indépendance contextuelle et plus précisément d'indépendance vis-à-vis des extractions précédentes, se traduit dans notre modèle par le fait qu'une expression indéfinie **crée systématiquement une nouvelle partition** à l'intérieur de son domaine d'interprétation. Cette nouvelle partition se justifie par un nouveau critère de différenciation (introduit par le prédicat de l'énoncé). A l'intérieur de la nouvelle partition, l'indéfini recrute son référent par l'extraction d'un élément quelconque qui sera distingué des autres éléments par un critère de différenciation reposant sur des informations prédicatives. Enfin, l'opération de restructuration se termine par une focalisation du nouveau référent. La forme générale des instructions de restructuration est donc celle de la Figure 36[76,77].

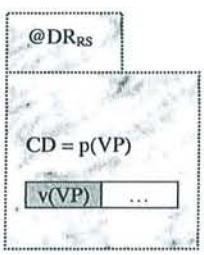


Figure 36 : Instruction de restructuration pour une expression indéfinie

L'application de ces principes à (146) conduit au résultat suivant : le référent de *une maison* sera l'une des deux maisons (n'importe laquelle) du domaine @DR_c de la Figure 37. Cette maison se distinguera de l'autre par le critère de différenciation *POSITION_HORIZONTALE*, véhiculé par la prédication et spécifiant qu'elle a la propriété d'être *À GAUCHE*.

[76] Nous notons par p(VP) une fonction qui calcule, à partir d'une prédication VP, la propriété p exprimée par cette prédication. De même, v(VP) retourne la valeur de cette propriété pour la prédication VP.

[77] Pour toutes les représentations graphiques des instructions de restructuration, nous utiliserons des traits pointillés pour les éléments du domaine qui ne sont pas affectés par cette opération.

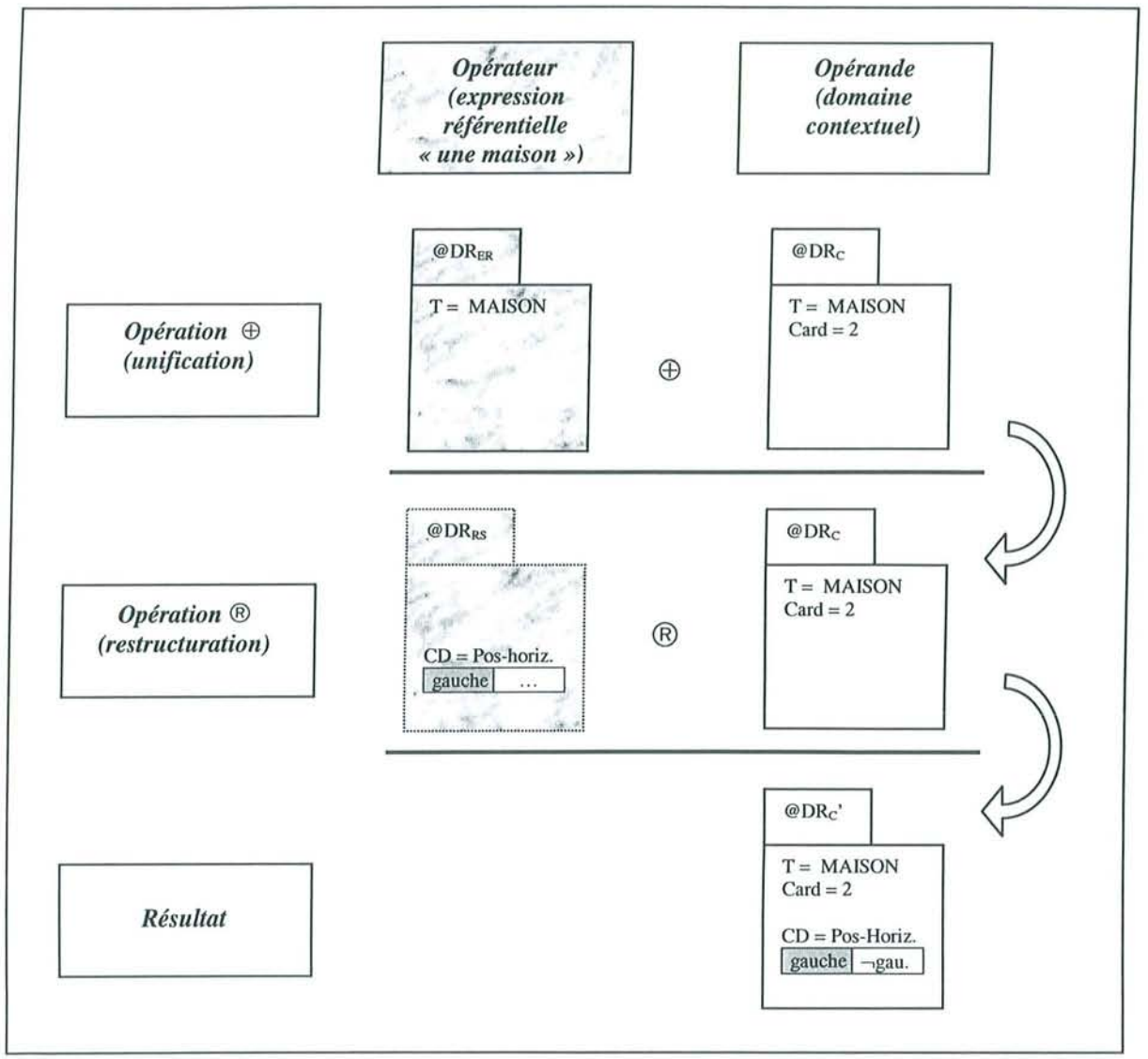


Figure 37 : Interprétation d'une expression référentielle indéfinie : « une maison » (exemple (146))

- *Résumé de l'interprétation des expressions indéfinies*

En résumé, l'interprétation pour une expression indéfinie est modélisée par :

a) la construction d'un domaine de référence sous-spécifié exprimant une contrainte sur le type des éléments contenus dans le domaine d'interprétation : il doit s'agir d'éléments de type *N*. Lorsque la tête nominale N est accompagnée par un modifieur P, des contraintes supplémentaires s'imposent en fonction de la sémantique des modifieurs (expression d'une propriété intrinsèque comme *bleu*, l'ordonnancement comme *premier* ou l'altérité comme *autre*). Pour une propriété intrinsèque par exemple, les éléments du domaine de référence doivent posséder cette propriété. Les algorithmes détaillés de la section 11.3.2 donnent plus de renseignements à ce sujet, mais les grandes lignes de la modélisation proposée peuvent être suivies indépendamment de ce point. La recherche d'un domaine compatible se fait prioritairement parmi les domaines du modèle contextuel. Si aucun domaine du modèle contextuel ne convient, on sélectionne le domaine générique correspondant à la classe des éléments requis.

b) la création d'une nouvelle partition dans ce domaine : cette nouvelle partition oppose, sur la base d'informations véhiculées par la prédication, le référent de l'expression aux autres éléments du domaine. Le référent de l'expression à interpréter sera focalisé.

8.3.2 Les expressions définies

- *Le domaine de référence sous-spécifié*

L'interprétation classique des descriptions définies s'appuie sur la présupposition existentielle et la condition d'unicité, dont il est généralement admis qu'elles sont réduites à un contexte correspondant à l'univers du discours : « *Definite descriptions often manage to select a unique referent through the combined forces of their own descriptive content and information supplied by the context in which they occur.*[78] » (Kamp et Reyle, 1993 : 253). A partir de là, on peut expliquer la préférence pour un traitement stipulant qu'une description définie réfère à un objet déjà introduit dans le discours. Les principaux problèmes que nous avons relevés sont d'une part l'impossibilité d'expliquer la différence entre pronoms et descriptions définies et la nécessité d'extensions théoriques pour tenir compte des relations non coréférentielles entre un antécédent et une description définie d'autre part (cf. section 2.3). Pour ces raisons, notre modélisation s'inspire des propositions de B. Gaiffe et al. (1997) qui partent de l'hypothèse inverse : les syntagmes définis créent un nouveau référent du discours et les cas de coréférence sont des exceptions. La présupposition existentielle s'applique donc non plus au référent du défini, mais à un contexte dont ce référent peut être extrait. Ce point de vue peut être largement rapproché de celui de F. Corblin (1987 : 244) : « *Le défini demande, pour être interprété, un domaine dans lequel son contenu soit en mesure de constituer un signalement singularisant.* »

Dans le cadre de notre modélisation, une expression définie cherche donc à s'interpréter dans un domaine de référence spécifique. Contrairement aux indéfinis, il doit s'agir d'un domaine effectivement disponible dans le modèle contextuel et non pas d'un domaine générique. A l'intérieur de ce domaine, un défini est supposé extraire son référent sur la base d'un critère distinctif. Cela signifie qu'il présuppose non seulement l'existence d'un domaine spécifique, mais encore d'un **domaine spécifique partitionné**. La structure exacte du domaine de référence sous-spécifié dépend de la forme de l'expression définie, mais dans tout les cas, la description contraint le type des éléments du domaine, le critère de différenciation de la partition présupposée et la valeur de ce critère qui permettra d'identifier le référent.

Pour une expression de la forme *le N P*, le type des éléments du domaine est donné par *N*. A l'intérieur de cet ensemble de *N*, l'interprétation de l'expression présuppose l'existence d'une partition sur le critère de différenciation $p(P)$[79]. Enfin, pour que l'expression puisse être interprétée, un des éléments de la partition doit pouvoir être isolé par sa valeur *P* pour ce critère de différenciation (Figure 38) :

[78] « Les descriptions définies sélectionnent souvent un référent unique à travers la combinaison de leur propre contenu descriptif aux informations fournies par le contexte d'occurrence. »

[79] En analogie à la fonction p(VP), nous supposons une fonction p(P) qui retourne, pour un modifieur P donné, la propriété correspondant à P. Notons également que des règles particulières s'appliquent lorsque P est une expression ordinale ou d'altérité (cf. section 11.3.2).

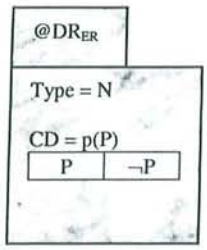


Figure 38 : Structure du DR$_{ER}$ des définis – « le N P»

(148) alors tu vas prendre *le gros rond* (C8Homme)

Ainsi, pour *le gros rond* en (148), on cherchera un domaine dont les éléments sont de type *ROND* et qui possède une partition organisée autour du critère de différenciation *p(GROS)*, c'est-à-dire la taille des objets. Enfin, à l'intérieur de cette partition doit exister un élément pour qui le critère de différenciation prend la valeur *GROS*. Ce domaine sous-spécifié est celui lequel est présenté graphiquement par @DR$_{ER}$ de la Figure 34, page 147.

Pour une expression de la forme *le N*, le type des éléments doit être un type pouvant englober un *N* dans sa partition. Il peut donc s'agir d'un super-type de *N* ou d'un type qui se compose entre autres d'un élément de type *N*[80]. A l'intérieur de ce domaine, l'interprétation de l'expression présuppose l'existence d'une partition avec un critère de différenciation *CATÉGORIE*, c'est-à-dire distinguant les élément de la partition selon leur type[81]. Enfin, pour que l'expression puisse être interprétée, un des éléments de la partition doit pouvoir être isolé par son type *N* (Figure 39) :

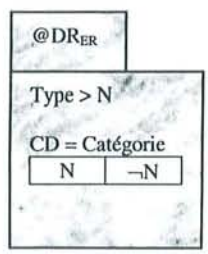


Figure 39 : Structure du DR$_{ER}$ des définis – « le N »

(149) I_1 donc l'autre dessin c'est *une personne*
I_2 assise donc sur *une chaise* et juste à côté *d'une table*
I_3 donc *la table* est représentée par un grand cercle (C9Homme)

A titre d'exemple, en (149), l'expression *la table* demande à s'interpréter dans un domaine de type > *TABLE*, permettant d'isoler, dans une partition selon la catégorie, un objet de type *TABLE*. La représentation graphique de ce domaine est donnée par @DR$_{ER}$ de la Figure 41, page 154.

Comme pour le domaine sous-spécifié des indéfinis, le domaine sous-spécifié des définis sert ensuite de motif de recherche pour un domaine compatible parmi ceux du contexte. On doit noter ici que la partition du domaine sous-spécifié peut ne pas être instanciée directement dans le domaine contextuel. Un domaine contextuel reste en effet compatible avec le domaine sous-spécifié, à condition que la partition présupposée puisse être récupérée à partir de la représentation mentale générique du domaine contextuel. Mais contrairement au processus d'interprétation des indéfinis, cette partition doit être récupérable indépendamment du prédicat de la phrase hôte de l'expression et avant

[80] Nous symboliserons les deux cas par la notation « Type > N ».
[81] Nous utilisons « type » et « catégorie » en tant que synonymes.

le processus d'interprétation proprement dit. C'est par exemple ce qui se passe pour l'interprétation de *la tête* en (150) :

(150) il y a donc une fillette qui se trouve à peu près juste au-dessous du soleil quoi, donc *la tête* avec un petit cercle (C9Eglise)

L'expression demande à s'interpréter dans un domaine de type > *TÊTE*, permettant d'isoler un objet de type *TÊTE*. Ce domaine est fourni ici par la représentation @F introduite pour *une fillette* : avant l'interprétation de *la tête*, celle-ci n'est pas en possession d'une partition. En revanche, elle peut, lorsque cela est nécessaire, récupérer des partitions génériques propres au concept *FILLETTE*. Parmi celles-ci, une partition donne accès aux composantes d'une fillette, dont un objet de type *TÊTE*. C'est cette partition qui remplit ici les contraintes du domaine sous-spécifié pour l'interprétation de *la tête*. Par conséquent, elle sera instanciée dans le domaine contextuel @F et c'est avec @F que va s'unifier le domaine de référence sous-spécifié pour l'interprétation de *la tête*.

- *L'opération de restructuration*

Une fois un domaine contextuel compatible identifié, l'expression définie restructure ce domaine par une opération d'extraction et de focalisation : elle en extrait l'élément qui possède la valeur recherchée (*N* ou *P*) pour le critère de différenciation et le focalise. La particularité de cette opération de restructuration consiste dans le fait qu'elle se réalise de façon préférentielle à travers un **changement de la structure focale de la partition** : cela signifie qu'une description définie focalise de préférence un élément non focalisé auparavant. Nous verrons dans la suite que c'est cette particularité qui permet de prédire la différence entre l'emploi d'une expression définie et celui d'une expression pronominale.

Figure 40 : Instruction de restructuration pour une expression définie

L'application de ce principe de restructuration à l'interprétation de *la table* de l'exemple (150) est représentée graphiquement dans la Figure 41, page 154. Cette expression s'interprète dans le domaine contextuel $@DR_C$ fourni par les figures du dessin : celui-ci regroupe des éléments de type *PERSONNE*, *CHAISE* et *TABLE*, dont le sur-type commun est *FIGURE*. De la partition *CATÉGORIE*, l'expression extrait l'élément de type *TABLE* et marque celui-ci comme focalisé. Étant donné que selon nos algorithmes de focalisation (7.3.4), *PERSONNE* était l'élément focalisé auparavant, l'interprétation passe bien par un changement de la structure focale de la partition.

- *Résumé de l'interprétation des expressions définies*

En résumé, l'interprétation d'une expression définie *le N(P)* est modélisée par

(a) la construction d'un domaine de référence sous-spécifié exprimant une contrainte sur le type des éléments : celui-ci doit être supérieur à *N* si l'expression est *le N*, et égal à *N*, si l'expression est *le N P*. Par ailleurs, le domaine exprime une contrainte sur l'existence d'une

partition avec un critère de différenciation : celui-ci porte sur la catégorie des éléments si l'expression est *le N,* et sur la propriété *P* si l'expression est *le N P*.

La recherche d'un domaine compatible se fait exclusivement parmi les domaines du modèle contextuel (exclusion des domaines génériques). En revanche, si la partition recherchée n'est pas déjà instanciée dans le domaine contextuel, il faut tester si elle peut l'être par héritage à partir de la représentation générique du domaine en question.

(b) l'opération de restructuration, qui consiste en l'extraction et la focalisation d'un élément de la partition : cet élément doit être isolable en vertu de sa valeur pour le critère de différenciation. Par ailleurs, il est de préférence non focalisé jusqu'alors.

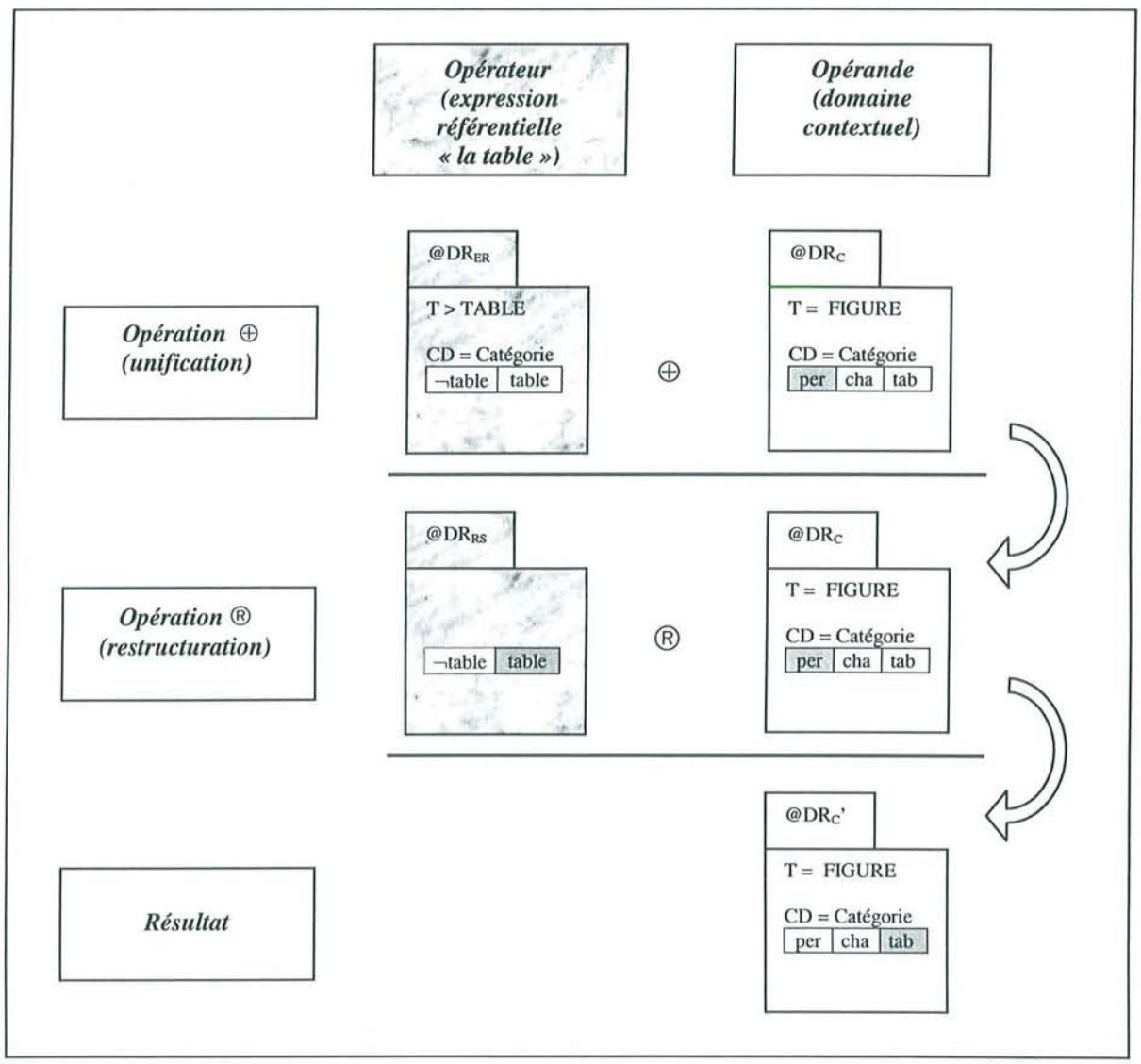


Figure 41 : Interprétation d'un défini : « la table » (exemple (149))

8.3.3 *Les expressions démonstratives*

- *Le domaine de référence sous-spécifié*

Contrairement aux définis, ce n'est pas le contenu linguistique de la description qui identifie le référent d'un démonstratif, mais son caractère *token-réflexif* (De Mulder, 1998) : par là, il faut entendre que les démonstratifs réfèrent à une entité nécessairement présente dans le contexte. Cela signifie que l'isolement du référent d'un démonstratif s'appuie par exemple sur une mention antérieure du référent ou sur un geste de désignation, guidant une focalisation perceptive (cf. la section 2.4).

Nous retenons donc pour notre modélisation que le référent d'une expression démonstrative *ce N* est identifié par des facteurs externes et non pas en raison de sa référence virtuelle. Par conséquent, le domaine de référence sous-spécifié d'un démonstratif ne contraint pas le type de ses éléments[82]. En revanche, il présuppose l'existence d'une partition, à critère de différenciation indifférent, mais à l'intérieur de laquelle un élément est mis en évidence par des facteurs externes à l'interprétation du démonstratif : cela se matérialise par un **domaine contenant une partition focalisée**, soit en raison du dialogue antérieur, soit en raison d'un geste de désignation. La représentation graphique de cette structure de domaine est donnée dans la Figure 42 :

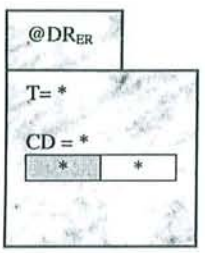


Figure 42 : Structure du DR_{ER} des démonstratifs – « ce N »

(151) I_1 au dessus de ça tu vas mettre un rond qui correspondra au soleil
I_2 et *ce rond* il faut qu'il soit au dessus de la deuxième pyramide (C8Egypte)

Appliqué à *ce rond* de l'exemple (151), le principe de construction du domaine sous-spécifié mène à un domaine $@DR_{ER}$ (Figure 44, page 156) qui a la même structure que celui de la Figure 42.

- *L'opération de restructuration*

Le principe d'identification externe du référent – modélisé à travers la récupération de l'élément focalisé – permet ensuite d'attribuer au contenu linguistique d'une expression démonstrative un rôle « quasi-attributif » : celui-ci peut être une nouvelle description ou même une reclassification du référent. Selon certains auteurs (Corblin, 1987), le pouvoir reclassificateur des démonstratifs serait sans contraintes linguistiques. Nous montrerons dans la section consacrée aux prédictions dérivées de notre modélisation que cette hypothèse doit être nuancée : le pouvoir reclassificateur semble se définir non seulement par rapport au référent présumé, mais plus généralement par rapport à la topographie focale de leur domaine de référence (cf. section 9.4).

Plus généralement, la section consacrée au fonctionnement des démonstratifs (2.4) a montré que l'effet interprétatif d'un démonstratif est d'instaurer une rupture par une mise en relief du référent. Cette rupture, qui marque la différence entre les expressions démonstratives et pronominales, se manifeste de différentes façons : reclassification du référent, introduction d'un nouvel état du référent,

[82] L'absence de contraintes est noté par « * ».

changement du statut thématique du référent ou insertion de celui-ci dans un nouvel univers du discours. Dans le cadre de notre modélisation, cet effet interprétatif est modélisé par une opération de restructuration qui consiste en **l'insertion du référent dans un nouveau domaine de référence** : ce nouveau domaine de référence est, pour une expression de la forme *ce N*, un domaine composé d'éléments de type *N*, à l'intérieur duquel le référent s'oppose aux autres éléments du domaine par un contraste interne, relevant (comme pour les indéfinis) des informations fournies par la prédication. La représentation graphique de cette opération de restructuration est donnée dans la Figure 43 :

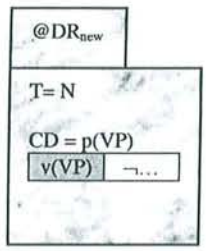


Figure 43 : Instruction de restructuration pour une expression démonstrative

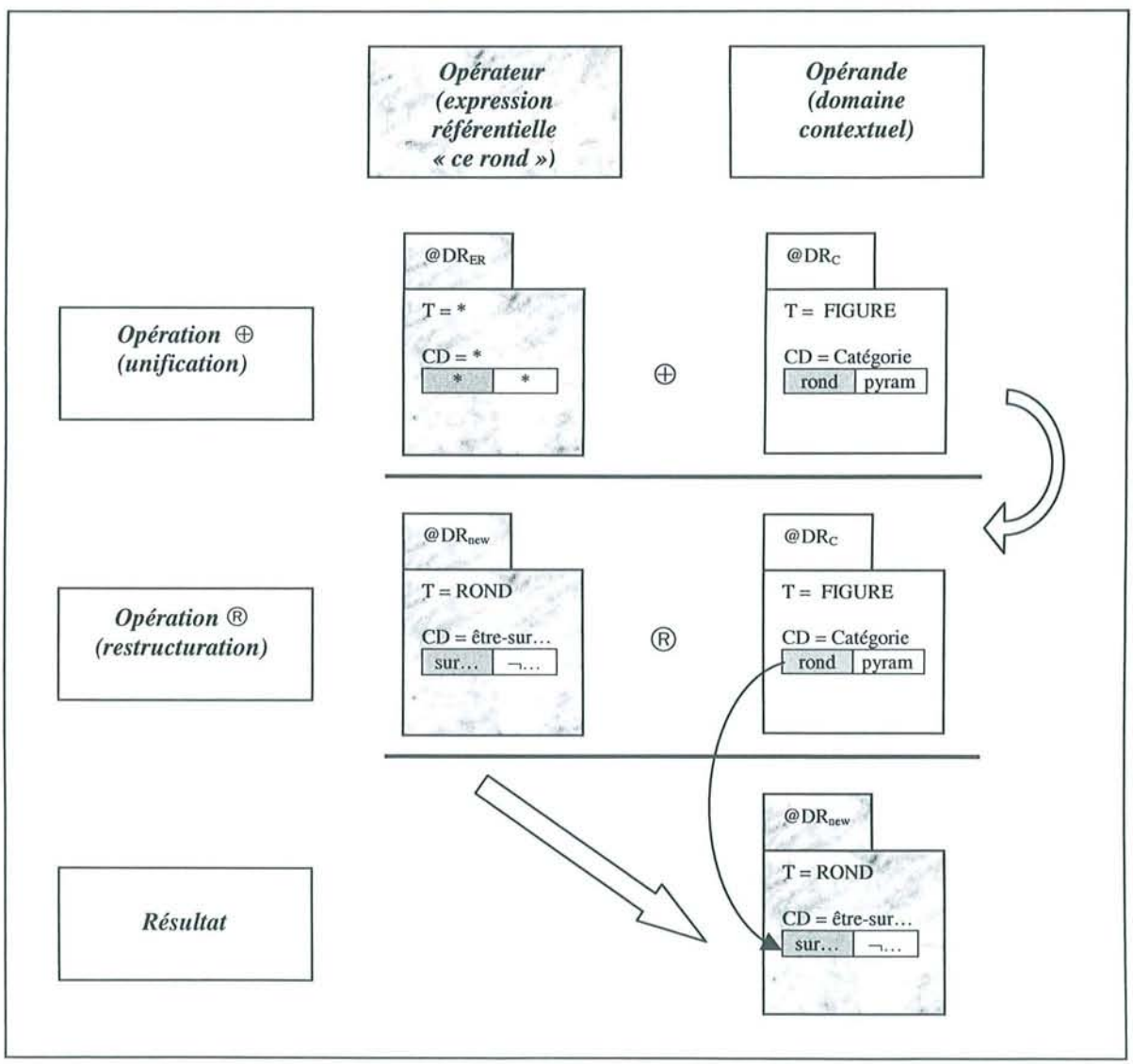


Figure 44 : Interprétation d'un démonstratif : « ce rond » (exemple (151))

L'application de cette opération de restructuration à l'interprétation de l'expression *ce rond* de l'exemple (151) est illustrée dans la Figure 44 : le domaine issu du contexte précédent est $@DR_C$: il regroupe des figures (un rond et une pyramide) et focalise la cible de l'action de positionnement qui est le rond. Le domaine sous-spécifié pour *ce rond* ($@DR_{ER}$) est compatible avec $@DR_C$ et donc unifié avec celui-ci. L'élément focalisé de ce domaine en est alors extrait comme référent. Il est inséré dans un nouveau domaine de type *ROND*, opposant le référent, qui reste focalisé, aux autres éléments de la partition. Celle-ci est justifiée par des informations prédicatives, ici la position verticale. Il est à noter que le complément de cette partition peut rester sans instanciation effective : c'est le cas ici, car il n'y pas d'autres ronds disponibles dans le contexte.

- *Résumé de l'interprétation des expressions démonstratives*

En résumé, le principe d'interprétation d'une expression démonstrative *ce N* est modélisé par :

(a) la construction d'un domaine de référence sous-spécifié sans contrainte sur le type des éléments[83], mais imposant l'existence d'une partition contenant un élément focalisé. La recherche d'un domaine compatible se fait parmi les domaines disponibles du contexte.

(b) la restructuration qui consiste à construire, autour du référent focalisé, un nouveau domaine d'éléments de type *N*, opposant le référent aux autres éléments du nouveau domaine en vertu d'informations prédicatives. Comme pour l'indéfini, la présence d'un modifieur P impose des contraintes supplémentaires sur la construction de ce nouveau domaine. Dans la mesure où de tels démonstratifs n'ont pas été relevés dans le corpus Ozkan, nous laissons pour l'instant ces cas de côté.

8.3.4 Les expressions pronominales de 3ième personne

- *Le domaine de référence sous-spécifié*

Dans la DRT, le pronom est traité par une mise en égalité de la variable correspondante au pronom avec celle d'une entité existante de l'univers discursif. Pour la sélection de cette entité, aucun mécanisme particulier n'est proposé : il doit s'agir simplement d'un « *suitable discourse referent*[84] » (Kamp et Reyle, 1993 : 70). La théorie du Centrage (Grosz et al., 1995) peut être considérée comme partiellement complémentaire à la DRT dans la mesure où elle explore des contraintes formulées sous forme d'enchaînements préférentiels. Celles-ci permettent de choisir l'antécédent d'un pronom parmi les entités introduites dans la phrase précédente, essentiellement sur la base d'un ordonnancement de ces dernières. Mais la théorie du Centrage ne traite que les pronoms à antécédent textuel et ne considère pas des discours dépassant une suite de deux séquences (cf. notre critique en 4.2.1).

Par rapport à ces modélisations, l'approche linguistique de G. Kleiber (1990), présentée en 2.5, a plusieurs avantages : d'une part, elle tient compte des différents usages du pronom (référent présent ou absent dans la situation, usage générique et non-générique) et d'autre part, elle postule une spécificité du pronom par rapport aux autres marqueurs référentiels. G. Kleiber considère en effet que le pronom est une expression référentielle possédant son propre mode de donation référentielle : son rôle spécifique est d'indiquer que l'on va parler d'un référent déjà saillant lui-même ou présent dans une situation saillante et que l'on va en parler en continuité avec ce qui l'a rendu saillant. Le pronom

[83] Il s'agit de la version extrême, modélisant le fait qu'un démonstratif est capable de reclassifier une entité sans limites. En réalité, il nous semble que cette vision des choses doit être nuancée. Nous y reviendrons dans la section 9.4.
[84] référent du discours convenable

renvoie alors à un référent conçu comme classifié et comme occupant une place d'argument dans une situation saillante.

La modélisation que nous proposons pour l'interprétation du pronom personnel essaie alors d'intégrer les contraintes dégagées par G. Kleiber (1990, 1994). Puisque *il* renvoie à un référent saillant et classifié, son référent doit être un élément focalisé dans un domaine comportant des éléments classifiés par l'appartenance à un type. Cela signifie que le domaine de référence sous-spécifié du pronom est un **domaine comportant un partition focalisée**, dont le type et le critère de différenciation ne sont pas spécifiés en avant. Une représentation graphique de ce domaine sous-spécifié est donnée par la Figure 45 :

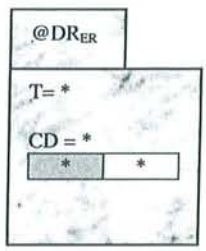


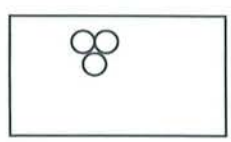
Figure 45 : Structure du DR_{ER} des pronoms personnels

En ne mettant aucune contrainte sur le type des éléments du domaine sous-spécifié, nous simplifions le traitement en laissant de côté les contraintes d'accord. Si l'on voulait intégrer ces contraintes, il faudrait tenir compte des réflexions suivantes : pour les non-humains[85], l'accord en genre doit chercher à se faire par rapport à la classification des éléments du domaine. Pour un certain nombre d'exemple du corpus Ozkan, on observe pourtant que cet accord ne se fait pas par rapport à une mention textuelle (immédiatement) antérieure, mais par rapport à une saillance situationnelle ou simplement par rapport à d'autres classifications possibles : ainsi, en (152), le pronom *le* en I_3 s'accorde en genre, non pas avec la classification en tant que *LIGNE VERTICALE*, mentionnée auparavant, mais en tant que *TRAIT*, choisie par le locuteur dans un énoncé qui se situe à quinze interventions en amont.

(152) I_0 et puis on va prendre le grand trait vertical
...
I_1 et tu vas reprendre encore une ligne verticale
M_1 une grande ?
I_2 une grande oui
I_3 et tu vas *le* placer en parallèle avec un petit écart (C5 Route)

Dans l'exemple (153), le pronom *la* (I_8) réfère à un élément introduit en I_6 comme *un autre (rond),* donc masculin. Mais au cours du dialogue antérieur s'est installé une double classification de ces objets, d'une part comme *ROND* (en tant qu'objet à extraire de la palette) et d'autre part comme *BOULE* (en tant que objet faisant partie de la scène en construction).

(153) I_1 ensuite prendre un autre rond
I_2 et le mettre euh sous les deux boules
M_3 ah
I_4 et je crois qu'il en faut quatre autres
I_5 et les mettre autour de celle qui est au milieu
I_6 quatre autres petits ronds enfin un autre
M_7 autour de laquelle ?
I_8 voilà, tu *la* colles contre (C11Lampe)



[85] En ce qui concerne les humains, la variation pour le genre est disponible, à condition que cela apporte de l'information (cf. la section en 2.5).

Dans ces cas, la contrainte en genre porte donc non par sur la classification par une mention antérieure immédiate, mais sur une autre classification, déductible de l'histoire dialogique, ou, à défaut, de la hiérarchie de types.

En ce qui concerne l'accord en nombre, il s'applique sur la cardinalité de la représentation du référent présumé. Les glissements au générique (aucune occurrence dans le corpus Ozkan) sont traitables en admettant une extension de la contrainte en nombre à la classe dont le référent est une instance.

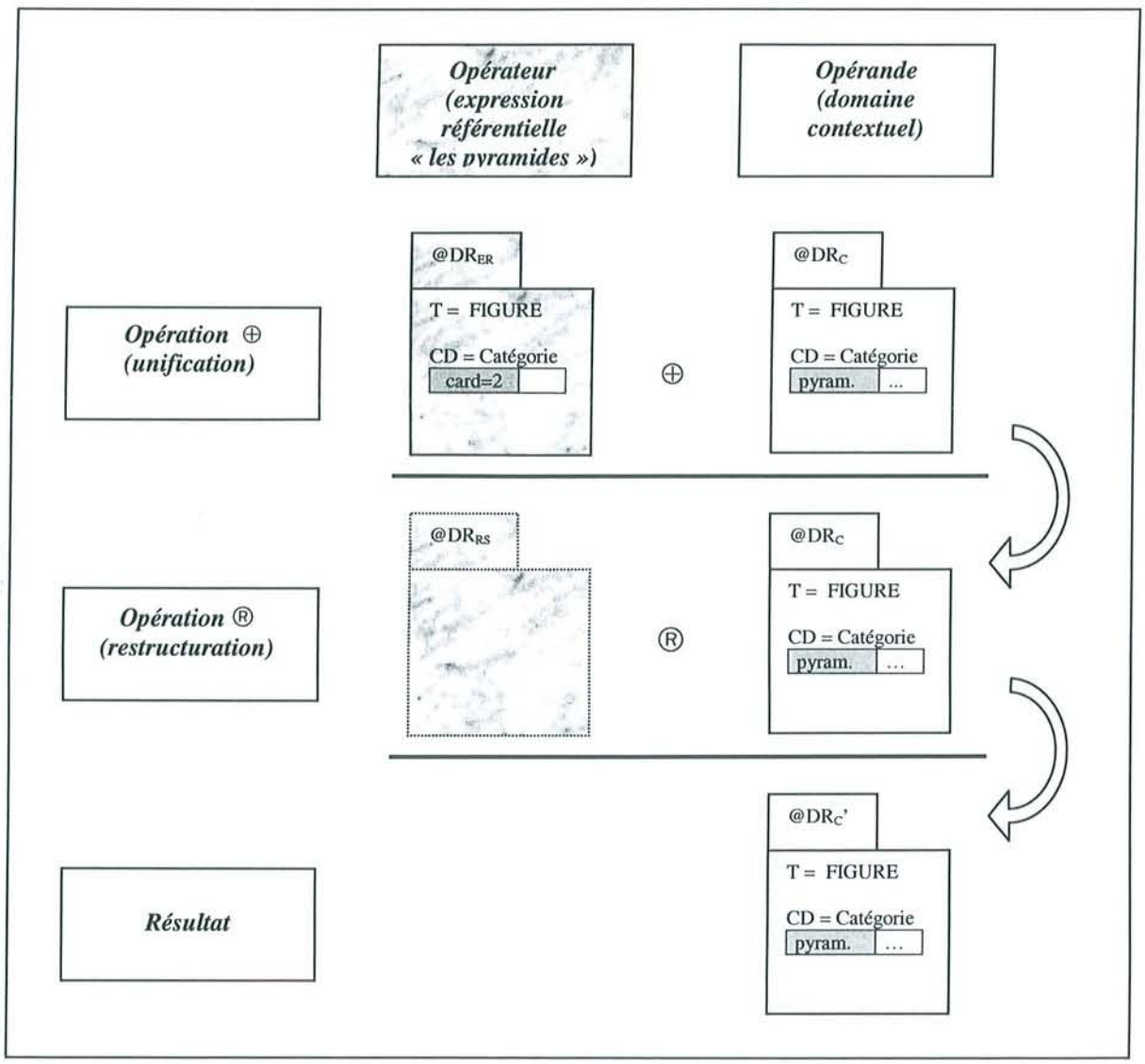


Figure 46 : Interprétation d'un pronom : « les » (exemple (154))

- *L'opération de restructuration*

La particularité du pronom consiste à indiquer que l'on parle d'un référent saillant et ceci en continuité avec ce qui l'a rendu saillant. Cette particularité est exprimée par la modélisation de l'opération de restructuration propre aux pronoms : celle-ci **ne modifie pas le domaine d'interprétation**. Cette modélisation assure la compatibilité de notre proposition avec tous les travaux qui considèrent le pronom comme une marque de continuité thématique (Kleiber, 1990 ; Grosz et al., 1995 ; Cristea et al., 1998 et d'autres).

A titre d'illustration, les principes exposés sont appliqués au pronom *les* de (154) :

(154) I_1 ouais tu décales les pyramides, c'est-à-dire tu *les* descends (C8Egypte)

Le domaine sous-spécifié pour *les* est donné par le domaine $@DR_{ER}$ de la Figure 46. Par ailleurs, le modèle contextuel contient un domaine de type *FIGURE* dont un ensemble de deux pyramides a été extrait et focalisé auparavant ($@DR_C$). Ce domaine est compatible avec les contraintes imposées par le domaine sous-spécifié $@DR_{ER}$. Par ailleurs, l'élément focalisé de ce domaine remplit la contrainte de cardinalité. Il est alors ré-identifié comme référent et ceci sans changement de structure de $@DR_C$.

- *Résumé de l'interprétation des expressions pronominales*

En résumé, l'interprétation d'une expression pronominale est modélisée par

(a) la construction d'un domaine de référence sous-spécifié sans contrainte sur le type des éléments. En revanche, il impose l'existence d'une partition contenant un élément focalisé et, sous certaines conditions, des contraintes supplémentaires sur le genre ou le nombre de l'élément focalisé. La recherche d'un domaine compatible se fait parmi les domaines disponibles du contexte.

(b) l'opération de restructuration qui est vide : on n'opère aucun changement sur la structure du domaine de référence identifié.

8.4 Synthèse : le tableau complet

A travers la modélisation proposée dans les quatre sous-sections précédentes, nous avons exprimé une relation forte entre l'emploi de différents types d'expressions référentielles (indéfinis, définis, démonstratifs et pronominaux) et différentes structures contextuelles avant et après l'interprétation référentielle. Cette relation a été formulée, type d'expression par type d'expression, par des domaines de référence sous-spécifiés (contraintes imposées sur la structure d'un domaine d'interprétation avant l'interprétation) et par des instructions de restructuration (mise à jour du modèle contextuel). Le Tableau 20, construit au fur et à mesure de l'exposition de notre modèle, donne un aperçu synthétique de la compatibilité de différents types d'expressions référentielles avec différentes structures contextuelles. Les lignes contiennent les représentations schématiques pour les trois structures contextuelles fondamentales, introduites au chapitre précédent : domaine sans partition (ligne 5), domaine partitionné sans élément focalisé (ligne 6) et domaine partitionné avec élément focalisé (ligne 7). Les colonnes concernent chacune un type d'expression référentielle : conformément aux sous-sections précédentes, elles fournissent la représentation schématisée de la structure du domaine de référence sous-spécifié (ligne 3) et celle des instructions de restructuration (ligne 4).

Ce tableau montre bien que toutes les structures introduites lors de la modélisation des domaines de référence sont exploitées pour le calcul référentiel : aussi bien les contraintes sur les domaines compatibles que les opérations de restructuration font intervenir le type du domaine, les partitions et/ou les structures focales. Une comparaison des contraintes et opérations de restructuration est particulièrement révélatrice à cet égard :

Type d'expression	*Contrainte sur un domaine compatible*	*Opération de restructuration*
Démonstratif	Existence d'un domaine partitionné et focalisé	Création d'un nouveau domaine
Indéfini	Existence d'un domaine	Création d'une nouvelle partition
Défini	Existence d'un domaine partitionné	Création d'une nouvelle structure focale
Pronom	Existence d'un domaine partitionné et focalisé	Aucune

Tableau 19 : Comparaison des contraintes et restructurations par type d'expression

Pour résumer, voici brièvement les principales caractéristiques : un indéfini contraint uniquement le type des éléments de son domaine d'interprétation, dont il extrait son référent par la création d'une nouvelle partition. Un défini contraint le type des éléments de son domaine et impose l'existence d'une partition. Il extrait son référent de cette partition, de préférence non focalisé jusqu'alors. Un démonstratif demande l'existence d'un domaine contenant une partition avec un élément focalisé. L'opération de restructuration consiste à insérer cet élément, qui sera le référent, dans un nouveau domaine de référence correspondant au type de l'expression démonstrative. Un pronom demande, comme le démonstratif, un domaine à partition focalisée, mais se distingue du démonstratif par une opération de restructuration nulle, assurant la continuité domaniale.

Le corps du Tableau 20 indique pour chaque type d'expression référentielle le résultat de l'opération de restructuration en fonction de la structure des domaines disponibles du contexte. Son commentaire fera l'objet du chapitre suivant. Mais auparavant, nous nous arrêterons brièvement sur les groupes nominaux sans noms pour montrer que le traitement de ceux-ci s'intègre de façon régulière dans notre modélisation.

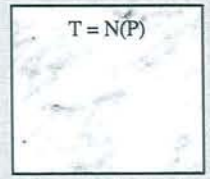
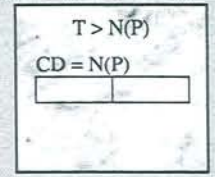
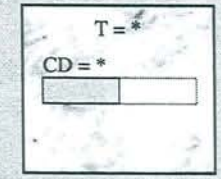
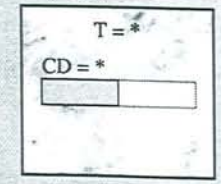
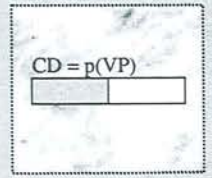
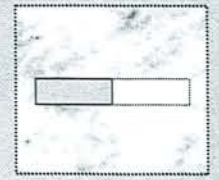
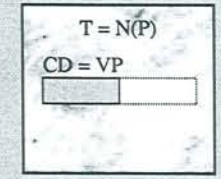
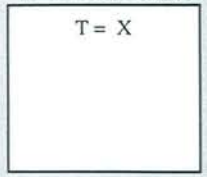
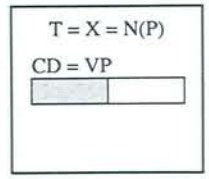
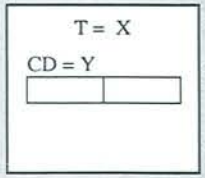
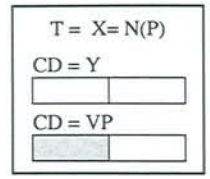
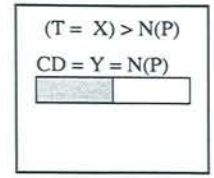
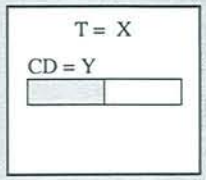
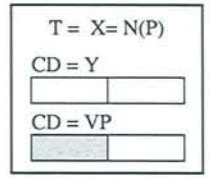
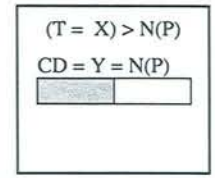
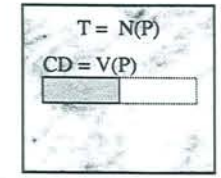
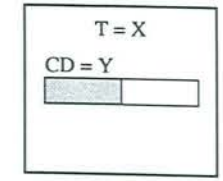
Type d'expression *Structure abstraite : Exemple*	*Expression indéfinie* un N(P) VP. : *Colorie une figure.*	*Expression définie* le N(P) VP. : *Colorie le triangle.*	*Expression démonstrative* ce N(P) VP. : *Colorie ce triangle.*	*Expression pronominale* il VP. : *Colorie le.*
Domaine de référence sous-spécifié	Contrainte sur le type des éléments 	Contrainte sur l'existence d'un partition 	Contrainte sur la focalisation 	Contrainte sur la focalisation 
Opération de restructuration	Création d'une nouvelle partition 	Extraction et focalisation 	Insertion dans un nouveau domaine 	Aucune
Prends deux triangles. 				
Prends un triangle et un carré. 				
Mets un triangle sur un carré. 				

Tableau 20 : Compatibilité entre structures contextuelles et expressions référentielles

8.5 Intégration des groupes nominaux sans nom

Au chapitre consacré à la description linguistique des marqueurs référentiels, nous avons présenté une classe à part : les groupes nominaux sans nom (section 2.6). Cette mise à part a été justifiée par des caractéristiques morpho-syntaxiques et interprétatives. D'un point de vue morpho-syntaxique, il s'agit de syntagmes nominaux sans tête nominale et plus précisément

de groupes nominaux indéfinis à *en* quantitatif : *j'en veux un bleu, j'en prends un, ... ;*

de groupes nominaux elliptiques : *le rouge, le grand, ... ;*

de pronoms possessifs : *la sienne, le tien,... ;*

de mentionnels : *le premier - le second, l'un - l'autre* ;

de composés de *celui* : *celui qui..., celui de..., ... ;*

de *celui-ci.*

La particularité interprétative de ces constructions découle de leur forme morpho-syntaxique : puisque leur tête nominale n'est pas fixée *in situ*, elle doit être récupérée dans le contexte textuel ou situationnel et ceci avant le calcul du référent de la totalité du groupe.

Un aspect intéressant de notre modélisation est qu'elle permet d'intégrer ces marqueurs dans un cadre tout à fait régulier. Leur interprétation suit en effet les principes propres aux déterminants qui les accompagnent, avec une particularité commune à toutes ces expressions : le type des éléments du domaine de référence sous-spécifié n'est pas précisé. Cette donnée demande donc à s'instancier en fonction des domaines disponibles dans le modèle contextuel. On notera que cette particularité n'est pas une exclusivité de cette classe : les pronoms (et les démonstratifs[86]) possèdent la même caractéristique. Par ailleurs, les particularités des mentionnels et de *celui-ci* (cf. les sections 2.6.2 et 2.6.3) sont dues seulement à la spécificité du critère de différenciation entrant en jeu : au lieu d'être dérivé d'une propriété intrinsèque comme la couleur, il repose sur l'ordonnancement discursif pour les mentionnels et sur une saillance situationnelle pour *celui-ci.* Dans la suite, nous présenterons brièvement la construction des domaines sous-spécifiés et les opérations de restructuration pour chacune de ces expressions, en les ordonnant selon le schéma de base instancié.

- *Instanciation du schéma des indéfinis*

L'interprétation des groupes nominaux indéfinis à *en* quantitatif (*j'en veux un bleu ; j'en prends un*) mène à la construction d'un domaine de référence sous-spécifié sur le modèle des indéfinis, à l'exception de la spécification du type des éléments. Comme pour les indéfinis, certaines propriétés des éléments du domaines sont contraintes en fonction des attributs de la composante indéfinie (*bleu*). Contrairement aux indéfinis à tête pleine, l'opération de sélection d'un domaine compatible se fait uniquement parmi les domaines contextuels (à l'exclusion des domaines génériques) : cela s'explique par l'absence d'indications sur le type recherché. Enfin, la restructuration est celle des indéfinis : elle mène à la création d'un nouvelle partition, opposant le référent extrait aux autres éléments du domaine.

[86] Nous mettons les démonstratifs entre parenthèses, car si – dans la version que nous donnons de leur interprétation – ils possèdent effectivement cette propriété, nous nous attachons tout de même à faire remarquer que cette vision de l'interprétation des démonstratifs doit être nuancée (cf. la section 9.4).

- *Instanciation du schéma des définis*

Les groupes nominaux définis elliptiques (*le rouge*), mentionnels (*le premier*), possessifs (*le sien*) et les composés de *celui* (*celui de droite*) hors *celui-ci* sont traités sur le modèle des expressions définies. Cela signifie qu'ils construisent un domaine de référence sous-spécifié (sans contraintes sur le type des éléments) possédant obligatoirement une partition, dont le critère de différenciation correspond au modifieur : il peut s'agir d'une propriété physique comme la couleur (*rouge*), d'une propriété relationnelle comme la position (*de droite*), d'un ordonnancement (*premier*) ou encore d'une relation d'appartenance (*sien*).

La seule particularité des mentionnels *le premier – le second* consiste dans la manière dont le critère de différenciation de type *ORDONNANCEMENT* est saturé : contrairement à des ordonnancements externes à la matérialité du texte (ordre spatial, ordre temporel), il s'agit ici d'un ordre discursif, donc interne à la matérialité du texte.

En ce qui concerne la paire *l'un – l'autre*, *l'autre* est traité de façon régulière par rapport au principe des définis : il présuppose une partition existante. Comme seule particularité, le critère de différenciation repose cette fois-ci sur l'altérité : *autre* demande une partition ayant déjà isolé un (autre) référent, c'est-à-dire une partition obligatoirement focalisée. L'opération de restructuration s'en ressent : conformément aux définis, il s'agit d'un changement de la structure focale, mais contrairement aux définis, celui-ci est obligatoire pour *autre*[87]. L'exemple (155) illustre ce principe :

(155) I_1 donc au bord de cette route, il y a deux maisons,
I_2 donc une maison qui se trouve à gauche de cette route et *l'autre* à droite. (C9Forêt, modifié)

Le domaine contextuel, avant l'interprétation du groupe nominal sans nom *l'autre* (I_2) est celui issu de l'interprétation de l'exemple (146), page 148 : elle est reproduite par le $@DR_C$ de la Figure 47. Conformément aux principes appliqués aux groupes nominaux sans nom, le domaine de référence sous-spécifié pour *l'autre* $@DR_{ER}$ ne contraint pas le type des éléments du domaine d'interprétation. Conformément aux principes propres définis, il pose une contrainte sur l'existence d'une partition. Enfin, conformément à la sémantique de autre, cette partition doit être focalisée. $@DR_{ER}$ est donc compatible avec $@DR_C$, avec lequel il sera unifié. L'opération de restructuration est celle des définis, c'est-à-dire le changement de la structure focale de la partition, ce qui donne $@DR_{C'}$ comme résultat de l'interprétation.

Seul *l'un* sort du paradigme des définis : cette expression semble ne pas présupposer l'existence d'une partition (*J'ai vu deux médecins. L'un...*). Pour l'instant, nous la traitons donc comme un groupe nominal indéfini *en... un*. Pourtant, elle s'en différencie par le fait que cette construction semble incomplète sans une suite en *l'autre* ou *le second* (*J'ai vu deux médecins. L'un était incompétent.*). La modélisation de son fonctionnement demanderait donc d'intégrer en plus une contrainte sur le parcours obligatoire de la suite de la partition amorcée.

- *Instanciation du schéma des démonstratifs*

Enfin, le marqueur *celui-ci* s'interprète selon le schéma des démonstratifs : il demande un domaine dont le type des éléments n'est pas spécifié et ayant mis en relief un référent par des principes extérieurs, comme la mention la plus récente ou un geste de désignation. Comme pour un démonstratif, l'opération de restructuration est la construction d'un nouveau domaine. Celui-ci

[87] Cette présentation du fonctionnement de *autre* est nécessairement brève, mais nous renvoyons le lecteur intéressé au chapitre 10, exclusivement consacré à la validation du fonctionnement de *autre* sur les données du corpus Ozkan.

instaure un contraste interne entre le référent et d'autres éléments de son type, ou entre le référent et d'autres mentions antérieures (en mentionnel).

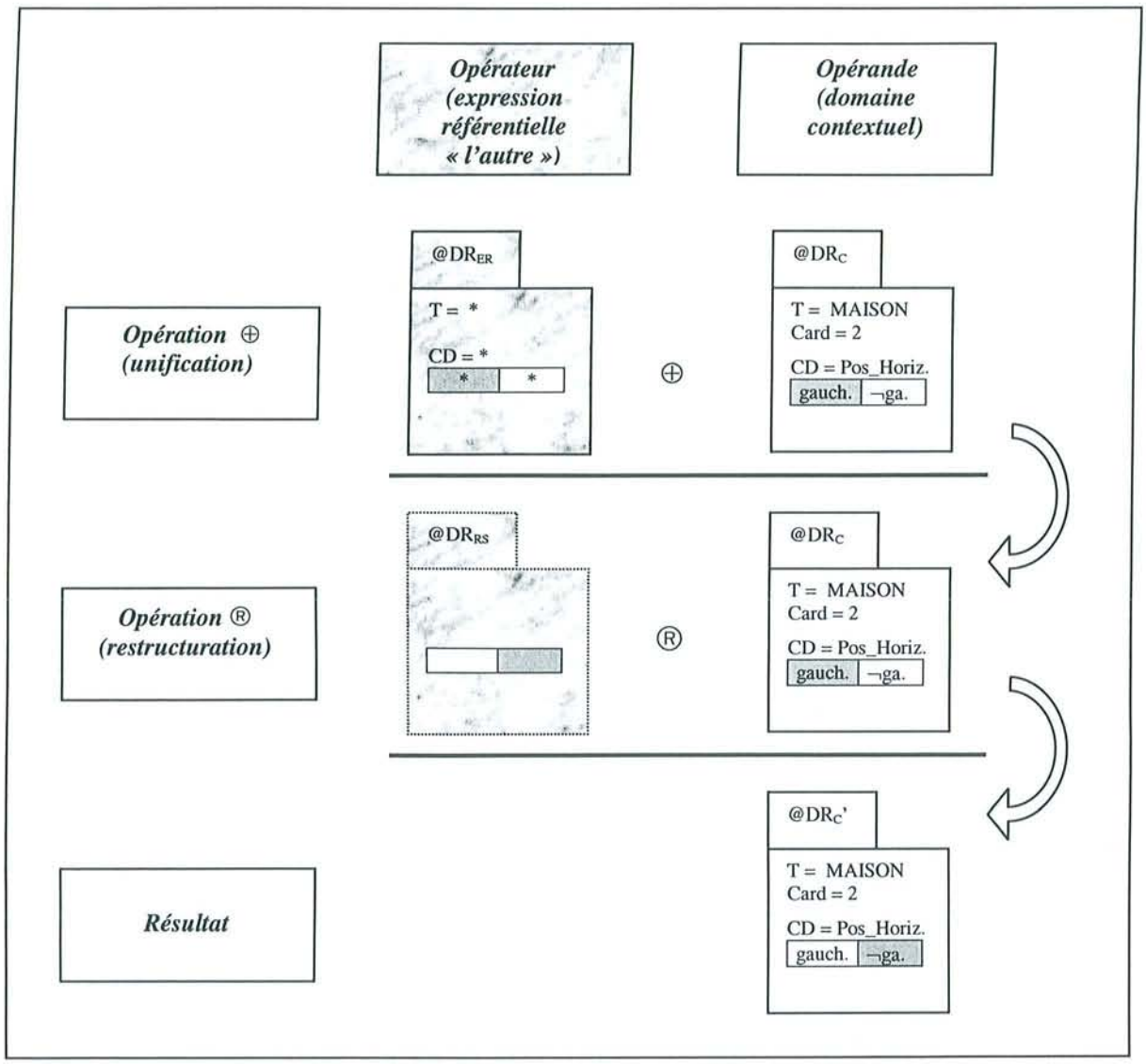


Figure 47 : Interprétation d'un groupe nominal sans nom : « l'autre », exemple (155)

- *Résumé du principe d'interprétation des groupes nominaux sans nom*

En résumé, l'interprétation d'un groupe nominal sans nom est modélisée par :

(a) la construction d'un domaine de référence sous-spécifié qui suit les procédures spécifiques à la détermination du groupe et à la sémantique des constituants, avec une particularité : il n'y pas de contraintes sur le type des éléments du domaine ;

(b) l'opération de restructuration qui correspond à celle définie pour le type de détermination du groupe.

9 Mise en œuvre du modèle : prédictions et applications

9.1 Introduction

Ce chapitre est consacré à la discussion des conséquences du tableau Tableau 20 (page 162), dont le cadre – les structures contextuelles possibles et les instructions attachées aux différents types d'expressions référentielles (domaine sous-spécifié et opération de restructuration) – a été introduit au chapitre précédent. Le corps du tableau indique pour chaque type d'expression référentielle le résultat de l'opération de restructuration en fonction de la structure du domaine d'interprétation. A titre d'exemple, le croisement de la ligne 5 avec la colonne 2 se lit de la façon suivante : l'interprétation d'un indéfini *un N* dans un domaine d'éléments de type *X* sans partition préalable (*Prends deux triangles. Colorie une figure.*) mène, à condition que les types *N* et *X* soient compatibles[88], à la création d'une nouvelle partition dans le domaine contextuel, à l'extraction du référent de cette partition et à la focalisation du référent. C'est sur cette base que nous sommes capable de formuler un certain nombre de prédictions sur l'emploi des différents types d'expressions référentielles.

Premièrement, à partir des connaissances sur les pré-conditions d'interprétation formulées par les domaines de référence sous-spécifiés, ce tableau permet de prédire si une expression donnée est interprétable dans tel ou tel domaine contextuel. A partir de la structure du domaine sous-spécifié des pronoms, on déduit par exemple qu'un pronom ne peut pas être interprété en (156), car il demande un domaine partitionné (157) .

(156) ? Prends deux triangles. Colorie *le*.

(157) Prends deux triangles. Agrandis celui de gauche. Colorie *le*.

Deuxièmement, le tableau permet de prédire, à travers l'opération de restructuration propre à chaque marqueur référentiel, des effets interprétatifs plus ou moins optimaux. Ainsi, en supposant que le défini extrait d'une partition existante un élément de préférence non focalisé, nous sommes capable de prédire les intuitions que l'on peut avoir sur l'emploi du défini en (158) et en (159).

(158) Prends un triangle. Colorie *le triangle*.

(159) Prends un triangle et un carré. Colorie *le triangle*.

La présentation détaillée ainsi que la discussion critique de ces prédictions feront l'objet de ce chapitre. De plus, nous confronterons notre modèle à des données réelles, issues du corpus Ozkan. Pour montrer que notre modélisation s'applique au-delà des brefs exemples traités jusqu'ici, nous l'illustrerons sur des extraits plus conséquents du dialogue *C5Egypte*, déjà présenté et commenté au chapitre introductif (section 1.3.2, page 19). Pour faciliter le rappel, nous reproduirons les extraits en question avant leur traitement.

[88] La compatibilité entre types est formulée de façon précise dans la section présentant les algorithmes de fusion entre domaines sous-spécifiés et domaines contextuels (11.3.3) : très brièvement, deux types sont compatibles, lorsqu'ils sont égaux, lorsque le type du domaine sous-spécifié est un sur-type de celui du domaine contextuel ou lorsque le type du domaine sous-spécifié n'est pas spécifié (noté « * »).

9.2 Les expressions indéfinies : prédictions et exemple de traitement

9.2.1 Prédictions

- *L'interprétation d'un indéfini est indépendante du contexte.*

Cette caractéristique fondamentale des indéfinis, mise en avant par les études linguistiques (section 2.2), est intégrée dans notre modélisation par le fait que l'interprétation d'un indéfini *un N* est indépendante de l'existence d'une partition préalable de son domaine d'interprétation. Cela se traduit dans le Tableau 20 par la possibilité d'interpréter un indéfini quelle que soit la structure du domaine contextuel (absence de cases vides dans la colonne des indéfinis). Par ailleurs, il y même la possibilité d'une interprétation indépendante de tout domaine contextuel, car le domaine de référence d'un indéfini peut être, à défaut d'un ensemble contextuel, la classe générique des *N*.

Pour isoler son référent à l'intérieur du domaine compatible (donné par le contexte discursif ou visuel), l'indéfini crée systématiquement une nouvelle partition dont il extrait un élément quelconque. Le schéma de la Figure 48 le montre pour l'interprétation de *un rond*, en I_3 de l'exemple (160). Cette extraction aléatoire est un deuxième indice de l'indépendance contextuelle : quelle que soit l'origine ou la structure du domaine, le référent en est extrait au hasard.

(160) I_1 ensuite prendre *un rond* enfin un petit cercle
M_1 hum [geste de saisie]
I_2 et le mettre à gauche de la barre en haut
I_3 et prendre *un rond* (C5Lampe, modifié)

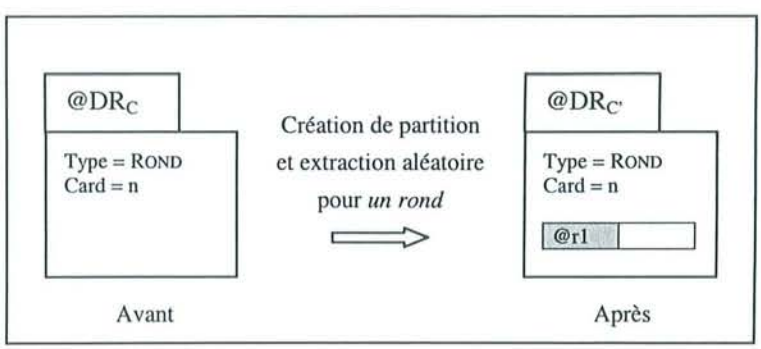


Figure 48 : Extraction d'un domaine non partitionné par création d'une nouvelle partition – exemple (160)

Cependant, la création systématique d'une nouvelle partition ne signifie par pour autant qu'un domaine pré-partitionné soit incompatible avec l'interprétation d'un indéfini (lignes 6 et 7 du Tableau 20). Mais dans ce cas, une deuxième partition est créée à l'intérieur de ce domaine et c'est cette nouvelle création qui traduit l'exclusion de toute connexion avec les extractions précédentes : dans un exemple tel que (160), l'expression *un rond* en I_3 ne permet effectivement pas de conclure ni à l'identité ni à la disjonction référentielle avec le référent de *un rond* en I_1[89] :

[89] Les connaissances spécifiques à la tâche ainsi qu'une préférence pour une lecture disjonctive en l'absence de marques d'identité (Corblin, 1987) laisseraient penser qu'une interprétation en disjonction référentielle est préférée, mais la formulation telle qu'en (160) nous semble rester douteuse.

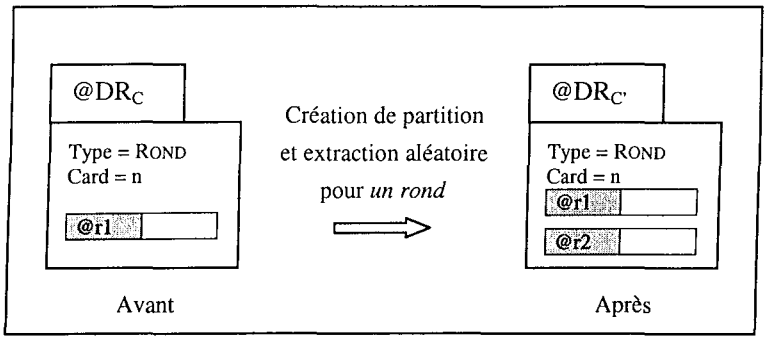


Figure 49 : Création d'une nouvelle partition dans un domaine pré-partitionné – exemple (160)

En effet, le domaine partitionné des éléments de type ROND, issu de l'interprétation de I_1, est re-partitionné une seconde fois par l'extraction d'un deuxième élément au hasard (Figure 49). Cet élément peut tout aussi bien correspondre à celui extrait précédemment qu'être un élément différent du premier. Si notre modélisation n'exclut pas, dans l'absolu, de tels coréférences (dont aucune occurrence réelle n'a été trouvée dans le corpus Ozkan), elle permet de prédire pourquoi celles-ci donnent le sentiment d'être difficilement acceptables : la création et le maintien d'une deuxième partition d'un même domaine semblent impliquer des efforts cognitifs démesurés par rapport à l'effet interprétatif, qui est en fait une incertitude sur l'identité référentielle de l'objet extrait. Considérer « *la condition de nouveauté comme défaut* » (Corblin, 1994), c'est-à-dire préférer une lecture disjonctive, permettrait alors d'éviter la création d'une nouvelle partition . Dans ce cas, l'indéfini serait systématiquement extrait de la partition existante, traitement qui rejoint en effet la modélisation que nous proposons pour les indéfinis assortis d'un indice explicite de disjonction référentielle (*autre, second* etc.), comme nous le verrons ci-dessous.

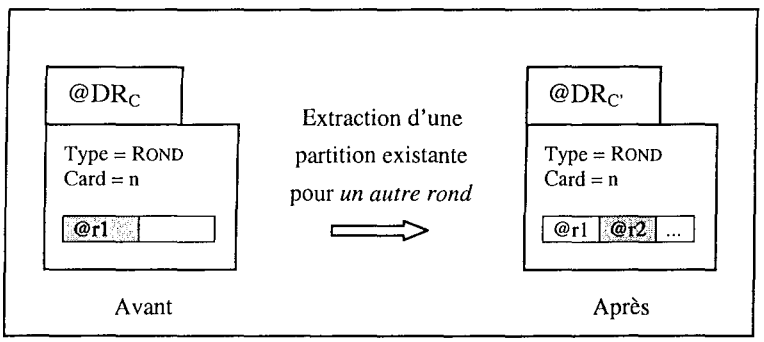
- *La dépendance contextuelle d'un indéfini doit être explicitement marquée.*

Les efforts cognitifs supplémentaires liés à la création et la maintien d'une deuxième partition en cas d'incertitude référentielle (cf. (160) ci-dessus) nous permettent de prédire que la dépendance contextuelle d'un indéfini, c'est-à-dire son interprétation relativement à des extractions précédentes, doit être explicitement marquée. L'indication explicite contribue en effet à éviter ces efforts : au lieu de créer une nouvelle partition, l'indéfini s'interprétera dans la partition existante. Par la suite, nous allons étudier plus en détail les deux possibilités d'interprétation dépendante du contexte : l'interprétation en disjonction référentielle et l'interprétation par coréférence.

Les marques de disjonction référentielle se manifestent souvent par des attributs à sémantique particulière. Il s'agit d'adjectifs tels que *autre* ou *deuxième*, dont le contenu sémantique impose des contraintes particulières aux domaines de référence sous-spécifiés (cf. 8.5) : ils doivent comporter une partition préalable dont un élément a déjà été extrait. Cette partition amorcée par une première extraction est alors réutilisée pour une deuxième extraction. Cela assure non seulement la disjonction référentielle des éléments extraits, mais évite aussi les efforts cognitifs supplémentaires liés à la création et le maintien d'une deuxième partition. Ainsi, en (161), l'expression indéfinie *un autre rond* (I_3) comporte une marque explicite de disjonction référentielle, présupposant une partition existante. Conformément aux principes des indéfinis, *un rond* extrait alors son référent de façon aléatoire, mais cette fois-ci non pas d'une nouvelle partition, mais du complément de la partition pré-existante (Figure 50).

(161) I_1 ensuite prendre un rond enfin un petit cercle
M_1 hum [geste de saisie]
I_2 et le mettre à gauche de la barre en haut
I_3 et prendre *un autre rond* (C5Lampe)

Figure 50 : Extraction en disjonction référentielle – exemple (161)



En plus des attributs mentionnés, d'autres marques indiquant une disjonction référentielle ont été observées dans le corpus Ozkan : il s'agit de marques portant non pas directement sur la différence des éléments d'une partition, mais sur la disjonction référentielle des événements dont les entités sont des participants :

(162) I_1 et tu vas prendre une grande barre que tu vas mettre entre les deux triangles
I_2 voilà et tu vas *reprendre une barre* pour continuer dans les autres (C8Egypte)

En (162), *reprendre une barre* dénote en effet l'itération d'un même type d'événement (*PRENDRE_UNE_BARRE*), dont les instances peuvent être considérées comme différentes, en particulier en raison de leur caractéristiques temporelles non identiques. Cette non-identité événementielle n'exclut pas dans l'absolu l'identité de leur participants, mais elle pourra peut-être être considérée comme un indice pour la disjonction référentielle de leur participants. Comme une modélisation de ces effets demande d'abord l'intégration de la représentation des événements dans la modélisation, ces cas ne sont pas traités pour l'instant.

Contrairement à une interprétation contextuellement dépendante par disjonction référentielle, les cas de coréférence (c'est-à-dire d'identité référentielle) entre indéfinis semble être rares. Comme l'ont montré un certain nombre de travaux, ils sont soumis à des conditions discursives particulières (Corblin, 1994 ; Danlos, 1999 ; Danlos et Gaiffe, 2000). Une piste possible est fournie par les travaux récents de L. Danlos sur la coréférence événementielle : il en ressort que des relations discursives de généralisation ou de particularisation entre deux propositions impliquent une coréférence événementielle entre les événements relatés. De cette coréférence événementielle découle alors une coréférence entre les participants, comme en (163) :

(163) Aurore a modifié *une figure*. Elle a colorié *un triangle*.

Nous avons tenté d'appliquer ces principes à certains exemples du corpus Ozkan, tels que (164) et (165) :

(164) I_1 et tu prends *une deuxième barre*
I_2 *une petite* (C5Egypte)

(165) I_1 alors, il va falloir que tu fasses *un toit*
I_2 il faut que tu mettes *un grand triangle*

Le cadre formel du traitement proposé (Salmon-Alt, Gaiffe et Romary, 2000) suit L. Danlos et B. Gaiffe (2000) : il repose sur la S-DRT, dont nous avons relevé, au chapitre 4, un certain nombre de

problèmes jugés centraux par rapport à notre préoccupation principale. Étant donné la fréquence relativement faible de ces configurations (6,3 % des indéfinis du corpus Ozkan), nous avons pour l'instant laissé de côté leur traitement, qui devra reposer sur une intégration des événements et propositions dans notre modélisation.

- *Les interprétations génériques découlent de notre modélisation.*

Une dernière prédiction sur les indéfinis concerne leur interprétation générique. Nous avons vu au chapitre 2 (section 2.2.2) qu'un indéfini *un N* s'interprète de manière générique, lorsque la propriété qui en a été prédiquée s'applique à toutes les occurrences de type *N*. Nous avons également constaté, au chapitre 4, que les possibilités d'interprétation générique étaient largement absentes des modélisations courantes du calcul référentiel.

Notre modélisation arrive à faire cette prédiction par l'opération de restructuration propre aux indéfinis : cette opération consiste en la création d'une nouvelle partition dans le domaine de référence. Or, une partition se justifie seulement sur la base d'un critère de différenciation, c'est-à-dire d'une propriété permettant de caractériser individuellement les éléments de la partition. Dans le cas des indéfinis, nous avons retenu comme critère de différenciation la propriété attribuée au référent par le prédicat de l'énoncé. Il s'ensuit alors que cette propriété ne peut plus fonctionner comme critère de différenciation dans un domaine où tous les éléments possèdent cette même propriété[90]. Par conséquent, l'extraction individuelle d'un élément échoue. Dans ce cas, sont retenus comme référent de l'expression tous les éléments vérifiant cette propriété, c'est-à-dire l'ensemble des éléments du domaine disponible :

(166) Un carré a quatre côtés.

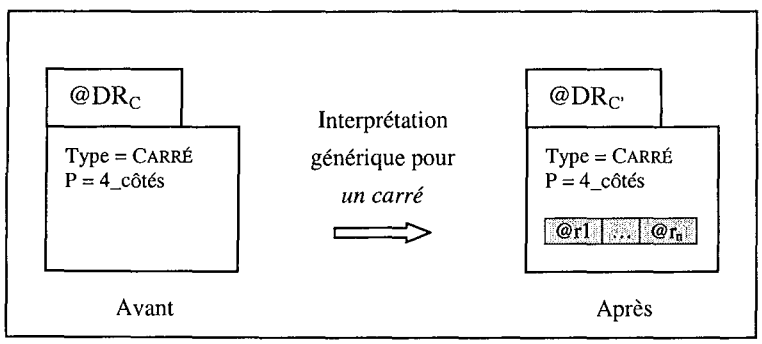


Figure 51 : Interprétation générique d'un indéfini – exemple (166)

9.2.2 *Exemple de traitement*

Afin d'illustrer la mise en œuvre de notre modélisation sur un dialogue réel, nous proposons de traiter un extrait qui porte sur une série référentielle d'indéfinis désignant des barres horizontales nécessaires à la composition d'une ligne d'horizon[91] :

(167) I_{14} et puis maintenant faudrait faire la ligne d'horizon
I_{15} il faut prendre une grande euh *une grande horizontale* [...]
I_{18} et tu *en* prends *une deuxième*
I_{19} *une petite* [...]
I_{21} et tu *en* prends *une autre petite* [...]

[90] Lorsqu'il s'agit de propriétés définitoires (« avoir quatre côtés » pour un objet de type *CARRÉ*), celles-ci sont stockées dans les représentations mentales génériques (cf. 7.2.3).
[91] Le contexte visuel du dialogue est celui présenté dans la section 1.3.2, page 19.

I_{24} et tu prends *une autre petite verticale*
M_6 *une autre petite verticale* ?
I_{25} euh *horizontale* pardon

Avant toute interprétation, nous supposons la structure contextuelle de la Figure 53a (page 173) : le domaine @P représente la palette des figures disponibles. Celui-ci est partitionné selon le type des figures, elles-mêmes formant des domaines partitionnés selon la taille des figures. Le domaine @H représente ainsi le domaine des objets de type *HORIZONTALE* et @GH celui des *GRANDES HORIZONTALES*. Il est à noter que nous traitons les objets de la palette en tant que domaines génériques, c'est-à-dire en tant que domaines représentant une classe dont les objets extraits seront des instances. Cela est justifié par la spécificité de la tâche, mettant à disposition, dans la palette, des types d'objets, plutôt que des objets directement manipulables. Un clic sur un objet de la palette mène en effet à la création d'une nouvelle figure relevant de ce type.

Parallèlement à la représentation du contexte fourni par la palette, la structure domaniale de la scène courante est celle de la Figure 59a (page 184). Dans la mesure où cette structure n'intervient pas dans l'identification des objets qui nous intéressent ici, nous la laisserons de côté pour alléger les représentations graphiques de la modélisation.

L'expression *une grande horizontale* (I_{15}) mène à la construction d'un domaine sous-spécifié tel que dans la Figure 52a (page 173). Conformément au principe des indéfinis, un domaine compatible est alors recherché, d'abord parmi les domaines spécifiques (ceux de la scène en construction), puis parmi les domaines génériques (ceux fournis par la palette). Aucun domaine spécifique n'étant compatible, ce domaine sous-spécifié sélectionnera le domaine générique @GH, remplissant tous les critères de compatibilité. Dans ce domaine, l'indéfini crée une nouvelle partition, reposant sur l'événement dénoté par le prédicat *prendre*. Sur la base de cet événement et en particulier ses propriétés spatio-temporelles, un élément de type *GRANDE HORIZONTALE* est aléatoirement extrait de ce domaine. Il est opposé aux autres par son premier rang dans l'ordre temporel des extractions, puis focalisé en tant que dernier référent activé (Figure 53b).

L'indéfini suivant est *en... une deuxième*. Selon les principes valables pour les indéfinis, pour les groupes nominaux sans nom ainsi que pour la sémantique particulière de *autre*, le domaine sous-spécifié correspondant à cette expression est celui de la Figure 52b. Il ne contraint pas le type de ses éléments (en dehors d'un type compatible avec une entrée lexicale de genre féminin), mais présuppose l'existence d'une partition amorcée. Dans le contexte de la Figure 53b, ce domaine est compatible avec le dernier domaine activé, c'est-à-dire @GH. De ce domaine, l'expression pourrait alors extraire un deuxième élément ce qui mènerait à l'identification d'une autre grande barre.

Or, ce n'est pas ce que vise le locuteur. Par conséquent, il reformule l'expression en *une petite*. Cette expression conduit à la construction d'un nouveau domaine sous-spécifié tel que dans la Figure 52c, précisant une propriété particulière des éléments du domaine, sans en préciser le type. Le dernier domaine contextuel activé @GH (celui des grandes barres) n'est alors plus compatible avec ce domaine sous-spécifié. La recherche se poursuit donc dans des domaines moins actifs, en commençant par @H. Celui-ci donne effectivement accès, à travers sa partition sur le critère de différenciation *TAILLE*, à un domaine d'éléments ayant la propriété d'être *PETIT*. C'est de ce domaine @PH qu'un élément sera extrait, de la même façon que pour l'expression précédente. La structure contextuelle résultante est celle de la Figure 53c.

L'expression indéfinie suivante est *en... une autre petite*. A partir de cette expression est construit le domaine sous-spécifié de la Figure 52d. Celui-ci ne contraint pas le type des éléments du domaine d'interprétation, leur impose une propriété particulière concernant leur taille et présuppose l'existence

d'une partition amorcée. Le dernier domaine activé @PH remplit tous ces critères. Du complément de sa partition, créée à l'occasion de l'extraction du référent précédent, est alors extrait un deuxième élément qui sera référent de l'expression courante (Figure 53d).

En (I_{24}), le locuteur emploie l'expression *une autre petite verticale*. Cette expression mène à un domaine sous-spécifié tel que dans la Figure 52e. L'expression demande, pour s'interpréter, un domaine d'éléments de type *VERTICALE*, ayant la propriété d'être *PETIT*. Par ailleurs, comme pour les expressions précédentes, *autre* impose l'existence d'une partition amorcée. Ce domaine sous-spécifié n'est instanciable avec aucun des domaines disponibles. Même si un domaine générique de type *PETITE VERTICALE* est accessible à travers la palette, celui-ci n'est pas partitionné et ne convient donc pas. L'expression n'est donc pas interprétable.

Notre modélisation est validée par le comportement de l'interlocuteur (M_6) qui demande confirmation. Suite à cette intervention, le locuteur corrige en effet son expression en modifiant le type des éléments du domaine. Le nouveau domaine sous-spécifié est alors celui de la Figure 52f. Contrairement au domaine précédent, ce domaine est compatible avec le domaine actif de la représentation contextuelle (@PH). Celui-ci est de nouveau sollicité pour l'extraction d'un troisième élément, toujours du complément de la partition amorcée (Figure 53e).

En résumé, le traitement de cet extrait montre que notre modèle est capable de traiter, dans un cadre unifié, un certain nombre de phénomènes ayant fait l'objet d'heuristiques spécifiques ou étant simplement oubliés dans d'autres modélisations : cela concerne en particulier les expressions elliptiques, combinées ou non à des expressions d'altérité et ordinaux. Ces expressions sont par ailleurs très fréquentes dans le corpus examiné et donc probablement dans le type de dialogues qui nous intéresse. Enfin, il est intéressant de noter que nous avons pu prédire correctement les deux expressions problématiques (I_{18} et I_{24}) pour lesquelles le locuteur a été obligé de fournir des reformulations.

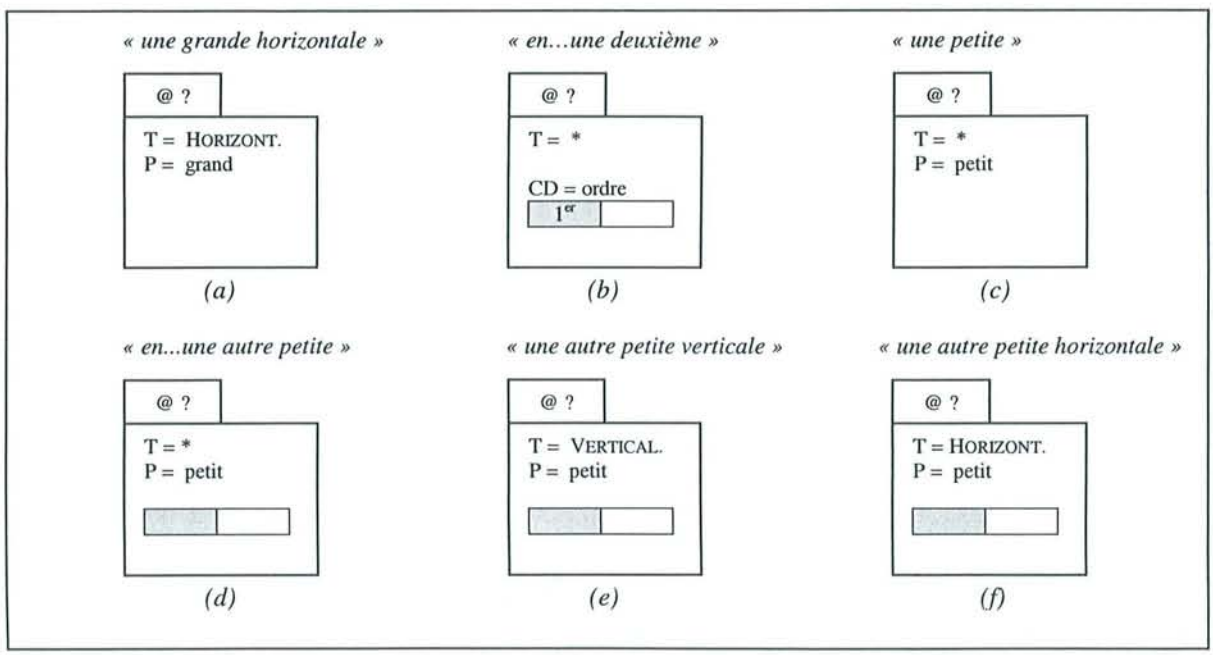


Figure 52 : Domaines de référence sous-spécifié de l'exemple (167)

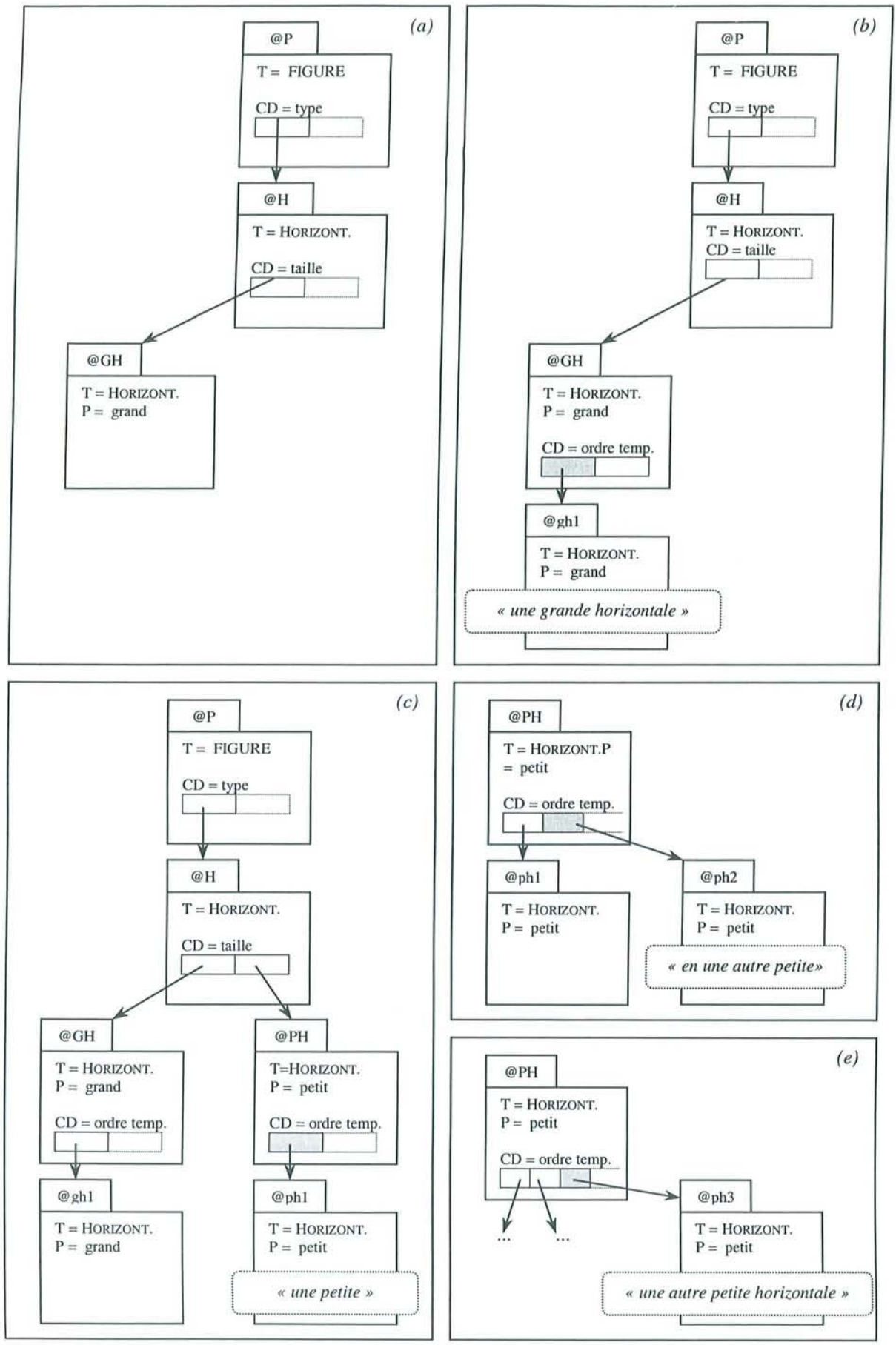


Figure 53 : Interprétation de l'exemple (167)

9.3 Les expressions définies : prédictions et exemple de traitement

9.3.1 Prédictions

- *Les descriptions définies nécessitent un domaine de référence partitionné.*

Une première prédiction sur les descriptions définies concerne la case vide au croisement de la ligne 5 et de la colonne 3 du Tableau 20 (page 162). L'absence d'un domaine résultant vient du fait que les descriptions définies sont incompatibles avec la structure domaniale de la ligne 5. Elles nécessitent en effet un domaine partitionné pour s'interpréter. Cette prédiction reflète l'intuition d'une acceptabilité décroissante des exemples (168) à (171) :

(168) Prends un triangle et un carré. Colorie *le triangle*.

(169) Prends un triangle. Colorie *le triangle*.

(170) Prends un triangle. Colorie *le triangle rouge*.

(171) Prends deux triangles. Colorie *le triangle*.

L'exemple (168) sert de base pour une comparaison. Son interprétation (relevant du croisement de la ligne 6 et de la colonne 3 du Tableau 20) n'est pas problématique : *le triangle* peut effectivement être extrait du domaine partitionné introduit par le syntagme coordonné *un triangle et un carré* :

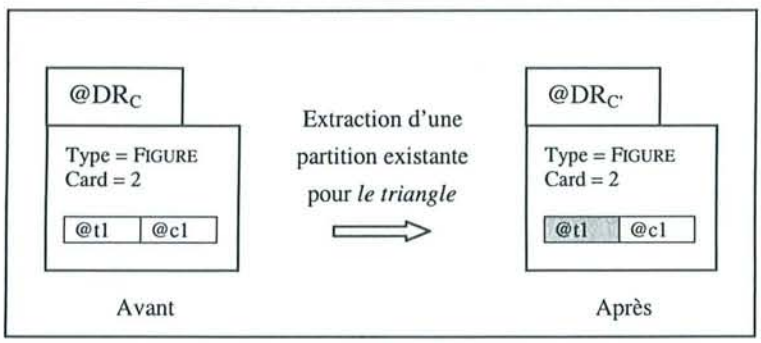


Figure 54 : Extraction de « le triangle » d'un domaine partitionné – exemple (168)

Contrairement à (168), le défini *le triangle* de l'exemple (169), s'il ne pose pas de problèmes d'interprétation, semble en tout cas moins approprié qu'un pronom dans le même contexte. Cette impression de sous-optimalité vient du fait que l'emploi d'un défini présuppose l'existence d'un domaine partitionné à l'intérieur duquel un élément puisse s'opposer, en vertu d'une propriété particulière, à d'autres éléments. Or, un tel domaine n'est pas disponible dans le contexte de l'exemple (169).

Le processus d'interprétation est alors celui de la Figure 55. Sont disponibles dans le contexte, avant l'interprétation du second énoncé, deux représentations : le domaine générique de la classe *TRIANGLE* ($@DR_{C1}$) et la représentation pour *un triangle* ($@DR_{C2}$), extrait de cette classe. Aucune des deux représentations ne remplit les contraintes du domaine sous-spécifié correspondant au défini *le triangle* ($@DR_{ER}$). Ce domaine est effectivement un domaine d'éléments de type *FIGURE* avec une partition selon le type de figures, dont une de type *TRIANGLE*. En revanche, une des deux représentations ($@DR_{C2}$) remplit au moins les critères de l'élément à isoler : elle est de cardinalité 1 et de type *TRIANGLE*. La stratégie de réparation consiste alors à construire, autour de cet élément, un domaine « virtuel » ($@DR_{C2'}$) de type *FIGURE*, l'insérant dans une partition supposée opposer un triangle à d'autres types de figures. Le sentiment de sous-optimalité de l'exemple (169) vient alors du

fait que le complément de cette partition ne peut pas être saturé par des éléments disponibles du contexte.

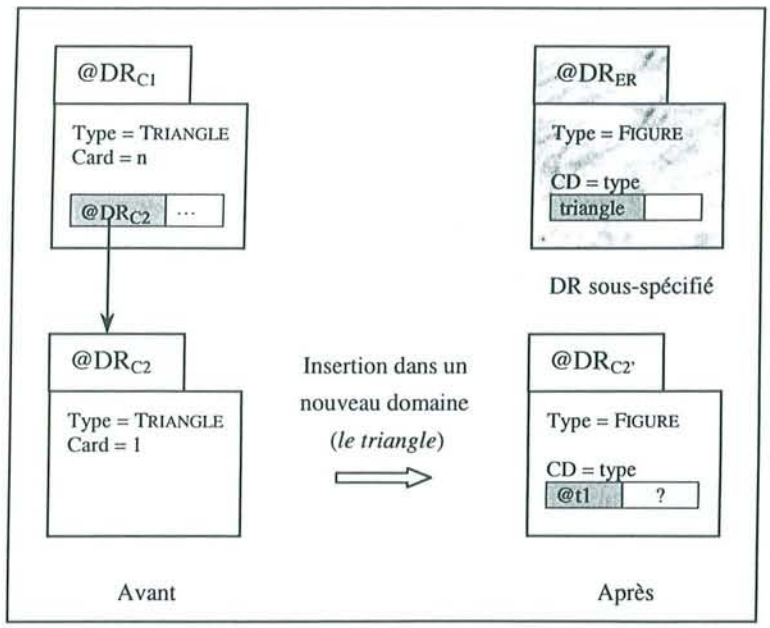


Figure 55 : Interprétation de l'exemple (169)

Les exemples (170) et (171) ne sont plus interprétables dans les conditions contextuelles données par le seul énoncé introductif. Cet échec s'explique par rapport à la stratégie de réparation mise en œuvre pour l'exemple précédent (169) : la stratégie présentée repose en effet sur la possibilité d'identification, dans le contexte, d'un référent convenable ($@DR_{C2}$), fût-ce à l'intérieur d'un domaine inapproprié. Or, cela repose sur une compatibilité des propriétés, dont le type et la cardinalité. Ce sont précisément ces deux caractéristiques qui posent problème en (170) et (171). Pour l'exemple (170), il s'agit de la propriété *COULEUR* : rien ne permet en effet d'inférer que le triangle isolé auparavant a la propriété d'être rouge. Si l'on voulait modéliser l'acceptation d'une lecture coréférentielle en (170), il faudrait admettre cette inférence[92]. Pour l'exemple (171), il s'agit de la cardinalité : la cardinalité de la représentation disponible dans le contexte est de 2, ce qui la rend incompatible avec la singularité du référent présumé. Ici, plus aucune stratégie de réparation ni d'inférence ne peut être mise en œuvre.

- *L'association n'est pas un mode d'interprétation exceptionnel.*

Une deuxième prédiction sur les expressions définies concerne le mode de référence en anaphore associative. Nous avons effectivement vu que les anaphores associatives ont tendance à être traitées comme des cas exceptionnels dans les modélisations courantes (cf. Bosch et Geurts, 1990 ; Bos et al., 1995). Nous avons également mentionné que ce point de vue est contraire aux observations de M. Poesio et R. Vieira (1998), qui ont relevé un usage fréquent des définis non anaphoriques dans un corpus journalistique de l'anglais.

Dans le cadre de la modélisation proposée, nous considérons que l'usage associatif des descriptions définies n'est pas un cas spécifique. Il relève de façon parfaitement régulière du principe d'interprétation propre aux définis et nous n'avons pas besoin d'introduire de mécanisme à part pour traiter cet usage. En effet, un défini suppose toujours un domaine partitionné d'où il peut extraire son

[92] Le même type d'inférence serait nécessaire pour admettre des références au défini à l'aide d'un hyponyme : « Prends une figure. Mets le triangle à gauche. »

référent sur des critères distinctifs. Ce qui distingue une interprétation en anaphore fidèle d'une interprétation en anaphore associative n'est pas le mode de construction du domaine de référence sous-spécifié, ni l'opération de restructuration. C'est l'origine de la partition du domaine d'interprétation. Nous avons effectivement prévu que la partition puisse être donnée en extension par le contexte linguistique ou visuel, mais aussi par des connaissances de nature encyclopédiques, auquel cas la partition est héritée d'une représentation mentale générique. C'est cette dernière possibilité qui est mise en jeu lors d'une interprétation en mode associatif.

Une deuxième propriété intéressante par rapport aux interprétations associatives est le fait que notre modélisation arrive à refléter correctement l'intuition selon laquelle l'anaphore associative « *roule* » dans un sens tout – partie (Kleiber et al., 1994). Un défini suppose en effet la représentation de son domaine d'interprétation (et non pas la représentation d'une entité qui est une composante de son référent) et c'est seulement à l'intérieur de ce domaine qu'il est capable d'isoler son référent.

(172) *Le pied* est abîmé, mais *la chaise* est toujours solide. (A. Azoulay, repris par Kleiber et al., 1994)

Dans les exemples d'anaphores associatives cataphoriques tels que (172), l'interprétation de l'expression *le pied* est effectivement soumise à l'activation d'un domaine donnant accès à un élément de type *PIED*. En l'absence d'éléments contextuels permettant d'instancier ce domaine, l'interprétation du défini reste sous-spécifiée, dans la mesure où le locuteur est en effet incapable de dire de quel *PIED* il s'agit. En revanche, l'interprétation de l'expression *la chaise* est indépendante de celle de *le pied.* Elle demande un domaine permettant d'isoler un objet de type *CHAISE* d'autres objet du même domaine et ce n'est qu'ensuite que le référent de la représentation sous-spécifiée pour *le pied* pourra être identifié de façon définitive, en se fusionnant avec un des pieds de la chaise en question. Ainsi, le « bon sens du roulement de l'anaphore associative » est maintenu.

- *Le défini cherche à stabiliser une partition amorcée.*

Une autre série de prédictions sur les définis est liée au rôle que nous leur attribuons dans la restructuration du contexte. Nous avons effectivement considéré que l'opération de restructuration consiste, pour un défini, en l'extraction et la focalisation du référent dans une partition donnée. Pour aller un peu plus loin, nous supposons même que l'emploi d'un défini est optimal lorsque celui-ci stabilise la partition. Par stabilisation d'une partition, nous entendons le parcours des items d'une même partition active[93]. Ce parcours stabilisateur se traduit dans le cadre de notre modélisation par une opération de changement d'élément focal à l'intérieur d'une même partition.

En admettant cette hypothèse, on pourrait s'attendre à ce qu'un défini puisse être employé de façon sous-optimale dans deux cas :

> lorsqu'il s'interprète dans la partition active, mais sans la parcourir, c'est-à-dire lorsqu'il extrait et focalise un élément qui était déjà focalisé ;
>
> lorsqu'il ne s'interprète pas dans la partition active, c'est-à-dire lorsqu'il extrait et focalise un élément d'une partition non active.

Par la suite, nous détaillerons ces deux prédictions sur des exemples.

La première prédiction s'applique à des exemples tels que (173) et plus précisément l'emploi de l'expression *le grand triangle* dans le troisième énoncé de (a).

[93] Cette idée de stabilisation peut peut-être être rapprochée de l'hypothèse de G. Kleiber, selon laquelle le défini prolonge la cadre d'évaluation posé par l'énoncé introducteur (cf. le « paradoxe de la reprise immédiate », chapitre 1).

(173) a. Prends un petit et un grand triangle. Colorie le grand triangle. Agrandis *le grand triangle*.
b. Prends un petit et un grand triangle. Colorie le grand triangle. Agrandis *le petit triangle*.
c. Prends un petit et un grand triangle. Colorie le grand triangle. Agrandis *le*.

La structure contextuelle à l'issue des deux premiers énoncés est celle de la partie gauche de la Figure 56. Elle fournit un domaine compatible avec le domaine sous-spécifié de l'expression *le grand triangle* (demandant un domaine de type *TRIANGLE* et permettant d'un isoler un sur le critère *TAILLE*). Mais l'opération de restructuration consiste alors en l'extraction et la focalisation de l'élément @grand, ce qui nous ramène exactement à la structure initiale du domaine. Le défini a donc bien agi sur la partition active (ici, la seule disponible du domaine), mais il n'en a pas changé la structure focale. Contrairement à (a), le défini *le petit triangle* en (b) change la structure focale de cette partition : il focalise l'élément @petit. Nous prédisons alors ce sentiment d'étrangeté face à la variante (a) par un emploi qui ne correspond pas à l'aspiration « stabilisatrice » du défini. Ce sentiment est encore renforcé par le fait que le pronom personnel en (c) provoque le même effet interprétatif que (a) et ceci, conformément au principe des pronoms, sans aucune opération de restructuration.

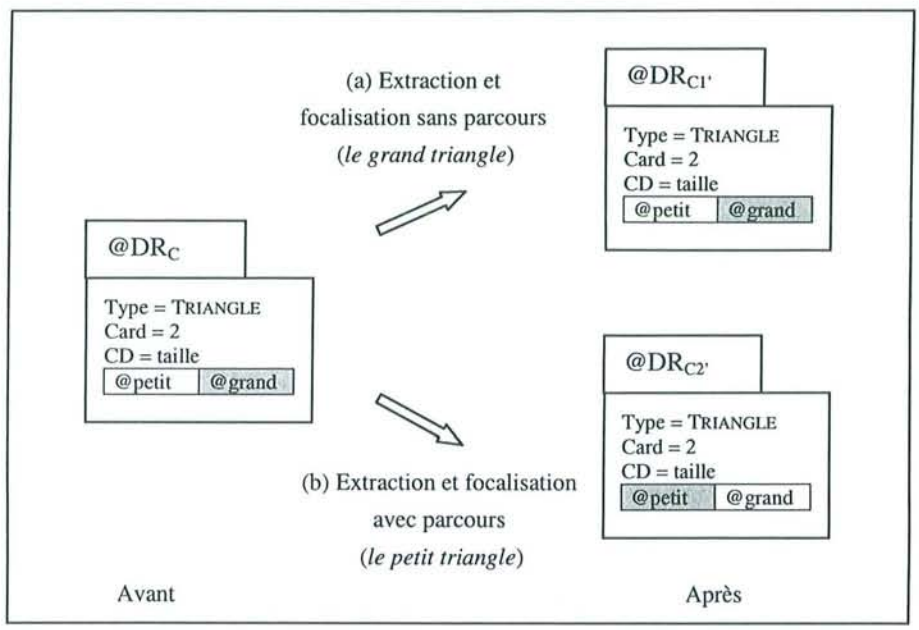
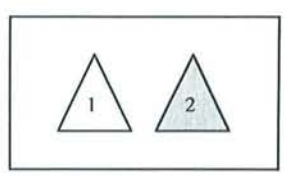


Figure 56 : Interprétation de l'exemple (173)

La deuxième prédiction est illustrée à travers l'exemple (174) :

(174) a. Agrandis le triangle blanc.
Supprime *le triangle de droite*.

b. Agrandis le triangle blanc.
Supprime *le triangle noir*.



Ici, *le triangle de droite* en (a) demande à s'interpréter dans un domaine de type *TRIANGLE*, dont un élément peut être isolé sur un critère de différenciation qui est la *POSITION HORIZONTALE*. Or, le domaine contextuel est celui de gauche de la Figure 57. Il s'agit effectivement d'un domaine de type *TRIANGLE*, mais dont la partition active, c'est-à-dire celle qui a été impliquée dans l'opération d'extraction précédente (*le triangle blanc*), est une partition selon la propriété *COULEUR*. Le domaine convient malgré cela, dans la mesure où une partition sur la position horizontale est instanciable, à partir des connaissances visuelles. L'expression *le triangle de droite* extrait donc son référent d'une partition nouvellement instanciée, mais cette procédure est plus coûteuse que l'interprétation du défini *le triangle noir* dans les mêmes circonstances (b). Nous expliquons alors l'optimalité moindre de (a)

face à (b) par le fait qu'en (a), le défini ne stabilise pas la partition active, mais en active une autre, fait qui est contraire à sa vocation en tant que « stabilisateur » de partition.

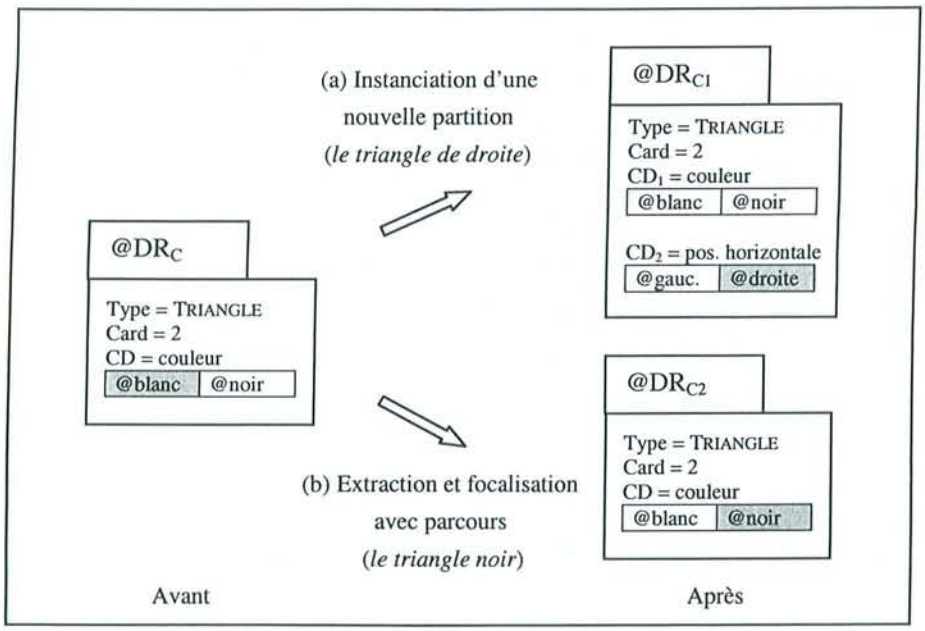


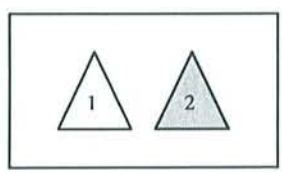
Figure 57 : Interprétation de l'exemple (174)

Enfin, un cas extrême combine les deux prédictions précédentes : lorsqu'un défini identifie, dans une nouvelle partition, un élément focalisé précédemment dans une autre partition, son emploi semble très difficile : c'est le cas de la variante (a) de l'exemple (175) :

(175) a. Déplace le triangle blanc.
Supprime *le triangle de gauche.*

b. Déplace le triangle blanc.
Supprime *le.*

c. Déplace le triangle blanc.
Supprime *ce triangle de gauche.*



En (a), *le triangle de gauche* crée non seulement un nouvelle partition au lieu de parcourir la partition active, mais il ré-identifie aussi l'élément focalisé au préalable (*le triangle blanc*). Dans ce cas, l'emploi du défini ne se justifie plus par rapport au pronom (b), ni même par rapport au démonstratif, mieux adapté pour changer de point de vue sur un même référent (c). Concernant la variante discursive de ce type d'enchaînements, F. Corblin (1987) exclut même la possibilité de coréférence dans des exemples tels que (176) :

(176) *Cette rose rouge* me gêne. Je vais jeter *la rose fanée.* (Corblin, 1987)

- *Les interprétations génériques des définis découlent de notre modélisation.*

Enfin, une dernière prédiction concerne l'emploi générique des définis : celui-ci est le plus souvent négligé dans les modélisations courantes, bien que considéré par F. Corlin (1987) comme mode d'interprétation fondamental. Notre modélisation arrive en effet à prédire de tels emplois par le fait qu'en absence d'un domaine contextuel convenable, l'espèce (la représentation mentale générique) correspondant au type de l'expression *le N* peut toujours être isolée dans la hiérarchie des types. La seule condition est que la prédication portant sur *N* remplit les critères exposés dans la section 2.3.4,

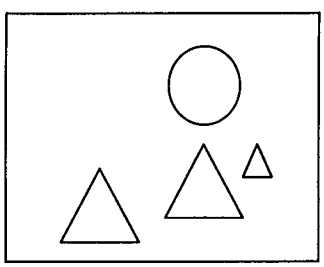
dont la vérification collective du prédicat sur l'espèce en question C'est le cas de *les égyptiens* de l'exemple (177) :

(177) [à propos des pyramides...]
M_1 tu comprends pourquoi *les égyptiens* en ont bavé pour les faire (C12Egypte)

L'absence d'un ensemble d'égyptiens particulier et la vérification collective du prédicat sur l'espèce en question mène ici à une interprétation générique. Pour l'instant, les procédures de vérification de la validité (collective ou distributive) d'un prédicat sur l'ensemble d'une classe ne sont pas implémentés, mais elles peuvent être ajoutées sans changer les principes de la modélisation : il suffira d'augmenter les représentations génériques par ce type de connaissances, en séparant les informations valides collectivement des informations valides distributivement sur une espèce.

9.3.2 Exemple de traitement

L'illustration de la modélisation du traitement des descriptions définies se fera sur le même extrait de dialogue : cette fois-ci, nous nous concentrerons sur le traitement d'une série de références aux objets de la scène et en particulier aux trois triangles (deux grands et un petit), disposés comme dans la reproduction de la scène ci-dessous :



(178) I_{14} et puis maintenant faudrait faire la ligne d'horizon
I_{15} il faut prendre une grande euh une grande horizontale
I_{16} et la placer à *la pointe des triangles des deux grands triangles*
M_3 ah mais elle va pas être horizontale alors
M_4 parce qu'il y *en* a *un* qui est plus haut que *l'autre*
[geste : M pose une horizontale]
I_{17} voilà comme ça
I_{18} et tu en prends une deuxième
I_{19} une petite
I_{20} tu la places à gauche de *la euh pyramide de gauche*
[geste : M pose une horizontale]
I_{21} et t'en prends une autre petite
M_5 oui
I_{22} et tu la places à droite de *la pyramide de droite*
[geste : M pose une horizontale]
I_{23} voilà comme ça
I_{24} et tu prends une autre petite verticale
M_6 une autre petite verticale ?
I_{25} euh horizontale pardon
M_7 ah
I_{26} et puis tu la places dans la même lignée
I_{27} à droite de *la petite pyramide*
[geste : M pose une horizontale]

Nous partons de la représentation du contexte telle que dans la Figure 59a (excepté la focalisation de @gt) : la scène courante @s se compose d'un ensemble de triangles @t et d'un grand rond @gr. Les triangles sont regroupés par taille : @pt1 représente un petit triangle, @gt représente un ensemble de deux grands triangles. Cet ensemble est lui-même partitionné selon la position horizontale de ses éléments. Nous distinguons le triangle de gauche @gt1 de celui de droite @gt2.

La première expression définie à traiter est *la pointe des deux triangles* (I_{16}). Il s'agit d'une expression de forme *le N1 de N2*, forme dont l'interprétation se fait en deux étapes : dans un premier temps, le référent de *N2* doit être identifié. Dans une deuxième étape, celui-ci fournit le domaine d'interprétation pour *N1*.

N2, *les deux triangles*, mène à la construction du domaine sous-spécifié de la Figure 58a. Celui-ci contraint le type des éléments du domaine d'interprétation à un sur-type de *TRIANGLE*, c'est-à-dire *FIGURE* et impose l'existence d'une partition permettant d'isoler un ensemble d'éléments de type *TRIANGLE* dont la cardinalité est deux. Étant donné qu'un tel domaine n'est pas disponible, le locuteur précise qu'il s'agit de *deux grands triangles*. Cela modifie la structure du domaine sous-spécifié qui contient maintenant un ensemble d'éléments de type *TRIANGLE* partitionné selon la taille (Figure 58b). Ce domaine est compatible avec le domaine contextuel @t, dont l'élément @gt est identifié et focalisé comme référent de l'expression *les deux grands triangles* (Figure 59a).

Ce référent sert lui-même de domaine de référence pour l'interprétation de *N1*, à savoir *la pointe*. Cela signifie que ce domaine doit posséder une partition donnant accès à un élément de type pointe (Figure 58c). Un telle partition n'est pas disponible dans le domaine spécifique @gt. En revanche, ce domaine possède un pointeur vers son domaine générique, c'est-à-dire une représentation pour les entités de type *TRIANGLE*. Dans cette représentation générique existe une partition, donnant accès aux composantes d'un triangle, donc entre autres à un élément de type *POINTE*. Cependant, cette partition, propre à un élément de type *TRIANGLE*, ne peut pas être instanciée directement dans un ensemble de deux triangles comme @gt. Elle est donc recopiée séparément dans chacun des éléments de cet ensemble, c'est-à-dire dans @gt1 et @gt2. En revanche, les éléments identifiés dans ces partitions (les pointes) sont à leur tour regroupés et ceci dans un groupe sans partition, afin de bloquer l'accès individuel. C'est ce groupe @pe qui est le référent de l'expression *la pointe des deux grands triangles*. Le domaine actif est toujours celui des deux grands triangles @gt (Figure 59b).

L'expression suivante est *en...un*. Il s'agit d'un indéfini que nous avons choisi de traiter ici en raison de la relation étroite qu'il entretient avec son correspondant défini *l'autre*. Cette expression conduit au calcul d'un domaine sous-spécifié qui, en dehors du genre et du nombre, ne contraint pas le type des éléments et ne demande pas de partition préalable (Figure 58d). Le domaine actif, @gt, regroupe des éléments ayant été classifiés par un désignation masculine (*triangle*). Par ailleurs, il est possible d'y créer une nouvelle partition selon la position verticale des éléments du domaine visuel. Conformément au principe des indéfinis, l'expression indéfinie extrait un élément de cette partition et le caractérise par des informations véhiculées par le prédicat de l'énoncé (*être plus haut que x*), comme dans la Figure 59c[94]. Il est à noter que cette extraction se fait bien aléatoirement : s'il est probable qu'en présence de la scène visuelle, l'interlocuteur est capable d'identifier le triangle en question, cette identification ne conditionne en aucun cas la bonne compréhension de l'énoncé.

L'expression *l'autre* s'interprète comme suit : elle introduit le domaine sous-spécifié de la Figure 58e. Celui-ci n'impose pas de contrainte sur le type des éléments, mais demande l'existence d'une partition amorcée. Cette structure s'accorde avec celle du domaine actif qui est toujours @gt. De la partition active de @gt, *l'autre* extrait alors le complément non focalisé, c'est-à-dire le triangle restant (Figure 59d).

L'expression suivante est *la pyramide de gauche*. Son domaine sous-spécifié demande des éléments de type *TRIANGLE*[95] se différenciant selon leur position horizontale (Figure 58f). Là encore, le domaine actif est compatible : il fournit un ensemble de deux triangles dont un peut être identifié en raison de sa position horizontale à gauche. Le référent de l'expression est donc @gt1 et la partition active est celle portant sur la position horizontale des éléments du domaine (Figure 59e). Il est à noter que nous avons ici appliqué notre prédiction sur le rôle stabilisateur que nous attribuons au défini. En faisant abstraction des critères d'activation domaniale, l'expression à interpréter était en effet

[94] Pour des raison de place, nous représentons, en cas de plusieurs partitions d'un même domaine, que la partition active. En revanche, nous indiquons tous les critères de différenciation, en surlignant celui correspondant à la partition active.
[95] Dans le cadre ce dialogue, nous supposons que les entrées lexicales *triangle* et *pyramide* sont tous les deux disponibles pour référer à des éléments de type *TRIANGLE*.

ambiguë : elle pouvait également s'interpréter dans le domaine de l'ensemble des triangles, @t. Or, une interprétation dans ce domaine aurait non seulement mené à un changement de domaine actif, mais aussi à la création d'une deuxième partition du domaine @t. En plus de la partition existante selon la taille des éléments, il aurait fallu gérer une partition selon la position horizontale. Nous verrons dans la suite que la modélisation selon laquelle l'expression courante s'interprète bien dans le domaine @gt correspond au comportement des protagonistes du dialogue.

L'interprétation de *la pyramide de droite* conduit à domaine sous-spécifié qui, parallèlement à celui de l'expression précédente, impose un domaine de type *TRIANGLE*, pouvant être partitionné selon la position horizontale des éléments (Figure 58g). Là encore, l'interprétation en termes de domaines activés s'avère indispensable : hors de tout contexte discursif, cette expression désigne en effet la petit pyramide à droite (@pt1). Or, ce n'est pas ce qu'envisage le locuteur ni ce que comprend l'interlocuteur. Le domaine d'interprétation est bien celui activé précédemment (@gt) et le défini exerce de plein droit son rôle de stabilisateur de partition. Il parcourt en effet la partition active (position horizontale) du domaine actif (@gt) et le référent identifié est bien la plus à droite des deux grandes pyramides (@gt2). Le contexte courant est alors celui de la Figure 59f.

Enfin, la référence à *la petite pyramide* confirme les interprétations précédentes. Un accès référentiel à @pt1 passe bien la partition existante du domaine des triangles selon la taille, plutôt que par une nouvelle partition selon la position. Le domaine sous-spécifié de cette expression est celui de la Figure 58h. Ce domaine, n'étant pas compatible avec le domaine actif, déclenche une recherche en amont. Le domaine @t convient et là encore, l'interprétation du défini s'apparente à un parcours de partition, dans la mesure où le référent de l'expression est fourni par le complément de la partition pointant sur le domaine précédent. (Figure 59g).

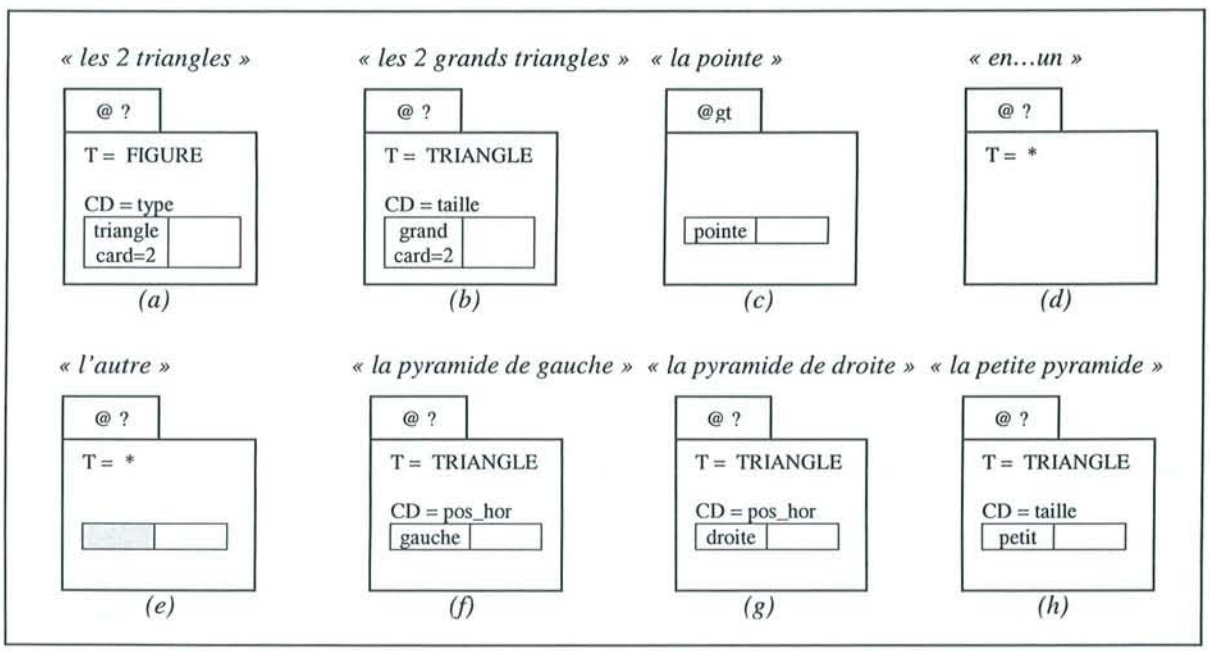
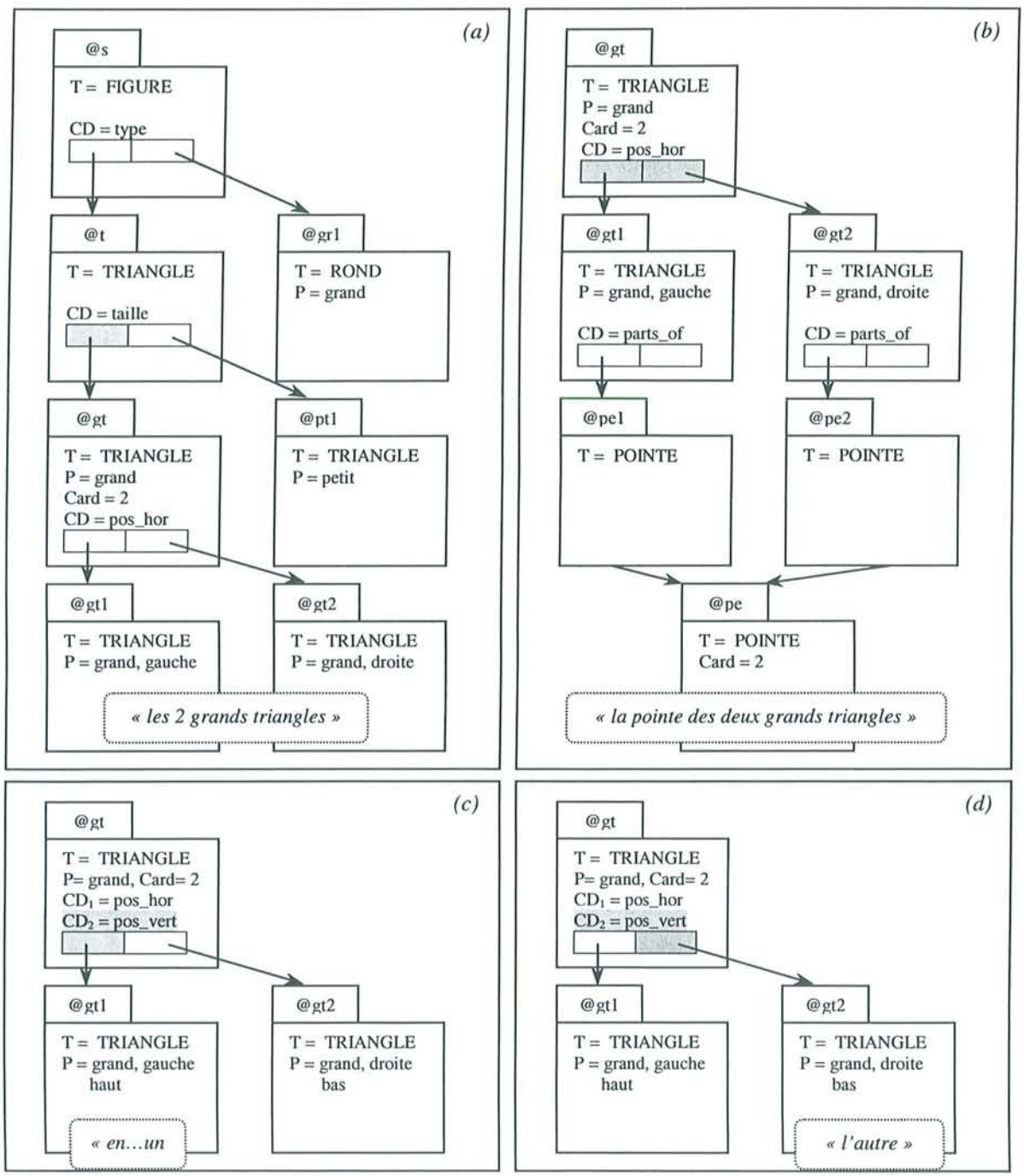


Figure 58 : Domaines sous-spécifié pour l'interprétation de l'exemple (178)

En résumé de cette série d'interprétation d'expressions définies, nous retiendrons deux choses : d'une part, le principe fondamental tel que nous l'avons modélisé – recherche d'un domaine partitionné et extraction d'un des éléments de la partition – permet de traiter de façon unifiée tous les emplois du défini. L'interprétation présentée ci-dessus l'a montré pour des emplois anaphoriques et associatifs et le traitement des emplois génériques (un seul cas dans le corpus Ozkan) a été discuté dans la section précédente. D'autre part, la prédiction selon laquelle un défini s'emploie de façon

optimale lorsqu'il stabilise une partition, a été confirmée : l'évolution des structures focales des représentations de la Figure 59 permet de s'en apercevoir. Bien plus, sans cette caractéristique, l'interprétation correcte de l'expression *la pyramide de droite* en I_{22} n'aurait pas été possible. Enfin, les réflexions autour de cette prédiction pourraient se prolonger par l'hypothèse que le cas contraire (non parcours d'une partition) a tendance à être marqué : cela nous semble être précisément la fonction de *même* : l'interprétation de *la même lignée* (I_{26}) que nous avons laissée de côté pour des raisons pratiques[96], semble aller dans ce sens.



[96] plus précisément, pour éviter les représentations successives de l'alignement des barres...

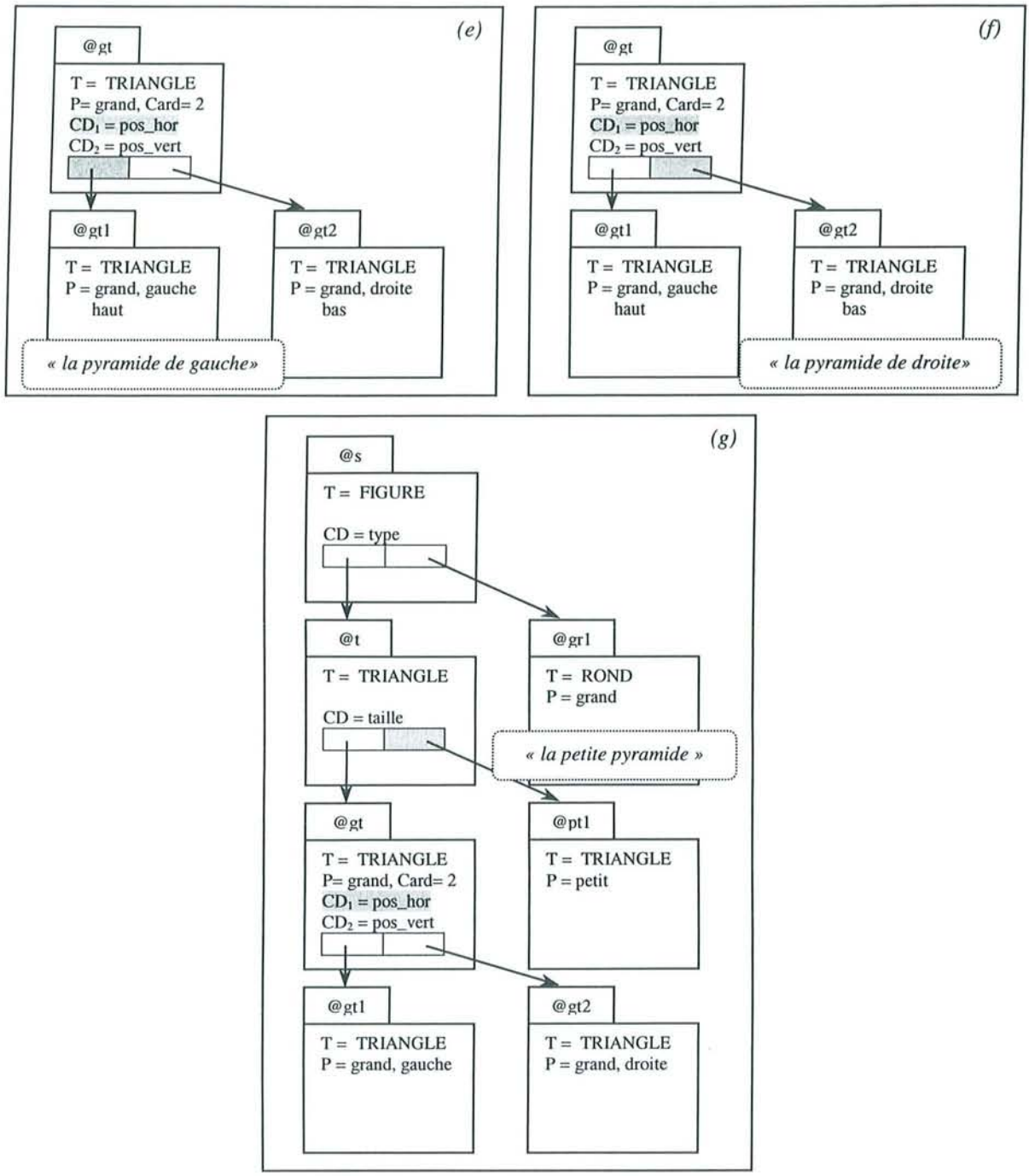


Figure 59 : Interprétation de l'exemple (178)

9.4 Les démonstratifs : prédictions et traitement d'un exemple

9.4.1 Prédictions pour les expressions démonstratives

- *Un démonstratif s'interprète dans un domaine focalisé.*

La première prédiction concernant les démonstratifs porte sur la structure des domaines de référence compatibles : il apparaît dans le Tableau 20 (page 162) qu'un démonstratif ne peut s'interpréter dans un domaine sans élément focalisé. Cette prédiction qui traduit le principe d'identification externe du référent, se trouve effectivement confirmée dans des exemples comme (179) ou (180).

(179) * Prends deux triangles. Colorie *ce triangle.*

(180) ? Prends un triangle et un carré. Colorie *ce triangle.*

Dans un domaine sans partition (179), l'interprétation de *ce triangle* échoue. Dans un domaine partitionné, mais sans élément focal (180), l'interprétation semble très difficile. Nous montrerons ci-dessous que la différence éventuelle entre (179) et (180) peut s'expliquer par une cohésion moindre et donc une accessibilité individuelle plus grande pour les éléments du domaine en (180).

Le domaine focalisé nécessaire à l'interprétation des définis peut être issu de l'historique discursif ou perceptif. L'élément focalisé peut en effet avoir été mis en relief par une mention préalable (181) ou par un geste de désignation concomitant (182) :

(181) I_1 alors tu vas prendre le gros rond
I_2 voilà et à gauche de *ce rond* tu vas prendre une petite barre (C8Homme)

(182) I_1 voilà donc *ce triangle-là* [+geste], il faut le mettre un peu plus à...
M_1 tu veux l'ajuster ? (C7Egypte)

- *Le démonstratif apporte du nouveau par la construction d'un nouveau domaine.*

Une deuxième prédiction concerne les conditions d'usage d'un démonstratif par rapport à son effet interprétatif : nous avons retenu des travaux de linguistique qu'un démonstratif doit « apporter du nouveau » (Kleiber, 1986, 1988, 1994 ; Corblin, 1995 ; de Mulder, 1998), que cela soit par une reclassification, un changement de thème ou l'instauration d'un autre type de rupture discursive.

Cette caractéristique du démonstratif a été modélisée par la construction d'un nouveau domaine. Pour un démonstratif *ce N*, nous avons en effet prévu d'insérer le référent, identifié sur des critères externes, dans un domaine d'éléments de type *N*. Cette modélisation permet de définir « l'apport de nouveau » par la différence entre l'ancien et le nouveau domaine du référent : plus ceux-ci se ressemblent, moins l'effet interprétatif du démonstratif se distinguera de celui du pronom et plus le démonstratif semblera employé de façon sous-optimale. Cette modélisation prédit également quand l'emploi d'un démonstratif est approprié :

lorsqu'il y a un changement du type des éléments du domaine ;

lorsqu'il y a un changement de la structure partitionnelle du domaine.

Ces deux cas sont effectivement observables dans le corpus Ozkan. Le premier cas s'accorde parfaitement avec une première mention d'un référent qui est saillant dans la situation de communication (par des caractéristiques visuelles ou un geste de désignation). Dans une telle

situation, le démonstratif permet de réduire l'incertitude inhérente à tout domaine non encore classifié :

(183) M_1 eh Nadine, tu nous fais pas dessiner un dromadaire
I_1 ça va être dur avec *ces formes-là* (C11Egypte)

Dans l'exemple (183), l'ensemble des formes possède en effet une représentation, mais son type est calculé par défaut (par exemple, *FIGURE*). Le démonstratif permet de réduire cette incertitude en insérant le référent dans un nouveau domaine dont le type sera celui de la mention *N*. « L'apport du nouveau » consiste alors en une information supplémentaire quant au type du référent.

Le deuxième cas, le changement d'une structure partitionnelle, est illustré en (184) :

(184) M_1 tu bougeras la pyramide après, c'est pas grave
I_1 voilà donc *ce triangle-là* [+geste], il faut le mettre un peu plus à...
M_2 tu veux l'ajuster ?
I_2 ouais je veux le mettre sur le bord (C7Egypte)

Ici, le domaine actif avant l'emploi du démonstratif est celui de trois triangles-pyramides alignés horizontalement, dont la position exacte est sujette à discussion. L'expression *la pyramide* en M_1 désigne le second triangle, alors que le démonstratif *ce triangle-là*, accompagné d'un geste, en désigne le troisième. Le type du nouveau domaine construit pour le démonstratif reste donc le même[97]. En revanche, la structure focale est différente, puisque l'élément focalisé est maintenant le référent du démonstratif. L'emploi d'un défini (*la pyramide de droite, la troisième pyramide*, ...) aurait eu le même effet interprétatif, mais le locuteur a dû juger le recours à un geste plus efficace.

Les deux emplois précédents étant expliqués de façon satisfaisante, il nous reste à fournir une explication pour l'emploi des démonstratifs dans les exemples suivants, où l'on est en droit de s'interroger sur l'apport informationnel du démonstratif :

(185) I_1 alors tu vas prendre le gros rond
I_2 voilà et à gauche de *ce rond* tu vas prendre une petite barre (C8Homme)

(186) I_1 eh donc le deuxième dessin représente une route
M_1 ouais
I_2 donc faite avec des barres verticales et euh
donc au bord de *cette route*, il y a deux maisons (C9Route)

En ce concerne l'exemple (185), la simple construction d'un nouveau domaine de type *ROND* (Figure 60) ne justifie effectivement pas l'emploi du démonstratif. Le type du référent est connu et de plus, le complément de la partition du nouveau domaine reste vide, puisque aucun autre rond n'est disponible dans le contexte. Le même type de raisonnement s'applique à l'exemple (186).

[97] En admettant que *pyramide* et *triangle* relèvent du même type, comme nous l'avons fait jusqu'ici. Ce changement entre domaine géométrique et domaine figuratif sera rediscuté au paragraphe suivant.

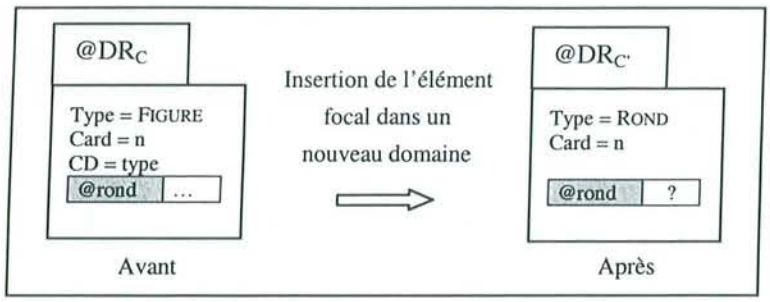


Figure 60 : Interprétation prédite pour l'exemple (185)

L'incapacité de prédire ces emplois vient du fait que pour l'instant, nous insérons systématiquement le référent dans un domaine dont le type correspond à *N* de la mention *ce N*. Pourtant, un examen attentif des données semble indiquer qu'il ne s'agit là que d'une première étape : le type du nouveau domaine ne peut pas toujours être calculé ainsi. Il faut aussi tenir compte d'autres indices linguistiques et même attendre, quelquefois, la suite du discours. Concernant les indices linguistiques, le rôle des prépositions spatiales est important : dans le corpus Ozkan, huit démonstratifs sur dix (dont les exemples (185) et (186)) fonctionnent en tant que site[98] dans un énoncé de positionnement. Or, dans la grammaire cognitive, les arguments spécifiant une préposition spatiale (le site et la cible) sont considérés comme formant un symbole conceptuel complexe, ce qui nous amène à les regrouper. Cela signifie que le référent de ces démonstratifs fera partie d'un nouveau domaine, non pas obligatoirement de type *N*, mais d'un type commun aux objets regroupés. Dans cette perspective, le démonstratif peut alors être considéré comme préparant le terrain à ce groupement : il détache son référent de l'ancien domaine, en créant un domaine provisoire de type *N*. Ce domaine sera alors restructuré à son tour en s'unifiant avec le résultat du groupement (Figure 61).

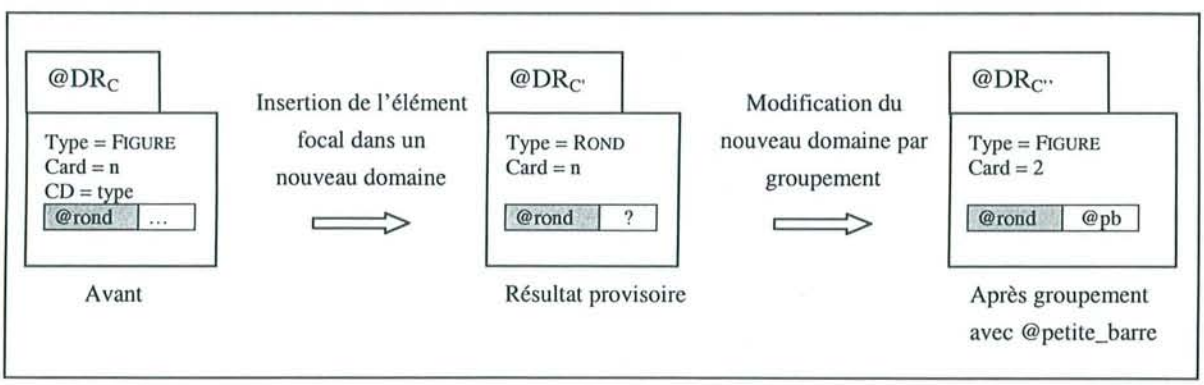


Figure 61 : Interprétation de l'exemple (185) en intégrant l'opération de groupement

En plus du facteur de groupement, d'autres facteurs semblent jouer un rôle pour la justification de l'emploi des démonstratifs dans le corpus Ozkan. Parmi ceux-ci, on trouve des changements entre les désignations en tant qu'objet géométrique ou en tant qu'objet figuratif (*triangle* vs. *pyramide*). Trois des dix démonstratifs du corpus (dont l'exemple (186)) traduisent en effet un changement de point de vue sur le référent. D'une description compositionnelle en termes géométriques (*faite avec des barres verticales*), on (re-)passe à l'objet figuratif (*cette route*).

[98] Le site ou *landmark* d'une préposition spatiale est l'objet de référence par rapport auquel une position est calculée. Cette entité est considérée dans la grammaire cognitive comme moins proéminent que la cible ou le *trajector*, qui représente l'objet à positionner.

- *Le pouvoir reclassificateur d'un démonstratif dépend de la topographie domaniale.*

La troisième prédiction que notre modélisation permet de faire sur les démonstratifs concerne leur pouvoir de reclassification : tel que formulé actuellement, le modèle traduit le point de vue selon lequel la référence virtuelle de l'expression n'intervient pas dans l'identification du référent. Celle-ci est alors disponible pour reclassifier le référent et ceci sans contraintes linguistiques, comme le suppose F. Corblin (1987) à propos d'un exemple tel que :

(187) Deux arbres encadraient l'entrée et *ces sentinelles* dormaient. (Corblin, 1987)

Dans les exemples (188) et (189), cette vision se trouve confirmée :

(188) I_1 alors tu vas prendre le gros rond
I_2 voilà et à gauche de *ce rond* tu vas prendre une petite barre (C8Homme)

(189) I_1 voilà donc *ce triangle-là* [+geste], il faut le mettre un peu plus à...
M_1 tu veux l'ajuster ? (C7Egypte)

Étant donné que le contexte linguistique ou gestuel ne fournit qu'un seul référent, la question de l'identification ne se pose même pas et le contenu linguistique des expressions démonstratives peut en effet être considéré comme entièrement disponible pour une reclassification.

Une observation de contextes plus complexes semble pourtant suggérer que la question de la reclassification ne se traite pas tout à fait indépendamment de celle de l'identification du référent. Lorsqu'il y a plusieurs référents accessibles, on observe que le démonstratif ne reclassifie pas toujours le référent qui, suivant notre modélisation, est supposé être l'élément focalisé du domaine actif.

(190) L'intégration d'un nouvel énoncé se fait sur la base de sa structure syntaxique. En fonction de *cette structure* s'appliquent des règles d'introduction de nouveaux référents et conditions. (ibid, page 60)

Ainsi, en (190), l'élément focalisé du domaine est le référent introduit par *l'intégration d'un nouvel énoncé*. Or, ce n'est pas ce référent qui est reclassifié par le démonstratif *cette structure* : ici, le démonstratif réfère plutôt à *sa structure syntaxique*.

A partir de cette observation, on pourrait tirer une des deux conclusions suivantes : soit, le calcul focal proposé est inadapté, soit un démonstratif n'identifie pas son référent en fonction de ce critère. Nous pensons que la solution se situe entre ces deux extrêmes. D'une part, le calcul focal retenu pour l'instant (celui de la théorie du Centrage) est simplifié, à la fois en ce qui concerne le calcul du focus et en ce qui concerne la représentation du focus par une valeur binaire. Mais d'autre part, l'hypothèse selon laquelle le démonstratif recrute son référent exclusivement en fonction du facteur focal semble être trop forte. Nous faisons alors une nouvelle proposition, considérant que le pouvoir reclassificateur d'un démonstratif est proportionnel au marquage focal du domaine d'interprétation :

le pouvoir reclassificateur d'un démonstratif *ce N* peut être exprimé par la valeur maximale de la différence acceptable entre le type associé à *N* et le type actuel du référent supposé ;

le marquage focal d'un domaine est une valeur numérique indiquant la probabilité avec laquelle il est possible d'isoler l'élément focalisé des autres éléments du domaine. Elle correspond à la différence positive entre la valeur focale de l'élément focalisé et la moyenne des valeurs focales des autres éléments du domaine ;

plus le marquage focal d'un domaine est élevé, plus la possibilité de reclassification du référent focalisé est grande. (Figure 62).

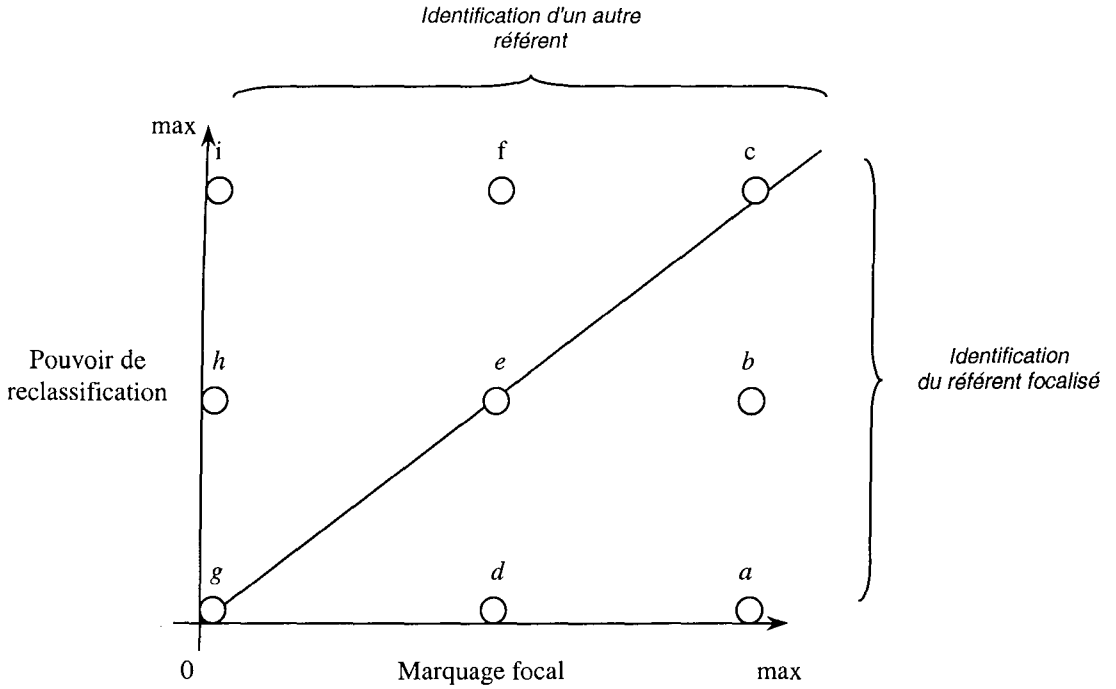


Figure 62 : Relation entre marquage focal d'un domaine et pouvoir reclassificataur d'un démonstratif

Lorsque le marquage focal d'un domaine est maximal, c'est-à-dire lorsque le référent focalisé peut être identifié sans ambiguïté, le pouvoir reclassificateur du démonstratif est également maximal (cas *a*, *b* et *c* de la Figure 62). C'est précisément le cas lorsqu'un référent est le seul à avoir été mentionné auparavant, ou encore lorsqu'il est identifié sans ambiguïté par un geste. C'est ce cas que considère F. Corblin lorsqu'il évoque une reclassification sans contraintes et c'est encore ce cas dont notre modèle tient compte pour l'instant. La modélisation proposée prédit alors un emploi optimal du démonstratif en tant que reclassificateur : étant donné qu'un démonstratif *ce N* restructure le contexte en insérant le référent dans un nouveau domaine de *N*, cette opération n'est justifiée que lorsqu'il y a apport d'information, c'est-à-dire lorsque ce nouveau domaine est différent du domaine d'origine. C'est cette différence entre les domaines avant et après l'interprétation qui permet de séparer le point *a* des points *b* et *c* de la Figure 62. Les exemples suivants illustrent ces trois cas :

(191) Prends un triangle. Colorie *ce triangle.* (*cas a*)

(192) Prends un triangle. Colorie *cette figure.* (*cas b*)

(193) Prends un triangle. Colorie *ce carré* (*cas c*)

Dans le cas *a*, l'opération de restructuration déclenchée par le démonstratif ne se justifie effectivement pas par rapport au pronom : un pronom mène au même domaine résultant et ceci sans opération de restructuration (Figure 63). Dans de tels cas, il est pourtant possible que l'opération de restructuration puisse être justifiée, soit a posteriori par une référence au complément de la partition du nouveau domaine (*ce grand triangle et pas un autre* ; cf. Corblin, 1995), soit par des facteurs discursifs dépassant le cadre actuel de notre modélisation (cf. de Mulder, 1998; Kleiber, 1986, 1988).

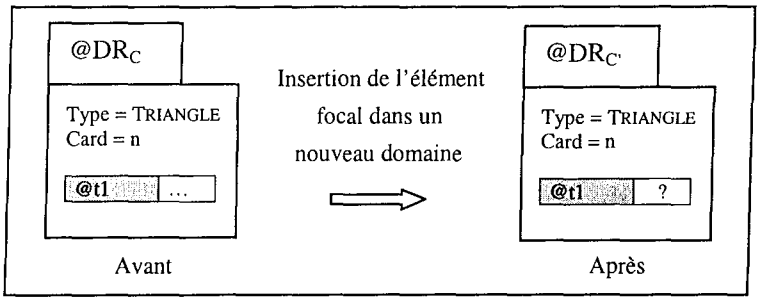


Figure 63 : Interprétation « ce triangle » – exemple (191)

Contrairement au cas *a*, les démonstratifs des cas *b* et *c* (exemples (192) et (193)) reclassifient tous les deux leur référent, mais *c* semble un cas limite : la différence entre le type du référent (*TRIANGLE*) et sa reclassification (*CARRÉ*) est effectivement maximale, dans la mesure où ces types sont mutuellement exclusifs. Mais même dans un tel cas, il n'y aura pas d'échec d'identification, le référent étant identifié sans ambiguïté dans un domaine à marquage focal maximal. En revanche, la reclassification maximale demande à être justifiée (par exemple par une erreur antérieure ou une transformation de l'objet). Autrement, elle peut être considérée comme une classification erronée.

Lorsque le marquage focal du domaine n'est pas maximal, la Figure 62 montre que le pouvoir reclassificateur du démonstratif diminue et ceci relativement à la focalisation des autres éléments du domaine (cas *d*, *e* et *f*). Sans chiffrer les valeurs focales respectives, nous estimons que le contexte introducteur des exemples (194) à (196) fournit un tel domaine : ici, notre calcul focal simplifié focalise en effet uniquement le référent de *un triangle*, mais celui de *un carré* ne nous semble pas pour autant inaccessible. Nous faisons donc l'hypothèse qu'il s'agit d'un domaine dont le référent focalisé peut être extrait avec moins de certitude que pour les cas *a*, *b* et *c*.

(194) Mets un triangle sur un carré. Colorie *ce triangle.* (*cas d*)

(195) Mets un triangle sur un carré. Colorie *cette figure* (*cas e*)

(196) Mets un triangle sur un carré. Colorie *ce carré.* (*cas f*)

Pour ces exemples, on constate effectivement qu'au-delà d'un certain seuil, la différence entre le type de l'élément focalisé (*TRIANGLE*) et le type de la reclassification rend difficile le maintien de l'élément focalisé comme référent à reclassifier. Si le cas *d* ne pose pas de problème pour ce qui est de l'identification de son référent, le cas *e* est déjà plus délicat : *cette figure* peut reclassifier les deux éléments du domaine et il semble difficile de les départager en tant que référents possibles de l'expression démonstrative, même si notre modèle prédit pour l'instant une reclassification de l'élément de type *TRIANGLE*. Enfin, le cas *f* ne fait pas de doute : ici, la référence virtuelle de *ce carré* participe clairement à l'identification du référent et sélectionne l'élément non (ou moins) focalisé du domaine. Au vu de ces exemples, il semble alors que dans les configurations à marquage focal moyen s'instaure un équilibre entre l'effort de reclassification du référent préfocalisé et une opération de changement focal, menant à l'identification (puis, éventuellement, à la reclassification) d'un autre élément du domaine. Nous pensons pourtant que le principe d'identification externe du référent d'un démonstratif peut être maintenu : dans la mesure où l'effort de reclassification ne peut être évalué que relativement à un référent supposé, celui-ci doit bien être fourni par le contexte. Nous faisons donc l'hypothèse que le référent supposé correspond systématiquement au référent focalisé et que ce n'est qu'en cas de reclassification trop coûteuse ou conflictuelle que les autres éléments du domaines sont évalués en tant que candidats référentiels. C'est ainsi qu'en (197), le démonstratif *cette structure* réfère non pas à l'élément focal premier (*l'intégration...*), mais bien à l'entité introduite par *structure syntaxique*.

(197) L'intégration d'un nouvel énoncé se fait sur la base de sa structure syntaxique. En fonction de cette structure s'appliquent des règles d'introduction de nouveaux référents et conditions.

Concernant la troisième série de points de la Figure 62 (les cas *g*, *h* et *i*), elle relève d'un domaine à marquage focal minimal. Dans des contextes introduisant des référents par une coordination (exemples (198) à (200)), il est en effet impossible d'extraire individuellement un des référents en raison de sa seule valeur focale. La seule différence avec des exemples tels que (179) consiste dans le fait que les éléments du domaine, bien qu'indifférenciables selon leur valeur focale, restent accessibles individuellement, ce qui n'est pas le cas pour des référents introduits au pluriel (*deux triangles*).

(198) Prends un triangle et un carré. Colorie *ce triangle.* (*cas g*)

(199) Prends un triangle et un carré. Colorie *cette figure.* (*cas h*)

(200) Prends un triangle et un carré. Colorie *ce carré.* (*cas i*)

Le marquage focal minimal du domaine explique alors pourquoi les exemples (198) et (200) sont interprétables, mais difficilement acceptables. Non seulement l'identification du référent est entièrement dépendante de la référence virtuelle du démonstratif (et ne se distingue donc plus du principe propre aux définis), mais en plus, le référent est extrait d'un domaine à cohésion particulièrement forte, ce qui rend d'autant plus difficile son insertion dans un nouveau domaine. En ce qui concerne l'exemple (199), il y a échec d'identification du référent, car celui-ci n'est ni identifiable par des critères de focalisation, ni même par la référence virtuelle du démonstratif.

Pour résumer ces réflexions sur la relation entre la focalisation d'un domaine et le pouvoir reclassificateur d'un démonstratif, nous retenons que la modélisation proposée dans les chapitres précédents prédit bien les cas relevant des configurations extrêmes (marquage focal maximal et minimal). Afin de pouvoir intégrer les configurations intermédiaires, il faudrait gérer un calcul focal plus fin, admettant des valeurs focales scalaires et implémenter une procédure de décision sur l'acceptabilité d'un démonstratif en fonction de la structure focale du domaine actif. Toujours est-il que le cadre de la modélisation proposée n'est pas incompatible avec une telle évolution.

9.4.2 Exemple de traitement

L'exemple que nous avons choisi pour illustrer la modélisation de l'interprétation des démonstratifs va au-delà du traitement du démonstratif lui-même. Contrairement aux indéfinis et définis qui forment des séries référentielles homogènes, les démonstratifs, déjà rares dans le corpus Ozkan (10 en tout), se trouvent encore plus rarement par série. Cela est tout à fait explicable par leur fonction qui est précisément le signalement de ruptures domaniales. Or, une rupture se définit par rapport à une structure canonique, ce qui signifie que la modélisation de l'interprétation des démonstratifs est inconcevable sans une prise en compte de la structuration contextuelle opérée en amont par d'autres marqueurs référentiels. Par conséquent, la modélisation de l'interprétation de l'exemple (201) reprend en partie des mécanismes introduits séparément pour le traitement des définis et des indéfinis, dont elle illustre aussi l'interaction.

(201) I_1 donc le deuxième dessin représente *une route*
M_1 ouais
I_2 donc faite avec *des barres verticales* et
I_3 donc au bord de *cette route*, il y a *deux maisons.*
I_4 donc *une maison* qui se trouve à gauche de euh de *cette route*
I_5 et *une autre* à droite. (C9Maisons)

La première expression est l'indéfini *une route.* Selon le principe des indéfinis, elle extrait un élément @r du domaine générique de *ROUTE* @R (cf. en haut à droite de la Figure 64a, page 205).

L'indéfini *des barres verticales* est interprété selon le même schéma : il extrait une représentation @v du domaine générique de *BARRES VERTICALE* @V (cf. en haut à gauche de la Figure 64a). Étant impliqués dans un état commun (*être_fait_avec*), les deux représentations spécifiques @r et @v forment un ensemble, représenté par @f_a. En tant que résultat de la composition, la représentation pour la route (@r) est l'élément focalisé de ce domaine (cf. en bas à gauche de la Figure 64a).

Le démonstratif *cette route* est interprété à partir de cette structure du contexte. Selon le principe que nous avons formulé, il identifie le référent sur des critères externes. En l'absence d'un geste de désignation, ce critère est la focalisation résultante de l'histique du dialogue. Ici, le dernier élément focalisé est la représentation de la route @r. Cette représentation est donc identifiée comme référent, extraite de son domaine d'origine et insérée dans un nouveau domaine @ de type *ROUTE*, à l'intérieur duquel il est identifiable par des informations apportées par le prédicat (*maisons_au_bord*). Le type du domaine ainsi que le complément de la partition (pour l'instant vide) sont considérés comme provisoires (cf. en bas à droite de la Figure 64a).

L'expression suivante est l'indéfini *deux maisons*. De la même façon que les indéfinis précédents, elle extrait son référent d'un domaine générique qui est ici *MAISON* @M (Figure 64b). Comme nous l'avons expliqué dans la section précédente, ce nouveau référent @m est regroupée avec la représentation de la route @r, en raison de la préposition spatiale *au bord de* qui les relie. Le résultat de ce groupement se superpose au domaine construit précédemment pour le démonstratif (@f), dont il change le type en *FIGURE* et instancie le complément de la partition par la représentation pour la maison @m. Par ailleurs, celle-ci sera l'élément focalisé de ce domaine, en raison de sa position en tant que cible (cf. la partie supérieure de la Figure 64c).

Ensuite, l'expression *une maison* extrait aléatoirement un élément du domaine @m et le focalise comme référent, en l'opposant au reste des éléments du domaine par sa position vis-à-vis de la route (@m1). Un nouveau démonstratif, *cette route*, doit alors être interprété : ici, notre principe selon lequel le démonstratif identifie systématiquement le référent focalisé échoue. Il y a non seulement improbabilité de reclassification d'un élément de type *MAISON* en *ROUTE*, mais encore impossibilité d'égalité référentielle entre le site et la cible de la préposition *à gauche de*. Ces facteurs devraient donc mener à l'identification d'un autre élément, @r en l'occurrence, comme référent du démonstratif. Celui-ci est alors extrait de son domaine d'origine (@f) et inséré dans un nouveau domaine @rm1, résultant du groupement de @r avec la maison précédemment extraite @m1, en raison de la préposition *à gauche de* (cf. la partie inférieure gauche de la Figure 64c).

En comparant ce nouveau domaine au domaine d'origine (@f), on constate que la seule différence consiste en l'extension du complément de la partition (avant : @m, représentation *deux maisons* ; après : @m1, représentant *une maison*) et en une sur-spécification du critère de différenciation (avant : position horizontale, exprimée par *au bord de* ; après : position horizontale, exprimée par *à gauche de*). Étant donné que @m implique @m1, le changement de l'extension de la partition n'apporte pas de nouvelles informations. Par ailleurs, la sur-spécification du critère de différenciation (*au bord de* vs. *à gauche de*) fait double emploi avec des informations sauvegardées ailleurs : ce critère a en effet déjà servi à distinguer @m1 du reste des deux maisons @m. De ces observations, il est alors possible de conclure à un emploi sous-optimal du démonstratif. Cela nous semble en effet être le cas : l'absence du démonstratif ne conduit pas à des problèmes de compréhension et le site de la préposition *à gauche de* aurait pu être inféré à partir de celui de l'expression *au bord de*.

Enfin, *une autre* extrait son référent du domaine des deux maisons, @m. Il est intéressant de noter ici que *autre* à l'indéfini n'est pas destiné à épuiser la partition du domaine. Cependant, étant donné que la cardinalité du domaine est connue, l'extraction d'un des éléments du complément de la partition

mène *de facto* à son épuisement. Cela ne pose pas de problèmes spécifiques, mais l'emploi du défini *l'autre* aurait été aussi, sinon plus approprié.

En résumé de l'application de notre modélisation à l'interprétation des démonstratifs, nous constatons qu'elle est capable de tenir compte des emplois des démonstratifs tant que leurs référents sont entièrement identifiables par des critères de focalisation ou des gestes de désignation. La modélisation des effets conjoints de la focalisation et de la référence virtuelle pour l'identification du référent, telle qu'en I_4, doit faire partie d'investigations futures. En revanche, il est intéressant de disposer d'un cadre de modélisation qui permet de prédire des emplois sous-optimaux de démonstratifs à partir d'une comparaison des structures contextuelles avant et après l'interprétation. D'après nos connaissances, aucun modèle computationnel du calcul référentiel n'a jusqu'ici intégré cet aspect.

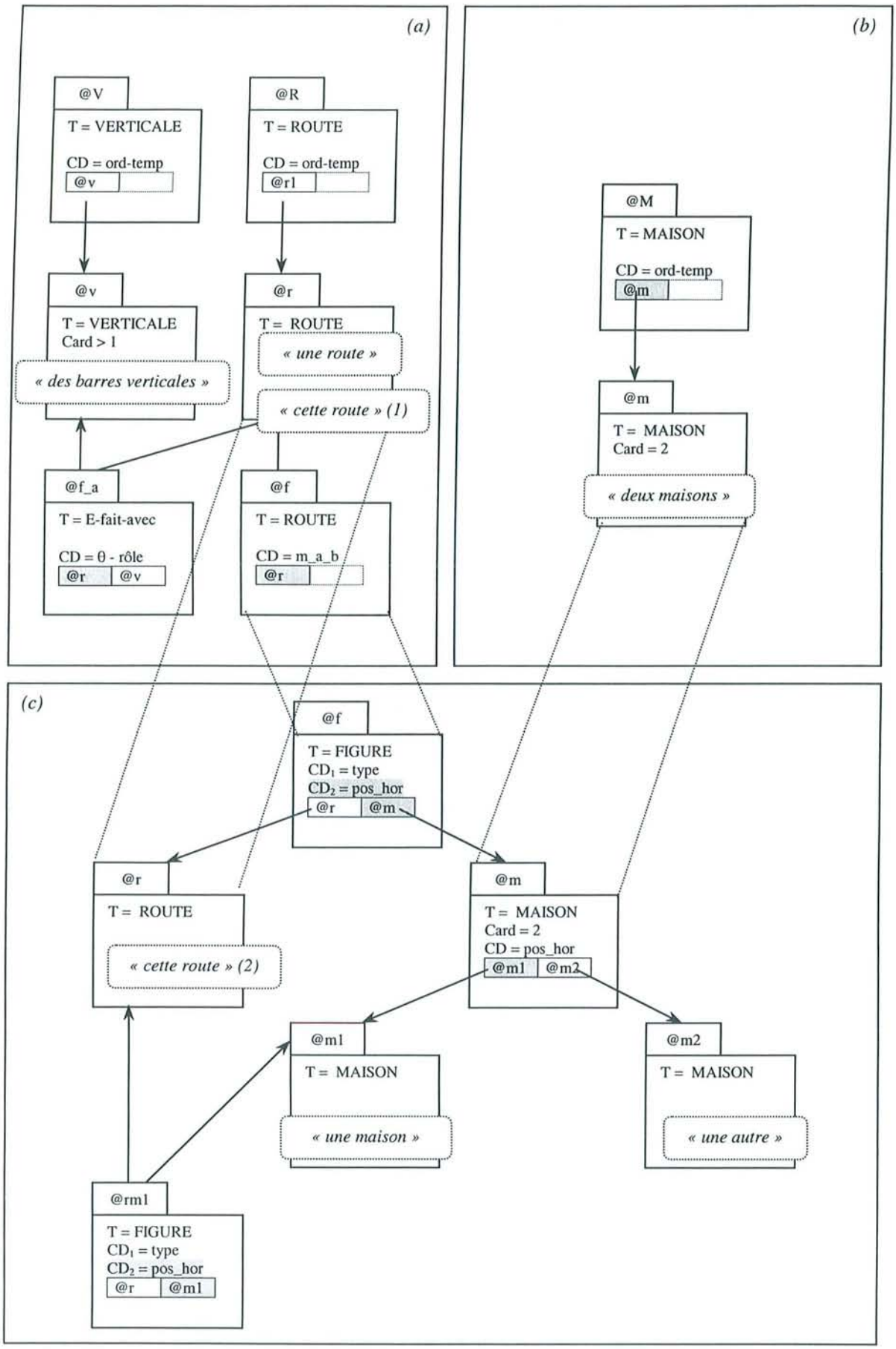


Figure 64: Interprétation de l'exemple (201)

9.5 Les pronoms : prédictions et traitement d'un exemple

9.5.1 Prédictions pour les expressions pronominales

- *Le pronom personnel de 3ème personne s'interprète dans un domaine focalisé.*

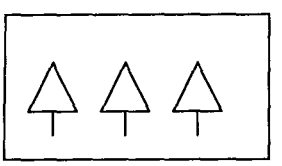
Les prédictions possibles pour l'emploi d'une expression pronominale sont comparables à celles des modélisations courantes consacrées au traitement des pronoms interphrastiques (comme la théorie du Centrage, Grosz et al., 1995). La modélisation prédit en effet qu'un pronom nécessite, pour s'interpréter, un domaine de référence ayant isolé et focalisé un référent. Cela exclut une interprétation au croisement des lignes 5 et 6 de la colonne 5 du Tableau 20, page 162.

La particularité de notre modélisation consiste dans le fait que la focalisation peut provenir d'autres sources que discursives. Suivant G. Kleiber (1990), la focalisation peut effectivement être le résultat d'une saillance situationnelle, impliquant le référent comme argument classifié. Cette situation se présente d'ailleurs fréquemment dans le corpus Ozkan, où l'objet courant à positionner peut être considéré comme actant principal d'une situation saillante. C'est ainsi qu'un certain nombre de pronoms sans antécédent discursif peuvent être interprétés : en (202), le pronom *la* renvoie sans ambiguïté à un référent perceptif. Il s'agit de l'objet manipulé, lui-même appartenant à un domaine d'éléments auparavant classifiés comme *PYRAMIDES* :

(202) I_1 bon, alors il faut faire des pyramides
I_2 [geste : prend un grand triangle]
I_3 donc je vais essayer de *la* faire quand même (C5 Egypte)

En (203), le pronom *ils* (I_6) fait référence à l'ensemble des figures de la scène. Mais contrairement à l'exemple précédent, les référents n'ont jamais été classifiés par une mention textuelle. Dans ce cas, on peut supposer que leur classification en tant qu'objets de genre masculin provient d'une catégorie de base, dans ce cas *ARBRE* ou *SAPIN*, déductible de la mention de *FORÊT* en I_1.

(203) I_1 il faut qu'on fasse une forêt, alors il va falloir une petite barre
I_2 et dessus tu vas y mettre un petit triangle
I_3 voilà et ça tu vas le répéter encore deux fois
M_1 je suis pas très bavard mais tout m'a l'air bien clair
I_4 ouais ben j'espère, je te fais travailler, hein
I_5 voilà on a fini
I_6 ah oui *ils* sont pas dans l'ordre, pardon, j'avais pas vu (C8Forêt)



- *Un pronom ne change pas la structure de son domaine.*

Cette deuxième prédiction reflète l'intuition de G. Kleiber (1990, 1994) selon laquelle un pronom est un marqueur référentiel original, car il indique que l'on parle du référent saillant en prolongeant la structure qui l'a rendu saillant. Pour notre modélisation, cela signifie qu'un pronom est toujours en concurrence directe avec un autre marqueur lorsque l'opération de restructuration associée à cet autre marqueur aboutit au même domaine de référence que le domaine avant l'interprétation. Nous avons déjà discuté quelques exemples dans les sections précédentes et nous les reproduisons donc juste pour mémoire :

(204) a. Prends un triangle. Colorie *le triangle*.
b. Prends un triangle. Colorie *le*.

(205) a. Prends un triangle. Colorie *ce triangle*.
b. Prends un triangle. Colorie le.

Par ailleurs, la continuité domaniale associée à l'usage d'un pronom s'illustre également à travers le fait que l'interprétation des expressions subséquentes n'est pas affectée par la présence d'un pronom, comme le montre une comparaison des variantes (a) et (b) de l'exemple (206) :

(206) a. Prends deux triangles. Mets en un sur la gauche. Colorie *le* en rouge. Supprime l'autre.
b. Prends deux triangles. Mets en un sur la gauche. Supprime l'autre.

9.5.2 Exemple de traitement

Le traitement des pronoms sera illustré sur le début de l'extrait ayant déjà servi à l'illustration de l'interprétation des indéfinis et des définis :

(207) I_{14} et puis maintenant faudrait faire la ligne d'horizon
I_{15} il faut prendre une grande euh une grande horizontale
I_{16} et *la* placer à la pointe des triangles des deux grands triangles
M_3 ah mais *elle* va pas être horizontale alors

Cela nous permet de reprendre comme domaine initial la structure contextuelle construite au cours de l'interprétation de l'indéfini *une grande horizontale* (I_{15}) : elle est reproduite dans la Figure 65a (page 197).

Le pronom *la* s'interprète dans le domaine @GH qui met à disposition un élément focalisé et classifié en tant qu'objet de type féminin (*BARRE*). Celui-ci est identifié comme référent et conformément à l'opération de restructuration propre aux pronoms, la structure du domaine @GH n'est pas modifiée.

Dans la suite de l'interprétation, le référent précédent @gh1 et @pe (référent de *la pointe,* cf. Figure 65b) seront regroupés en tant que participants du même événement de placement. A l'intérieur de ce nouveau domaine @f, le référent du pronom reste l'élément focalisé, car il joue le rôle de cible dans l'acte de positionnement (Figure 65c). Le second pronom, *elle*, est interprété dans ce nouveau domaine. Il en isole l'élément focalisé, toujours sans restructurer le domaine @f.

Pour l'instant, nous n'avons pas intégré dans la modélisation des contraintes supplémentaires, dues à des restrictions sélectionnelles ou à la tâche. Des tests effectués par S. Lappin et H. Laess (1994) ont montré que l'intégration de connaissances sur des co-occurrences de paires lexicales[99] améliore le système RAP[100] de 3% seulement. Ce taux relativement faible peut être dû à la nature du corpus, comportant majoritairement des pronoms à antécédent intraphrastique et à l'insuffisance des seules connaissances des co-occurrences lexicales. Dans un corpus comme le nôtre, des connaissances supplémentaires permettraient probablement d'améliorer le calcul référentiel. En (208), il serait par exemple possible d'exclure le référent de *le gros triangle* comme référent du pronom *le* : il est en effet impossible de placer un objet « décalé par rapport à lui-même » :

(208) I_1 ensuite on a les pyramides dans le désert
alors il faut que tu prennes les deux euh enfin le triangle
I_2 le gros triangle...
I_3 voilà et tu *le* reprends une deuxième fois un peu décalé par rapport au premier (C8 Egypte)

La prise en compte de ce type de connaissances passera donc par une augmentation des informations dans les représentations génériques. Le même type de connaissance permettrait par ailleurs de traiter des glissements aux génériques (bien qu'absents du corpus Ozkan), tels qu'en (209) :

(209) Aurore n'aime pas les triangles parce qu'*ils* ont trois côtés.

[99] Ces connaissances sont extraites statistiquement d'un corpus. Elle permettent, pour un pronom ambigu, de choisir comme antécédent le nom ayant la plus grande la probabilité d'être combiné au verbe associé au pronom.
[100] présenté dans la section 5.2.1

Dans ce cas, en l'absence d'un domaine mettant à disposition un référent focalisé à cardinalité convenable, il faudrait accéder au type de l'élément focalisé et tester si les informations prédicatives sont vérifiées sur ce type. La structure de nos représentations, par le pointeur sur le type et par la possibilité d'intégrer des connaissances encyclopédiques, permet en effet d'effectuer cette opération.

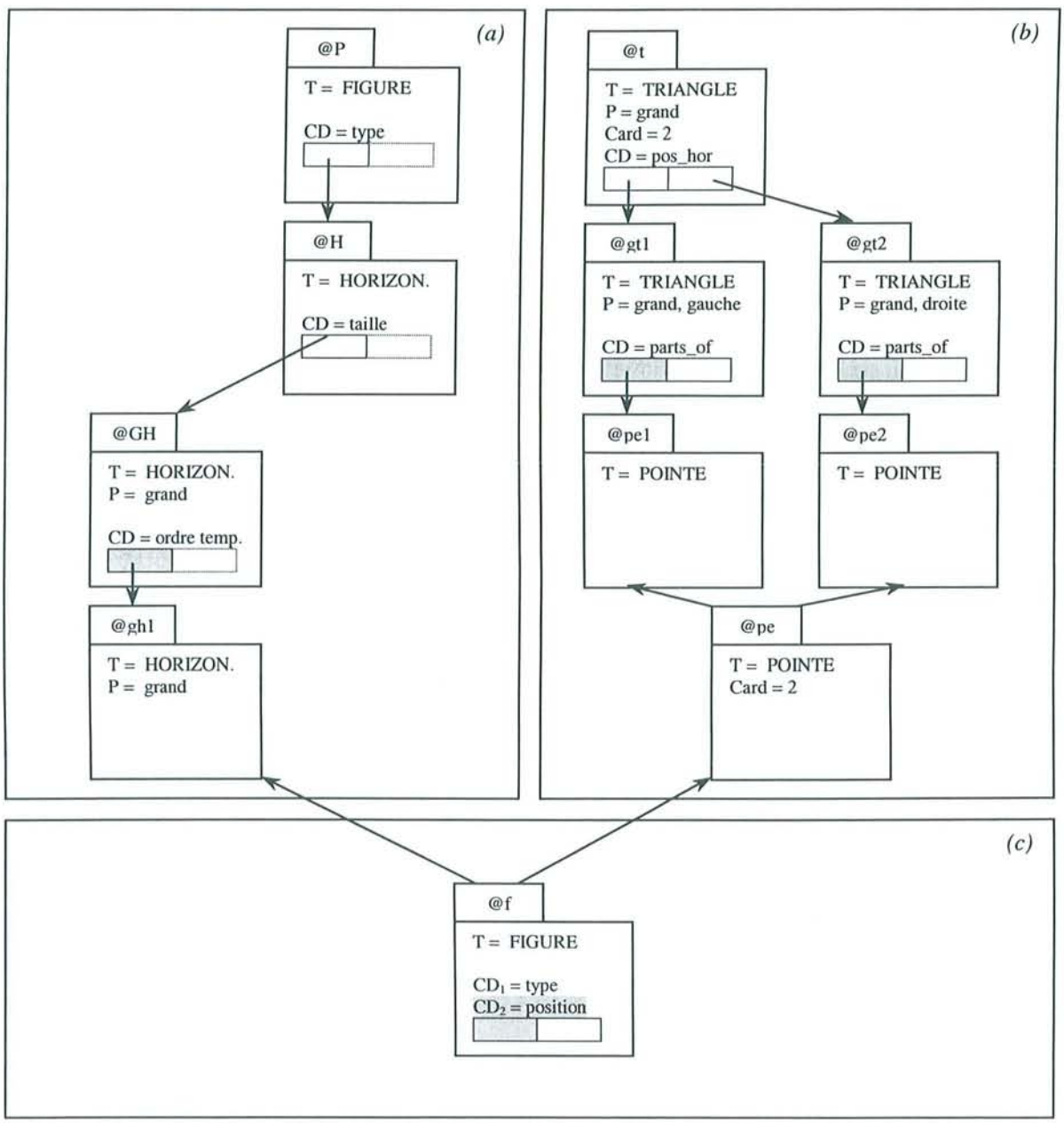


Figure 65 : Interprétation de l'exemple (207)

9.6 Bilan

En conclusion de ce chapitre, nous retenons tout d'abord qu'une grande partie des phénomènes référentiels du corpus examiné est prédictible et traitable dans le cadre de la modélisation proposée ici. De plus, parmi ces phénomènes, on trouve un certain nombre de références ayant été exclues ou considérées comme problématiques par les modèles existants. Nous mentionnerons ici les emplois génériques des indéfinis, définis et pronoms, les emplois associatifs des descriptions définies, les

démonstratifs accompagnés ou non d'un geste de désignation ou encore les groupes nominaux sans noms, combinés ou non à des expressions d'altérité ou d'ordre.

En plus de la couverture d'un plus grand nombre de phénomènes référentiels, nous proposons un modèle unifié pour le traitement de différents types d'expressions référentielles : un même mécanisme fondamental – le calcul d'un domaine de référence sous-spécifié ainsi qu'une opération de restructuration du domaine d'interprétation – est décliné pour tenir compte de l'interprétation des expressions indéfinies, définies, démonstratives et pronominales. A partir de ce mécanisme fondamental, nous sommes capable de formuler des prédictions sur la compatibilité entre des domaines contextuels et des expressions à interpréter : d'une part, cela nous permet de prédire l'inadéquation entre une expression et un contexte, ce qui contribue à éliminer des domaines d'interprétation incompatibles. D'autre part, nous sommes capable de choisir, parmi plusieurs domaines compatibles, celui à l'intérieur duquel l'expression aura un effet interprétatif maximal.

La présentation de ces prédictions, ainsi que leur mise en oeuvre sur des extraits de dialogues réels nous a par ailleurs amenée à discuter les limites actuelles de notre modélisation : pour l'instant, les structures utilisées sont construites essentiellement à partir des propriétés des objets et mises à jour par les expressions qui y réfèrent. Un certain nombre de phénomènes nécessiteraient pourtant une prise en compte d'informations supplémentaires.

Parmi celles-ci, on peut citer d'abord les informations concernant les événements auxquels les objets participent. Comme nous l'avons mentionné, ce type d'information permettrait de traiter les cas de coréférence entre indéfinis et de formuler des contraintes supplémentaires sur l'interprétation des pronoms. Par ailleurs, ces connaissances sont nécessaires afin de résoudre des anaphores associatives sur des participants d'une éventualité (*un meurtre – la victime*). Le traitement de ces phénomènes pourrait être intégré dans notre modélisation, à condition de disposer d'une représentation des événements et d'une définition des opérations possibles sur ces entités. Les travaux de A. Reboul (2000) ainsi que de O. Grisvard (2000) ont d'ores et déjà montré que notre cadre de modélisation, inspiré de la théorie des représentations mentales, est adapté à une telle extension.

Deuxièmement, l'augmentation des informations relatives aux propriétés des types permettrait d'intégrer facilement le traitement des références génériques. Il ne s'agit pas là d'un problème de modélisation – la prédiction théorique de ces références fonctionne bien –, mais d'un problème d'augmentation de la base des connaissances encyclopédiques. Les structures de représentation que nous utilisons prévoient un effet des champs dédiés à contenir ce type d'informations.

Enfin, lors du traitement des démonstratifs, nous avons abordé la question du calcul focal. Étant donné le grand nombre de travaux abordant la résolution référentielle quasi-exclusivement sous l'angle de la focalisation[101], une simplification à cet égard peut paraître peu orthodoxe. Nous l'assumons pourtant pleinement et ceci pour deux raisons : premièrement, nous faisons précisément l'hypothèse que le calcul focal (ainsi que plus largement le facteur de saillance) n'est qu'un facteur parmi d'autres intervenant dans les processus d'interprétation référentielle. Deuxièmement, il n'y a pas de consensus théorique (ni pratique) sur la manière de calculer le focus, comme en témoigne la variété des algorithmes proposés. Par conséquent, nous avons voulu proposer une modélisation indépendante d'une vision théorique particulière. Les structures de données proposées sont compatibles avec plusieurs algorithmes de calcul focal et s'adaptent par ailleurs sans problème à un remplacement du booléen (simplificateur de la valeur focale, en ce qu'il marque uniquement l'élément le plus saillant de la partition) par une échelle de valeurs focales. Par ailleurs, la simplification que nous avons adoptée n'empêche pas la mise en œuvre des principes référentiels que nous considérons comme

[101] cf. en particulier le chapitre sur les approches cognitives (chapitre 0) et celui sur les algorithmes et implémentations (chapitre 5).

fondamentaux : la modélisation proposée permet déjà de faire des prédictions intéressantes, non seulement pour l'emploi d'expressions pronominales (comme la majorité des modèles basés exclusivement sur des facteurs focaux), mais pour différents marqueurs référentiels.

Enfin, soulignons que notre tableau synthétique se lit dans les deux sens. Si nous nous sommes concentrée ici sur le sens de lecture correspondant à l'interprétation d'une expression (colonne → ligne), un deuxième sens de lecture (ligne → colonne) permet effectivement de faire des prédictions sur la génération : à partir de la structuration d'un domaine contextuel donné et des connaissances sur l'effet interprétatif des différentes expressions, il est en effet possible de trouver le « bon » type d'expression pour désigner un élément de ce domaine. A titre d'exemple, si le domaine courant possède une partition focalisée dont on souhaite désigner l'élément focalisé, on devrait alors générer, pour une efficacité maximale, un pronom.

10 Entre corpus et théorie : l'annotation (co)référentielle et l'interprétation de *autre*

10.1 Introduction

Ce chapitre porte sur la relation entre l'annotation de corpus et la validation de théories linguistiques, en abordant plus particulièrement le domaine de la (co)référence. Au premier abord, cette relation peut être perçue comme un « cercle vicieux » : la validation d'une théorie sur des données linguistiques présuppose un corpus annoté et l'annotation d'un corpus doit reposer sur des principes cohérents, donc eux-mêmes issus d'une réflexion théorique. Une façon de contourner ce problème consiste alors à n'annoter que des phénomènes pour lesquels on dispose d'une base théorique suffisamment consensuelle. Cette solution a par exemple été adoptée par A. Tutin et al. (2000) qui annotent un large corpus du français en relations anaphoriques, mais en excluant l'annotation de toutes les descriptions définies non elliptiques : « *Dealing with this kind of anaphoric expressions appeared premature given the lack of satisfactory formal description [...]* »[102] (Tutin et al., 2000 : 2). Par conséquent, le corpus annoté perd une partie de son intérêt. Il pourra certes servir à l'entraînement et l'évaluation d'outils de traitement automatique de la langue correspondant à l'état des connaissances au moment de l'annotation, mais il ne permet pas de valider des modèles théoriques portant sur les phénomènes les plus intéressants : c'est-à-dire ceux pour lesquels il n'y pas encore d'accord dans la communauté scientifique, mais qui devraient permettre, à terme, de développer des outils plus performants.

Notre modélisation des processus référentiels, développée dans les chapitres précédents, nous permet, entre autres, de faire des prédictions sur l'interprétation référentielle des expressions d'altérité en *autre*, prédictions que nous aimerions valider sur un corpus de dialogues. Or, ces recherches se situant précisément dans un des domaines à faible consensus théorique, les corpus annotés au niveau (co)référentiel, peu nombreux pour le français (Tableau 21), s'avèrent inappropriés à la validation de nos hypothèses concernant l'interprétation de *autre*.

Premièrement, la nature des corpus ne s'y prête pas : il s'agit de textes monologiques écrits, trop éloignés des conditions naturelles d'un dialogue oral. Deuxièmement, *autre* se combine avec tous les déterminants (*un autre N, l'autre N, cet autre N*), ce qui exclut les corpus annotés seulement pour des syntagmes à détermination particulière. Par ailleurs, une modélisation des phénomènes référentiels qui ferait l'impasse sur certains types d'expressions référentielles (telles que les descriptions définies), s'avérerait parfaitement inadaptée, non seulement du point de vue de l'opérationalité du système de dialogue résultant, mais aussi d'un point de vue cognitif. Troisièmement, concernant le type des liens annotés, *autre* entretient certes des liens avec d'autres entités (*un carré – un autre*), mais il apparaît aussitôt que ces liens ne sont pas de nature coréférentielle au sens strict : il n'y a pas identité des référents. Il suffit en effet d'observer des dialogues en situation naturelle pour se convaincre non seulement du fait que les liens entre expressions référentielles semblent loin d'être majoritairement coréférentiels[103], mais aussi du fait que bon nombre d'expressions ne réfèrent pas à des antécédents textuels, mais à des objets accessibles dans la situation de communication. En ce qui concerne *autre*, les liens peuvent s'établir effectivement avec des objets non linguistiques (un sujet manipulant un carré – « *et maintenant, il faut en prendre un autre* »).

[102] « Le traitement de ce type d'expressions anaphoriques semble prématuré, étant donné l'absence d'une description formelle satisfaisante [...]. »

[103] M. Poesio et R. Vieira (1998) ont montré que 70% des descriptions définies d'un corpus journalistique anglais ne sont pas utilisées au sens coréférentiel strict (même tête nominale, même référent).

Auteurs	Bruneseaux et Romary (1997)	Popescu-Belis et Robba (1998)	Clouzot et al. (2000)	Tutin et al. (2000)
Filiation	Loria, Nancy	Limsi, Orsay	Cristal-Gresec, Grenoble	Cristal-Gresec, Grenoble, XRCE Grenoble
Corpus	Extrait du *Père Goriot,* (Balzac)	Extrait de *Vittoria Accoramboni,* (Stendhal)	Le Monde Diplomatique	Textes journalistiques, scientifiques et littéraires
Nbre de mots	30000	10000	95000	1000000
Phénomènes annotés	certaines expressions référentielles, liens de coréférence large[104]	toutes les expressions référentielles, liens de coréférence stricte[105]	expressions anaphoriques pronominales	expressions anaphoriques, sauf descriptions définies
Schéma d'annotation	compatible MATE[106]	compatible MATE	format propriétaire (base TEI[107])	TEI
Accès recherche	oui	oui	?	non

Tableau 21 : Ressources françaises annotées au niveau coréférentiel

Au vu de cette situation, nous avons décidé d'annoter nous-mêmes un corpus de dialogues : il s'agit du corpus Ozkan, présenté dans la section. Cela nous a amenée à réfléchir, préalablement, à un schéma d'annotation adapté. Étant donné qu'il n'y a pas actuellement de consensus théorique autour des phénomènes (co)référentiels à annoter, notre objectif principal a été d'éviter l'écueil de la projection d'une théorie particulière sur le schéma d'annotation. Cela était non seulement nécessaire pour obtenir des résultats pertinents sur la validité de notre modèle, mais aussi (et surtout) pour rendre la ressource annotée réutilisable pour d'autres investigations. Nous avons donc voulu proposer un schéma qui soit compatible avec les schémas utilisés antérieurement, tout en étant modulaire et extensible à des phénomènes nouveaux et le plus indépendant possible d'un point de vue théorique particulier.

Le présent chapitre est organisé comme suit : sur la base de notre modèle de la référence et des travaux récents issus de la linguistique descriptive, nous proposons d'abord une modélisation des principes référentiels associés à *autre*. Cette modélisation nous permet de formuler des prédictions, allant au-delà des intuitions naturelles, que nous aimerions valider sur le corpus Ozkan. Une deuxième partie sera donc consacrée à l'élaboration d'un système d'annotation (co)référentielle qui soit non seulement adapté à nos besoins, mais aussi suffisamment générique pour couvrir les besoins en annotation coréférentielle apparaissant au travers d'autres travaux d'évaluation. Le schéma proposé est indépendant d'une théorie ou d'un objectif d'évaluation particulier, suit les recommandations européennes du projet MATE et de la TEI et donne une justification théorique du type des liens et des principes de codage en cas de conflit. Par ailleurs, il préconise une séparation entre annotation coréférentielle et annotation référentielle, mise en œuvre par une architecture modulaire, spécifiée en méta-schémas XML. Dans une troisième partie, nous rapporterons quelques problèmes rencontrés lors de l'application de ces principes de codage à notre corpus et nous présenterons des résultats qui confirment notre modélisation de l'interprétation référentielle de *autre*.

[104] Par coréférence large, nous entendons différents types de liens entre deux entités discursives (identité du référent, anaphore associative,...).
[105] Par coréférence stricte, nous entendons un lien entre deux entités discursives ayant le même référent.
[106] Multilevel Annotation Tools Engineering (Mengel et al., 2000)
[107] Text Encoding Initiative (Sperberg-McQueen et Burnard, 1994)

10.2 L'interprétation de *autre*

10.2.1 Approches linguistiques

L'expression *autre* a fait l'objet de différents travaux linguistiques : elle a, par exemple, été étudiée à l'intérieur du paradigme des expressions comparatives (Pottier, 1982), comparée aux expressions ordinales *premier/second* (Berrendonner et Reichler-Béguelin, 1996), reconsidérée par rapport à la classe des indéfinis (van Peteghem, 1996) ou replacée dans le couple *l'un/l'autre* (Schnedecker, 1998). Tous ces travaux abordent, sous des formes diverses, les deux questions suivantes :

Quelle est la particularité sémantique de l'expression *autre* ?

Comment cette particularité participe-t-elle à l'identification du référent ?

Concernant la première question, A. Berrendonner et M.-J. Reichler-Béguelin (1996) considèrent que la sémantique de *autre* **présuppose un objet-repère antagoniste, ontologiquement différent du référent recherché, mais ayant en commun avec ce dernier un attribut de catégorisation**. Ces deux propriétés-clé – le contraste avec un objet-repère O_R et l'appartenance à un ensemble homogène – sont également discutées dans d'autres travaux. M. van Peteghem (1996), en particulier, considère que « autre *présuppose l'existence d'une classe référentielle à laquelle appartiennent les deux constituants mis en relation, mais en même temps, il implique que le corrélé inférieur soit exclu comme référent potentiel du SN contenant* autre ». (Peteghem, 1996 :196)

Dans les cas les plus évidents, l'appartenance à une même classe référentielle se traduit par un même nom tête pour les deux syntagmes :

(210) I_1 euh maintenant il faut prendre *un grand triangle*
et le mettre à gauche de l'écran
I_2 maintenant prendre *un autre grand triangle...* (C11Eglise)

Mais il faut souligner que l'accès au référent du syntagme contenant *autre* ne dépend pas de façon cruciale de l'opération qui consiste à récupérer un nom tête dans le contexte linguistique (Schnedecker, 1998 : 189). Ce constat s'appuie sur deux observations : la possibilité de reclassifier l'ensemble par le nom tête du syntagme d'altérité (211) et l'impossibilité à rétablir, dans certaines constructions elliptiques, la catégorie à partir du seul contexte linguistique (212) :

(211) I_1 voilà donc l'église est faite là
I_2 et donc à côté à sa gauche il y a *une autre maison* (C9Eglise)

(212) [contexte visuel : un triangle et un rond à l'écran]
I_1 j'ai déplacé le rond... tu ne le vois pas sur l'écran ?
M_1 si si si...voilà ramène-le sur *l'autre* / sur **l'autre rond.* (C5Lampe)

Il semblerait donc que le nom tête de l'expression contenant *autre*, s'il existe, contraint la recherche en mémoire d'un objet-repère compatible avec cette catégorisation. Mais la caractéristique sémantique fondamentale et propre à *autre* resterait le principe de non identité : *autre* profilerait d'abord un référent non identique à un repère, que celui-ci soit accessible textuellement ou dans la situation de communication (Schnedecker, 1998 : 191)

Dans cette même perspective, M. van Peteghem (1996) considère que *autre* fait partie des « confrontatifs » (Damourette et Pichon, 1911-1940), ce qui permettrait en particulier un rapprochement avec d'autres constructions grammaticales corrélatives introduites par *que*, comme les consécutives (*tellement...que*) ou les comparatives (*plus...que*). La construction « prototypique » pour *autre* serait alors une construction de type *autre...que*, mais dans certains emplois, la subordonnée

manquerait et le terme de comparaison devrait alors être retrouvé dans le contexte. Si cette vision des choses est compatible avec la sémantique contrastive ou comparative de *autre* observée par ailleurs (Pottier, 1982), elle n'arrive pas à expliquer, à elle seule, pourquoi la construction corrélative fonctionne en (213), mais pas en (214).

(213) I_1 maintenant prendre *un autre triangle* que celui de gauche (C11Egypte, modifié)

(214) I_1 *maintenant prendre *l'autre triangle* que celui de gauche (C11Egypte, modifié)

Ces exemples nous amènent à notre deuxième question : l'identification du référent. La seule particularité sémantique de *autre* – l'expression de la différence par rapport à un objet-repère – laisse en effet de côté un aspect important des emplois *in situ* des expressions d'altérité : pour isoler un référent, la **sémantique de *autre* se combine avec celle des différents déterminants**. Or, cette combinaison donne lieu à des contraintes sur le contexte qui ne sont pas les mêmes pour tous les déterminants. Comme l'observent A. Berrendonner et M.-J. Reichler-Béguelin (1996), les parcours référentiels varient en fonction de la détermination de l'expression contenant *autre* : par rapport à un objet-repère O_R, la mention d'*un autre N* impliquerait une progression des parties vers le tout, ce qui signifie que l'existence de l'ensemble est le résultat de l'interprétation. *L'autre N*, en revanche, extrait la partie complémentaire d'un ensemble fini, dont une première partie, le O_R, a déjà été extraite. La progression va donc du tout vers la partie. Cette observation fournit aussi une explication aux problèmes des corrélatives en (213) et (214) : on comprend qu'en (213), la construction du tout à partir des parties est moins sensible à l'ordre d'introduction des parties. Pour (214), en revanche, on présuppose l'existence d'un ensemble à extension connue et cela explique que l'introduction d'un des éléments *a posteriori* (*celui de gauche*, en l'occurrence) fonctionne mal.

10.2.2 Modélisation par les domaines de référence

Étant donné ces repères linguistiques ainsi que le principe que nous avons énoncé à travers notre modélisation (en particulier au chapitre 8) – chaque type d'expression référentielle (a) impose des contraintes spécifiques sur la structure de son contexte d'interprétation et (b) extrait son référent de ce contexte par une opération de restructuration particulière – nous proposons une modélisation du fonctionnement de *autre* qui repose sur les hypothèses suivantes :

(a) Contraintes sur la structure du contexte d'interprétation (DR sous-spécifié) :

Autre présuppose un contexte d'interprétation dont a déjà été extrait un objet repère O_R. Cela signifie concrètement que l'on recherche, parmi les domaines de référence disponibles dans le modèle contextuel, un **domaine partitionné et focalisé**, ayant isolé antérieurement (que ce soit sur des critères perceptifs ou linguistiques) un référent-repère ou un ensemble de référents-repère O_R (cf. Figure 66a)

Pour les expressions en *autre* à tête nominale pleine, le type *N* identifié par ce nom doit être **compatible avec le type commun des éléments du domaine** (*T*). Cette compatibilité est soumise à des conditions particulières : identité de *T* et de *N*, domination de *N* sur *T* dans une hiérarchie de types, ou possibilité de recatégorisation de *T* par *N* selon des règles spécifiques au domaine d'application.

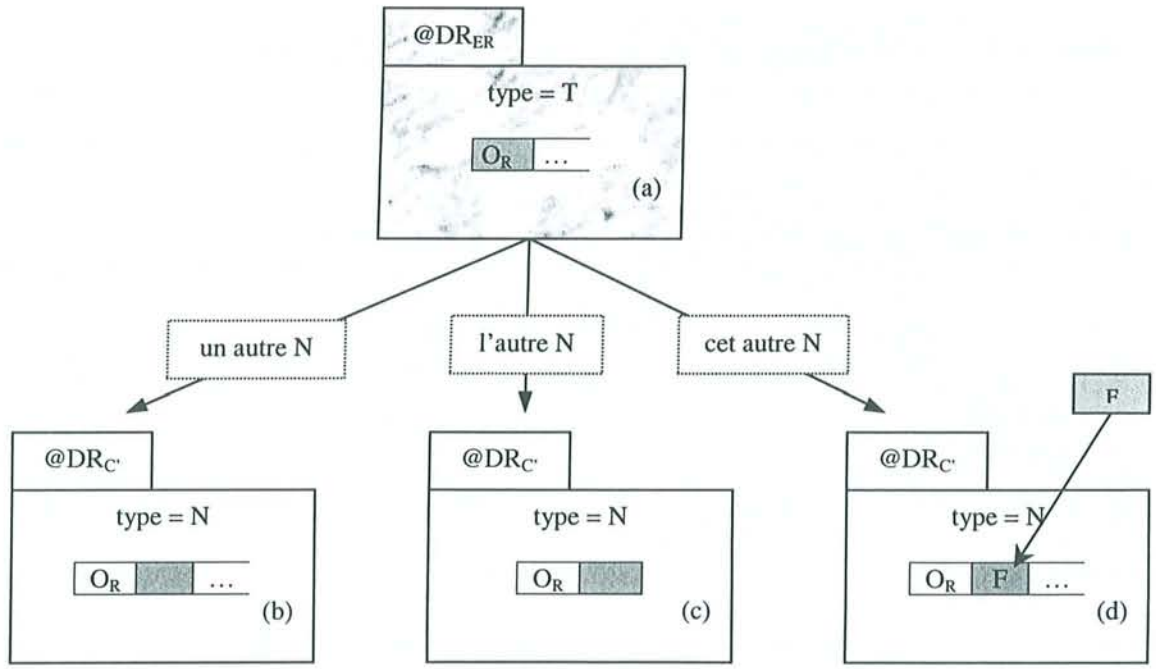


Figure 66 : Structure du domaine de référence sous-spécifié pour « autre » (a) et restructuration selon la détermination (b,c,d) ; O_R = objet-repère, F = référent préalablement focalisé

(b) Extraction du référent et opération de restructuration, en fonction de la détermination :

Déterminant indéfini : une expression de type *un autre N*, conformément au fonctionnement des indéfinis (cf. la section 8.3.1), **crée un nouveau référent** de type *N* extrait du domaine de l'objet-repère O_R, à condition que *N* et *T* (le type du domaine de l'objet-repère O_R) soient compatibles. Le nouveau référent se trouve regroupé avec O_R dans une même partition et reclassifie éventuellement le type des éléments de ce domaine par le type *N* (cf. Figure 66b).

Déterminant défini : conformément aux principes des définis (cf. la section 8.3.2), une expression de type *l'autre N* **extrait un élément d'une partition existante** et le focalise. En raison de la particularité sémantique de *autre*, cet élément doit être le complément de la partition existante du domaine de référence du référent-repère. Cela signifie en particulier que, contrairement à l'indéfini, le domaine de référence est supposé être connu dans son extension, de façon à ce que l'expression puisse en épuiser la partition. Enfin, comme pour l'indéfini, une reclassification des éléments du domaine par *N* est possible (cf. Figure 66c).

Déterminant démonstratif : une expression de type *cet autre N*, conformément aux principes d'interprétation des démonstratifs (cf. la section 8.3.3), identifie son référent d'abord sur des critères de saillance externe, traduits dans notre modélisation contextuelle par la focalisation (*F*). Ce référent est ensuite **replacé dans un nouveau domaine** de type N, à l'intérieur duquel il se regroupe avec son objet-repère (cf. Figure 66d).

10.2.3 Prédictions

A partir des hypothèses sous-jacentes à la modélisation proposée, nous sommes capable de formuler trois prédictions sur les emplois des expressions d'altérité. La suite de ce chapitre consistera à valider ces trois prédictions sur un corpus de dialogues.

- *Pour « l'autre N », l'extension du domaine d'extraction doit être connue et finie.*

Il découle de la modélisation du mode de fonctionnement des expressions d'altérité que seul l'emploi du défini *l'autre N* présuppose un domaine d'interprétation à extension finie dont l'expression puisse épuiser la partition. En effet, les énoncés b. et c. de l'exemple (215) n'ont pas les mêmes conditions d'interprétation : b. peut s'interpréter sans que l'on connaisse le nombre maximal de triangles disponibles pour la manipulation. En revanche, c. est inapproprié dans un contexte ou ce nombre serait inconnu ou supérieur à deux.

(215) a. I_1 euh maintenant il faut prendre *un grand triangle* et le mettre à gauche de l'écran
b. I_8 maintenant prendre *un autre grand triangle*
c. I_8 maintenant prendre *l'autre grand triangle* (C11Egypte, modifié)

- *Les possibilités d'isolement de l'objet-repère par un défini sont limitées.*

Le mode de donation de l'objet-repère n'est pas sans rapport avec un emploi subséquent de *autre*. De la façon dont nous modélisons le fonctionnement référentiel des définis découle en effet la prédiction que l'isolement de l'objet-repère par un défini contraint les possibilités d'un parcours référentiel en *autre*. Résumé très brièvement, un défini *le N* instaure, à l'intérieur de son domaine de référence, une partition opposant le seul élément de type *N* à des élément de type ¬*N*. Par conséquent, l'extraction d'autres éléments de cette partition au moyen d'une expression de type *l'autre N* remettrait en cause la contrainte d'unicité associée au défini (Russell, 1905), comme l'illustre l'exemple (216) :

(216) I_1 euh maintenant il faut prendre *le triangle* et le mettre à gauche de l'écran
I_8 ? maintenant prendre *l'autre triangle* (C11Egypte, modifié)

La seule possibilité de parcourir un tel domaine au moyen de *autre* est de le combiner avec un nom tête compatible avec le sur-type de N (217). C'est une sorte de « coup de force présuppositionnel » (Berrendonner et Reichler-Béguelin, 1996) qui demande tout de même à l'interlocuteur un effort supplémentaire, car il doit vérifier que ce type s'applique bien à l'objet-repère.

(217) I_1 euh maintenant il faut prendre *le triangle* et le mettre à gauche de l'écran
I_8 maintenant prendre *l'autre figure* (C11Egypte, modifié)

Enfin, l'absence d'un nom tête (*l'autre*) implique encore plus d'efforts pour l'interlocuteur, car il doit essayer d'inférer un type commun pour l'objet-repère et le référent de l'expression. Il semble que cette opération est facilitée par la mention explicite, dans le discours précédent, d'un sur-type commun (218), mais elle devient très difficile lorsqu'un sur-type doit être inféré (219) :

(218) I_1 euh maintenant il faut prendre *le grand triangle* et le mettre à gauche de l'écran
I_8 maintenant prendre *l'autre [sous-entendu : triangle]* (C11Egypte, modifié)

(219) I_1 euh maintenant il faut prendre *le triangle* et le mettre à gauche de l'écran
I_8 maintenant prendre *l'autre [sous-entendu : figure]* (C11Egypte, modifié)

- *L'emploi du démonstratif* cet autre N *n'est justifié que s'il y reclassification.*

Le principe de fonctionnement des différents marqueurs référentiels, tel que nous l'avons modélisé, prédit que l'emploi d'un démonstratif en *autre* n'est justifié que s'il y a reclassification des éléments

du domaine de référence de l'objet-repère. L'explication est la suivante : *cet autre (N)* ne crée pas de nouveau référent comme l'indéfini et n'identifie pas un référent à l'intérieur d'un domaine connu comme le défini. Il réfère, comme le pronom, à un référent déjà focalisé. Son emploi n'est donc justifié que s'il se distingue de celui du pronom. Et cette distinction se fait précisément par le changement de point de vue sur le référent, par l'intégration de celui-ci dans le domaine de référence de l'objet-repère. Or, comme pour la prédiction précédente, nous supposons que toute opération de reclassification est coûteuse d'un point de vue cognitif. L'exemple (220), tiré d'une dépêche d'actualité, illustre bien ce coût : ici, le journaliste a tout bonnement jugé nécessaire d'ajouter entre parenthèses le référent de l'expression *cet autre N.* Nous pensons alors que les locuteurs ont tendance à éviter ces constructions, en tout cas en ce qui concerne les dialogues qui nous intéressent ici.

(220) « Carlton est un précurseur du numérique terrestre. Comme nous souhaitons nous positionner comme leader dans *cette autre forme de distribution* (le numérique terrestre) à travers nos technologies et produits, on pense que [...] » a expliqué Frank Dangeard. (Dépêches Yahoo, 11/12/2000)

10.3 L'annotation (co)référentielle du corpus Ozkan

10.3.1 La ressource primaire

Ayant constaté l'absence de ressources linguistiques exploitables pour notre objectif – la validation d'une modélisation du fonctionnement référentielle de *autre* – nous avons choisi d'annoter le corpus de dialogues collecté par N. Ozkan (Ozkan, 1994). Malgré l'univers restreint de la tâche que les sujets avaient à accomplir (conception de dessins simples à partir de figures géométriques), ce corpus est extrêmement riche du point référentiel et illustre toute la variété des phénomènes référentiels dans un dialogue homme-homme multimodal combinant langage, perception visuelle et gestes : il contient tout type d'expressions référentielles et celles-ci entretiennent des liens beaucoup plus variés que simplement coréférentiels.

I1	alors ensuite il faut prendre euh faut faire une fleur
I2	alors tu prends le petit rond
M1	mmh
I3	tu vas le mettre à droite de l'écran
M2	à droite [+geste]
M3	au milieu en haut en bas ?
I4	vas-y monte encore je te dirai... voilà... stop
I5	et tu vas reprendre encore un rond
M4	un grand ou un petit ?
I6	un petit

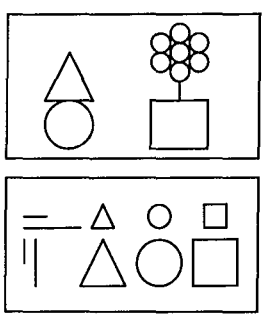


Figure 67 : Extrait du corpus Ozkan (transcription, dessin final et palette

Nous avons travaillé sur les 33 dialogues, répartis entre 7 couples de sujets et portant sur la reconstruction de 7 scènes différentes, pour lesquelles une transcription orthographique a déjà été réalisée à l'IMAG. Cette transcription est assortie de différentes annotations supplémentaires, portant sur les actions et gestes des sujets, sur la nature des actes dialogiques, sur la stratégie des interlocuteurs et sur certains implicites. A partir de ces informations, nous avons constitué de façon semi-automatique une ressource primaire, annotée en format XML. L'annotation s'inspire des recommandations de la TEI pour le codage de l'oral et des gestes. Elle permet un accès systématique aux énoncés (`<u>`[108]), aux gestes (`<kinesic>`), à l'agent de ces actions (attribut `who`)[109] et aux

[108] pour *utteranec* (énoncé)
[109] Cette information est particulièrement importante lors de la résolution de pronoms déictiques.

commentaires de différents types (`<note>`, attribut `type`). Ensuite, nous avons annoté manuellement les expressions référentielles[110], en gardant des informations sur le type de leur détermination (`<de>`[111], attribut `type`). La Figure 68 présente un extrait du corpus annoté, correspondant aux interventions I_3 et M_2 du dialogue présenté dans la Figure 67. La taille du corpus annoté est de 11500 mots, le nombre des interventions verbales s'élève à 1679 et celles-ci contiennent en tout 1344 expressions référentielles.

```
<u id="xsd:53" who="I">
        <seg> tu vas <de id="de32" det="PR_pers"> le </de> mettre à droite de <de id="de33" det="NP_def"> l'écran </de> </seg>
    <note type="speech_act"> requête </note>
</u>
<u id="xsd:54" who="M">
    <seg>à droite</seg>
    <note type="speech_act"> information répétée</note>
</u>
<kinesic id="xsd:55" who="M">
    <global_description> geste d'exécution</global_description>
    <note type="speech_act">action</note>
</kinesic>
```

Figure 68 : Annotation du corpus Ozkan

10.3.2 Examen critique des propositions pour le codage coréférentiel

L'étape suivante consiste à annoter les liens entre expressions référentielles. L'objectif est d'aboutir à un schéma d'annotation qui soit exploitable par rapport à nos préoccupations théoriques, c'est-à-dire la validation des prédictions de notre modèle sur les emplois de *autre*. Mais le souci de réutilisabilité de ces ressources linguistiques impose en même temps une annotation suffisamment générique pour être compatible avec des approches théoriques différentes de la nôtre. Nous allons donc évaluer brièvement quelques types de schémas d'annotation (co)référentielle, puis en proposer une généralisation qui remplisse ces deux critères.

- *Un schéma purement coréférentiel : le schéma MUC*

Il s'agit d'un schéma d'annotation issu d'une campagne d'évaluation de systèmes de compréhension automatique à travers sept conférences (MUC 1-7 : *Message Understanding Conference*, Chinchor et Hirschman, 1997), dont une des tâches d'évaluation était la reconnaissance automatique des expressions coréférentielles, c'est-à-dire des expressions renvoyant à une même entité extralinguistique. Par conséquent, le schéma est conçu pour identifier, dans un texte, des expressions référentielles (noms, syntagmes nominaux et pronoms) et annoter les liens coréférentiels entre celles-ci. Ne sont annotées que les expressions dont les référents sont effectivement en relation d'identité, comme en (221) :

(221) tu prends un petit carré et tu le mets au milieu (C12Route)

tu prends `<coref id= "1">` un petit carré `</coref>`
et tu `<coref id= "2" ref="1">` le `</coref>` mets au milieu... [112]

[110] Les choix concernant la nature et la délimitation des éléments à annoter sont comparables à ceux du schéma MATE ou de Popescu-Belis (1999) : annotation de toutes les expressions nominales et pronominales, à l'exclusion des expressions nominales non référentielles (impersonnels, usages attributifs), des pronoms relatifs et des pronoms ça/cela à contenu « indistinct » (Corblin, 1995).

[111] pour *discourse entity* (entité discursive)

[112] Pour une meilleure lisibilité, les exemples codés ne sont pas complets : ils ne comprennent que les balises nécessaires à notre démonstration. (omission de la détermination de <de>, des informations relatives aux énoncés, locuteurs et commentaires).

Par rapport à nos préoccupations, ce schéma strictement coréférentiel est insuffisant à deux égards : premièrement dans un exemple tel que (222), le référent de l'expression d'altérité (*un autre grand triangle*) n'est pas en relation d'identité avec le référent de l'expression « antécédente » (*un grand triangle*), appelée « objet-repère » par A. Berrendonner et M.-J. Reichler-Béguelin (1996). Se pose donc le problème du typage des liens et de l'identification d'un antécédent textuel approprié : est-ce la mention de l'objet-repère, la mention d'un ensemble dont le référent serait extrait, ou les deux ?

(222) I_1 euh maintenant il faut prendre *un grand triangle*
et le mettre à gauche de l'écran
I_8 maintenant prendre *un autre grand triangle...* (C11Egypte)

Ensuite, les données du corpus montrent que ces objets-repère ne sont pas toujours introduits discursivement. Une expression en *autre* peut également être interprétée par rapport un objet-repère disponible visuellement ou manipulé récemment (223) :

(223) I_1 oui la route alors c'est deux grands traits verticaux
I_2 [geste : manipule un premier grand trait]
I_3 on va chercher *l'autre morceau* (C7Route)

De ces observations découlent deux contraintes pour un schéma compatible avec nos objectifs : premièrement, il doit permettre d'annoter des relations avec des objets non linguistiques. Deuxièmement, il doit permettre d'annoter des liens autres que purement coréférentiels. Les recommandations du projet européen MATE (Poesio et al., 1999 ; Poesio 2000) ouvrent effectivement la voie à de tels schémas d'annotation.

- *Vers un typage des liens : les recommandations de MATE*

L'objectif premier du projet MATE (*Multilevel Annotation Tools Engineering,* Mengel et al., 2000) est d'augmenter la réutilisabilité de larges corpus à travers la définition de standards d'annotation et le développement d'outils facilitant l'acquisition et l'extraction de connaissances. Ce projet est consacré plus particulièrement au traitement de dialogues oraux et propose, en conséquence, des standards pour une annotation multi-niveaux – syntaxe, prosodie, coréférence et actes dialogiques – de telles ressources.

Dans ce cadre, le niveau coréférentiel (Poesio et al., 1999 ; Poesio, 2000) est destiné à couvrir l'annotation de relations anaphoriques. Ces relations sont définies comme une relation `<coref:link>` entre deux segments textuels `<coref:de>`. La balise `<coref:link>` contient un attribut `href` qui pointe sur l'élément anaphorique. Par ailleurs, elle comporte un élément <coref:anchor> qui pointe sur l'antécédent (224) :

(224) I_1 maintenant je vais faire *la maison*, faut mettre *un toit* dessus (C6Route)

maintenant je vais faire `<coref:de id="de1">` la maison `</coref:de>` faut mettre `<coref:de id="de2">` un toit `</coref:de>` dessus

```
<coref:link href="de2" type="poss">
   <coref:anchor href="de1"/>
</coref:link>
```

Contrairement au schéma MUC, cette relation n'est pas limitée à une relation d'identité entre les référents des deux segments. Il peut s'agir également d'une relation plus complexe, comme en (224) . A ce sujet, il est important de noter que la balise de lien `<coref:link>` contient un attribut obligatoire (`type`) afin de typer la relation entre l'antécédent et la mention subséquente. Les valeurs de cet attribut sont prises dans une liste qui inclut des relations d'identité, d'appartenance à un ensemble, d'inclusion entre ensembles, de possession ou d'instanciation.

- *Passage au niveau référentiel : Bruneseaux et Romary (1997)*

Parallèlement à l'introduction de liens autres que coréférentiels, la proposition MATE possède une deuxième caractéristique intéressante par rapport à nos objectifs : elle propose une prise en compte des liens de référence directe. Il s'agit de liens entre une expression référentielle et un objet qui n'a pas été mentionné auparavant, mais qui est accessible dans le contexte visuel. MATE suit en cela les propositions de F. Bruneseaux et L. Romary (1997) qui étendent les schémas d'annotation coréférentielle par des éléments supplémentaires permettant de coder des entités appartenant à des univers extralinguistiques. Ces entités peuvent ensuite fonctionner comme « antécédents » pour des expressions référentielles. Selon le schéma MATE, les entités non discursives sont codées par une balise `<coref:ue>`[113] et les univers référentiels par une balise `<coref:universe>`. Ce principe permet par exemple de coder la relation qu'entretient l'expression *le premier triangle* en (225) avec un des deux triangles visibles à l'écran : pour cela, on passe par le codage d'un univers d'éléments disponibles :

(225) [contexte visuel : deux triangles à l'écran]
I_1 donc tu vas commencer par prendre une petite barre
I_2 que tu vas mettre à gauche de la pointe *du premier triangle* (C8Egypte)

```
<coref :universe  id= "u1">
 <coref:ue  id ="ue1"> T1 </coref:ue>
 <coref:ue  id ="ue2"> T2 </coref:ue>
</coref :universe>

donc tu vas commencer par prendre une petite barre que tu vas mettre à gauche de la pointe
<coref:de id="de1"> du premier triangle </coref:de>

<coref:link href="de1" type="ident">
 <coref:anchor href="ue1"/>
</coref:link>
```

Le Tableau 22 résume le éléments pour le codage des relations anaphoriques selon les recommandations de MATE.

Elément	***Sémantique***	***Attributs***
<coref : de>	expression référentielle (chaîne de caractères)	id : identificateur unique
<coref:link>	relation anaphorique	href : pointeur sur mention à lier ; type : typage du lien
<coref:anchor>	antécédent discursif ou visuel	href : pointeur sur l'antécédent
<coref:universe>	situation visuelle	id : identificateur unique
<coref:ue>	éléments visibles, regroupés à l'intérieur d'un univers	id : identificateur unique

Tableau 22 : Éléments pour codage coréférentiel selon MATE

- *Quelques problèmes de codage*

Si nous avons déjà noté, à propos du schéma MUC, l'insuffisance d'un codage strictement coréférentiel, c'est-à-dire ne tenant compte que des relations d'identité, le typage des liens tel que proposé par MATE ne résout pas pour autant tous les problèmes. D'abord, l'absence d'une réelle

[113] *universe entity* (entité de l'univers)

justification théorique des valeurs proposées pour le typage des liens entraîne quelquefois des difficultés à choisir entre les douze relations de non-identité : en (226), la relation entre *oranges* et *some* est donnée comme étant de type instanciation, mais est-il réellement possible de la distinguer de `member`, `subset` ou `part_of` ?

(226) We need *oranges*. There are *some* at Corning. (MATE, 2000 : 176)

Ensuite, il y a des configurations qui ne sont pas prises en compte par ce schéma de codage. Cela concerne certaines anaphores nominales (Corblin, 1995), comme dans l'exemple (227), mais aussi les expressions d'altérité (228) :

(227) Slice *the green pepper.*
Now, remove the top of *the red one.* (Dale, 1992 : 221)

(228) I_1 euh maintenant il faut prendre *un grand triangle*
et le mettre à gauche de l'écran
I_8 maintenant prendre *un autre grand triangle* (C11Egypte)

Un deuxième problème se pose pour les expressions au pluriel. Lorsque celles-ci ne sont pas coréférentielles avec des antécédents pluriels, mais avec des antécédent dispersés (Schnedecker et Bianco, 1995) l'établissement des liens se complique. S'il est effectivement prévu de créer des antécédents complexes pour des syntagmes nominaux coordonnés, cela n'est plus possible pour des cas dans lesquels l'antécédent est construit autrement que par une coordination discursive. En (229) par exemple, *the vegetables* doit être lié à un ensemble de référents, construit dynamiquement par le discours. Or, ce type de liens ne peut pas être codé correctement.

(229) Peel and chop the onions and potato.
Crape and chop the carrots.
Slice the celery.
Melt the butter.
Add *the vegetables.* (Dale, 1992 : 237)

Enfin, un troisième problème est lié à l'introduction des univers référentiels, telle que proposée par F. Bruneseaux et L. Romary (1997) : si cela permet effectivement de coder des références à des éléments accessibles dans la situation de communication, d'autres questions se posent. L'une concerne les modalités précises de mise à jour de ces univers lorsque la situation de communication évolue (ce qui est le cas normal). Une deuxième question porte sur l'extension de ces univers : s'agit-il d'entités bien précises (objets impliqués dans une tâche limitée), ou les univers englobent-ils aussi toutes les connaissances, partagées ou non, des locuteurs ? Enfin, une troisième question, plus pratique, se pose à propos de l'exemple (230) :

(230) [contexte visuel : plusieurs figures visibles à l'écran]
maintenant il faut prendre *le grand triangle*
et *le* mettre à gauche de l'écran (C11Egypte)

Dans une telle situation, l'expression référentielle *le grand triangle* sera liée à une entité extralinguistique, disponible dans l'univers constitué par les figures à l'écran. Mais qu'en est-il pour *le* ? Entre-t-il dans une relation de coréférence avec l'expression précédente, dans une relation référentielle avec l'entité extralinguistique ou les deux ? Si on voulait, dans une optique résolument référentielle, systématiser les liens référentiels au détriment des liens coréférentiels (qui en seraient déductibles), cela supposerait un codage exhaustif du contexte extralinguistique, chose qui n'est pas envisageable de façon réaliste. Mais dans le cas contraire, on mélange deux niveaux d'annotation qui sont différents (référentiel vs. coréférentiel), ce qui mérite au moins une prise de conscience théorique.

Dans la section suivante, nous proposerons d'abord des solutions théoriques aux problèmes de typage et de l'insuffisance des liens, puis au problème de la confusion des niveaux d'annotation. Nous

montrerons ensuite comment ces solutions peuvent être mise en œuvre par l'utilisation de méta-schémas XML, solution qui a l'avantage d'être modulaire et adaptable en fonction de différentes approches théoriques

10.3.3 Vers une solution de codage cohérente et modulaire

- *Quels sont les liens pertinents pour une annotation coréférentielle ?*

Une annotation coréférentielle (à l'opposé d'une annotation référentielle) est, par définition, limitée à l'annotation de liens entre segments discursifs. Dans la section précédente, nous avons mis en évidence quatre points potentiellement problématiques lors d'une telle annotation :

> *Les liens de type `identité` sont nécessaires, mais insuffisants.*
> *Le choix entre différents liens autres que `identité` n'est pas toujours aisé.*
> *Certains exemples, dont les expressions d'altérité, ne peuvent pas être codés.*
> *Pour les expressions plurielles, un lien avec des antécédents discursivement disjoints est impossible.*

A partir de ce constat, nous proposons une solution autour des trois idées suivantes : garder un lien de type `identité`, introduire un lien d'extraction englobant toutes les relations de non-identité mentionnées auparavant et introduire un lien de codomanialité dont nous verrons qu'il est capable de couvrir les cas non pris en compte jusqu'alors (anaphores nominales, expressions d'altérité, expressions plurielles).

D'un point de vue théorique, ce triptyque repose sur une conception de l'acte référentiel en tant qu'ensemble limité d'opérations cognitives. Nous considérons que l'opération de base est une opération d'extraction (Figure 69a), isolant un référent `R` d'un contexte `C` donné. Mais cette extraction peut être « court-circuitée », dans deux cas particuliers : le premier cas est l'extraction par coréférence (Figure 69b). Il s'agit d'une ré-identification d'une entité déjà extraite. Cette ré-identification directe est en quelque sorte un « raccourci cognitif », évitant de repasser de nouveau par le domaine contextuel lorsque le référent est suffisamment accessible. Le deuxième cas est l'extraction par codomanialité (Figure 69c). Là aussi, il s'agit d'une extraction par « raccourci » : l'identification du référent `R2` se fait par rapport à un objet-repère `R1`, ce qui assure la disjonction référentielle entre l'objet-repère et le référent.

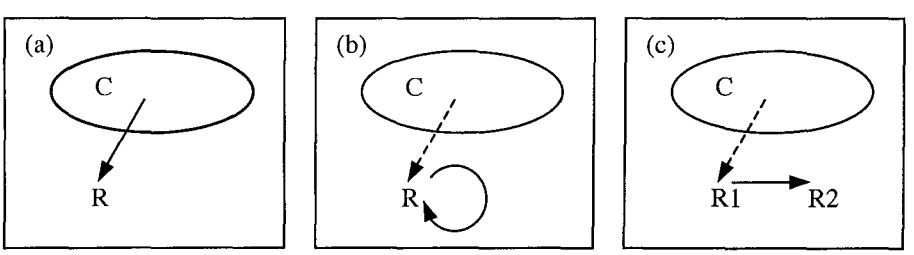


Figure 69 : L'identification référentielle comme (a) extraction simple, (b) extraction par coréférence et (c) extraction par codomanialité

Un telle conception des relations (co)référentielles a deux avantages : premièrement, elle explique les difficultés de la définition exhaustive d'une liste des relations de non-identité par le fait qu'il s'agit, à un niveau plus abstrait, d'une même opération de base, à savoir une extraction. Par ailleurs, cela n'est en aucun cas incompatible avec un choix pratique consistant à sur-typer des liens d'extraction, si l'approche théorique particulière le demande. Deuxièmement, cette position permet d'intégrer les cas problématiques – certaines anaphores nominales, expressions d'altérité et expressions plurielles – que

nous avons présentés plus haut. Ce qui caractérise ces exemples, c'est le fait que leur interprétation référentielle s'appuie sur des structurations locales du contexte. Or, ces structurations se font par des groupements locaux de référents. Notre proposition revient précisément à matérialiser les groupements par des liens de codomanialité entre certaines expressions. L'annotation de ces liens permet à la fois d'exprimer une relation de disjonction entre ces référents (anaphores nominales : *le grand triangle – le petit* ; expressions d'altérité : *un grand triangle – un autre* ; expressions ordinales : *un premier triangle – un deuxième*) et de créer des ensembles d'antécédents potentiels pour des expressions plurielles (*un premier triangle – un deuxième – les deux triangles*)[114].

En réponse à la question des liens pertinents pour une annotation référentielle au niveau discursif (annotation coréférentielle), on obtient donc le schéma du Tableau 23 : les liens possibles entre deux entités discursives `<coref:de>` peuvent être de type identité (`IDENT`), codomanialité (`CODOM`) ou extraction (`EXTRACT`).

<table>
<tr><td></td><td colspan="2">Antécédent discursif :
<coref:anchor href="de..."/></td></tr>
<tr><td rowspan="2">Expression à lier :
<coref:de></td><td>Relation d'identité :
<coref:link ...
type= "ident"></td><td>Relation de codomanialité :
<coref:link ...
type= "codom"></td></tr>
<tr><td colspan="2">Relation d'extraction :
<coref:link... type= "extract"></td></tr>
</table>

Tableau 23 : Schéma des liens référentiels pertinents entre entités discursives

Un lien d'identité s'établit entre deux entités discursives DE_1 et DE_2 dont le référent R est identique, comme en (231) :

`IDENT(DE1,DE2) ssi R(DE1) = R(DE2)` [R1]

(231) maintenant il faut prendre le grand triangle
et le mettre à gauche de l'écran (C11Egypte, modifié)

maintenant il faut prendre <de id="de1"> le grand triangle </de> et <de id="de2"> le </de> mettre à gauche de l'écran

```
<coref:link href="de2" type="ident">
    <coref:anchor href="de1"/>
</coref:link>
```

Un lien de codomanialité s'établit entre deux entités discursives DE_1 et DE_2 dont les référents R_1 et R_2 sont disjoints, à condition que l'identification de R_2 se fasse par rapport à un objet-repère R_1, comme en (232) :

$CODOM(DE_1,DE_2)$ ssi $R(DE_1) \cap R(DE_2) = \varnothing$ [R2]

(232) euh maintenant il faut prendre un grand triangle
maintenant prendre un autre grand triangle (C11Egypte)

euh maintenant il faut prendre <de id="de1"> un grand triangle </de>
maintenant prendre <de id="de2"> un autre grand triangle </de>
<coref:link href="de2" type="codom">

[114] Pour l'instant, l'annotation d'un lien entre une expression plurielle et des antécédents multiples passe par l'utilisation de plusieurs ancres <coref:anchor> à l'intérieur de l'élément <coref:link>. En pratique, le codage pourra être facilité par l'introduction d'un identificateur unique pour les groupements créés par les éléments en relation de codomanialité. Mais cela demande de coder systématiquement les relations de codomanialité, alors que nous n'avons codé, pour l'instant, seulement les liens de codomanialité relatifs aux expressions d'altérité.

```
    <coref:anchor href="de1"/>
  </coref:link>
```

Un lien d'extraction s'établit entre deux entités discursives DE_1 et DE_2, lorsque le référent de DE_1 est extrait du référent de DE_2 :

$EXTRACT(DE_1, DE_2)$ ssi $R(DE_1) \subset R(DE_2)$ [R3]

(233) maintenant je vais faire la maison
faut mettre un toit dessus (C6Route)

maintenant je vais faire <de id="de1"> la maison </de>
faut mettre <de id="de2"> un toit </de> dessus

```
<coref:link href="de2" type="extract">
     <coref:anchor href="de1"/>
</coref:link>
```

Étant donné ces définitions, il reste à déterminer l'ordre préférentiel des liens à coder, au cas où plusieurs possibilités se présentent. Cet ordre découle des relations entre les liens :

Une relation de type IDENT entre deux entités discursives DE_1 et DE_2 implique que celles-ci sont extraites du même domaine :

$IDENT(DE_1, DE_2) \Rightarrow \exists x\ EXTRACT(R(DE_1), x) \wedge EXTRACT(R(DE_2), x)$ [R4]

Une relation de type CODOM entre deux entités discursives DE_1 et DE_2 implique que celles-ci sont extraites du même domaine :

$CODOM(DE_1, DE_2) \Rightarrow \exists x\ EXTRACT(R(DE_1), x) \wedge EXTRACT(R(DE_2), x)$ [R5]

De ces relations d'implication, on déduit donc qu'en cas de conflit doivent être codés préférentiellement les liens de type COREF et CODOM, puisque les liens de type EXTRACT en découlent. L'inverse, en revanche, n'est pas vrai : de deux référents R_1 et R_2 extraits du même domaine D on ne peut conclure ni à leur identité (D : *trois triangles,* R_1 : *le triangle rouge,* R_2 : *le triangle vert*), ni à leur disjonction (D : *trois triangles,* R_1 : *le triangle rouge,* R_2 : *les triangles rouge et vert*). Par conséquent, l'exemple (234), dans lequel l'expression *une autre* (de3) peut être considérée comme étant codomaniale avec *une maison* (de2), mais aussi comme étant extraite de *deux maisons* (de1), se codera comme suit :

(234) donc au bord de cette route, il y a deux maisons,
donc une maison qui se trouve à gauche de cette route
et une autre à droite (C9Maisons)

donc au bord de cette route, il y a <de id="de1"> deux maisons </de>, donc <de id="de2">
une maison </de> qui se trouve à gauche de cette route et <de id="de3"> une autre </de> à
droite

```
<coref:link href="de2" type="extract">
     <coref:anchor href="de1"/>
</coref:link>
<coref:link href="de3" type="codom">
     <coref:anchor href="de2"/>
</coref:link>
```

Enfin, lorsqu'il y a une ambiguïté entre COREF et CODOM, la seule solution cohérente consiste à annoter les deux relations, car aucune n'est déductible de l'autre. Ainsi, en (235), l'expression *l'autre* (de4) est à la fois coréférentielle avec *une grosse pyramide* (de2) et codomaniale avec *un grand triangle* (de3) :

(235) I1 pyramides dans le désert
I2 alors une grosse pyramide
I3 [geste : I pose une grande pyramide]
I4 alors euh tu mets un grand triangle
sur la droite de l'autre (C12Egypte)

```
<de id="de1"> pyramides dans le désert </de>
alors <de id="de2"> une grosse pyramide </de>
[geste : I pose une grande pyramide]
alors euh tu mets <de id="de3"> un grand triangle </de>
sur la droite de <de id="de4"> l'autre <de>
<coref:link href="de2" type="extract">
    <coref:anchor href="de1"/>
</coref:link>

<coref:link href="de4" type="ident ">
    <coref:anchor href="de2"/>
</coref:link>

<coref:link href="de4" type="codom">
    <coref:anchor href="de3"/>
</coref:link>
```

- *Modulariser les niveaux d'annotation*

Dans la section précédente, nous avons exclusivement considéré le niveau d'annotation coréférentiel, c'est-à-dire instaurant des liens entre deux entités discursives. Or, nous avons déjà noté que cela pouvait être insuffisant, car certaines expressions ne trouvent pas leur « antécédent » dans le texte précédent, mais dans la situation de communication (236) :

(236) [contexte visuel : deux triangles à l'écran]
donc tu vas commencer par prendre une petite barre que tu
vas mettre à gauche de la pointe *du premier triangle* (C8Egypte)

Ce problème a motivé la proposition de F. Bruneseaux et L. Romary (1997), consistant à coder explicitement les entités de l'univers et à les faire fonctionner comme « antécédents ». En dehors des difficultés pratiques que pose un codage explicite des entités non discursives, nous avons noté également que cette solution, intégrée aux propositions de MATE, soulève des problèmes provoqués par une fusion de deux approches théoriques différentes : l'approche coréférentielle (lien entre deux entités discursives) et l'approche référentielle (lien entre une entité discursive et une entité non discursive).

Afin de remédier à ce problème, nous proposons de modulariser les niveaux d'annotation. Notre proposition peut être comprise comme un prolongement de l'approche adoptée par MATE. Ce projet sépare déjà différents niveaux d'annotation dialogique (syntaxe, prosodie, dialogue, coréférence), auxquels nous proposons d'ajouter un niveau référentiel. De plus, nous allons montrer comment le cadre formel des schémas XML permet de garder une cohérence globale des schémas d'annotation utilisés.

En ce qui concerne le niveau coréférentiel, nous gardons le schéma proposé dans la section précédente (balises `<coref:link>`). Par ailleurs, lorsque cela est nécessaire, nous introduisons un niveau d'annotation référentielle utilisant des balises `<ref :link>`, qui établissent des liens entre une entité discursive et une entité de l'univers. Pour des raisons pratiques, nous ne préconisons pas le codage exhaustif des univers du discours. Les entités de l'univers sont introduites uniquement lorsque

cela est nécessaires. Elles sont distinguées, pour l'instant, par un identificateur spécifique (commençant par « `ue` », pour *universe entity*).

Les types de lien pour l'annotation référentielle sont les mêmes que ceux du niveau coréférentiel (Tableau 24) : comme pour un antécédent discursif, le référent d'une expression référentielle peut se trouver en relation d'identité, de disjonction ou d'extraction par rapport à un élément de l'univers.

	Antécédent non discursif : `<ref:anchor href = "ue..."/>`	
Expression à lier : `<coref:de>`	Relation d'identité : `<ref:link... type = "ident">`	Relation decodomanialité : `<ref:link... type= "codom">`
	Relation d'extraction : `<ref:link ... type= "extract">`	

Tableau 24 : Schéma des liens référentiels pertinents entre une entité discursive et une entité non discursive

Un lien d'identité s'établit entre une entité discursive DE et une entité de l'univers UE lorsque UE est le référent R de DE, comme en (237):

$IDENT(DE,UE)$ ssi $R(DE_1) = UE$ [R6]

(237) [contexte visuel : deux triangles sur l'écran]
donc tu vas commencer par prendre une petite barre
que tu vas mettre à gauche de la pointe *du premier triangle* (C8Egypte)

donc tu vas commencer par prendre une petite barre
que tu vas mettre à gauche de la pointe
`<de id= "de1">` du premier triangle `</de>`

```
<ref:link href="de1" type="ident">
  <ref:anchor href="ue:triangle1"/>
</ref:link>
```

Un lien de codomanialité s'établit entre une entité discursive DE et une entité de l'univers UE si le référent de DE (R) et UE sont disjoints, à condition que l'identification de R se fasse par rapport à l'objet-repère UE, comme en (238) :

$CODOM(DE,UE)$ ssi $R(DE) \cap UE = \varnothing$ [R7]

(238) I_1 un petit trait de chaque coté de ...
M_1 [geste : M manipule un premier trait]
I_2 voilà et de *l'autre coté* (C7Egypte)

I_1 un petit trait de `<de id="de1">` chaque coté `</de>` de ...
M_1 [geste : M manipule un premier trait]
I_2 voilà et de `<de id="de2">` l'autre coté `</de>`

```
<ref:link href="de2" type="codom">
  <ref:anchor href="ue:trait1"/>
</ref:link>
```

Un lien d'extraction s'établit entre une entité discursive DE et une entité de l'univers UE lorsque le référent de DE est extrait du référent de UE, comme en (239) :

$EXTRACT(DE,UE)$ ssi $R(DE_1) \subset UE$ [R8]

(239) [devant un dessin « pyramides dans le désert »]
hé Nadine, tu nous fais pas dessiner un dromadaire (C12Egypte)

hé Nadine, tu nous fais pas dessiner <de id="de1"> un dromadaire </de>

```
<ref:link href="de1" type="excract">
 <ref:anchor href="ue:dessin1"/>
</ref:link>
```

En ce qui concerne les priorités d'annotation lorsque plusieurs liens référentiels sont possibles, nous adopterons, pour les mêmes raisons théoriques qu'au niveau coréférentiel, les mêmes conventions : IDENT est prioritaire sur EXTRACT, CODOM est prioritaire sur EXTRACT et IDENT et CODOM doivent être annotés tous les deux.

Enfin, pour certaines expressions, il est possible de trouver à la fois une relation coréférentielle et une relation référentielle. Dans ces configurations, nous proposons de donner la priorité à l'annotation des relations coréférentielles et ceci pour deux raisons : d'un point de vue pratique, ce choix permet d'obtenir des ressources linguistiques plus facilement réutilisables, dans la mesure où la majorité des besoins immédiats en ressources annotées se situe au niveau coréférentiel (cf. les conférences MUC). D'un point de vue théorique, ce choix se justifie par le fait que les relations référentielles sont déductibles des relations coréférentielles, alors que l'inverse n'est pas vrai :

Un lien d'identité coréférentielle entre une entité discursive DE_1 et une entité discursive DE_2 permet de déduire un lien d'identité référentielle entre DE_1 et l'unité extradiscursive UE_2 qui est le référent de DE_2 :

$$IDENT(DE_1, DE_2) \Rightarrow IDENT(DE_1, UE_2) \qquad [R9]$$

Un lien de codomanialité coréférentielle entre une entité discursive DE_1 et une entité discursive DE_2 permet de déduire un lien de codomanialité référentielle entre DE_1 et l'unité extradiscursive UE_2 qui est le référent de DE_2 :

$$CODOM(DE_1, DE_2) \Rightarrow CODOM(DE_1, UE_2) \qquad [R10]$$

Un lien d'extraction coréférentielle entre une entité discursive DE_1 et une entité discursive DE_2 permet de déduire un lien d'extraction référentielle entre DE_1 et l'unité extradiscursive UE_2 qui est le référent de DE_2 :

$$EXTRACT(DE_1, DE_2) \Rightarrow EXTRACT(DE_1, UE_2) \qquad [R11]$$

En (240), *le grand triangle de droite* (de2) peut être considéré comme extrayant son référent soit d'une entité non discursive (l'ensemble des deux triangles à l'écran), soit d'une entité discursive (*les deux triangles*, de1). En appliquant ces principes, il sera de préférence relié, par un lien coréférentiel d'extraction, à l'entité discursive.

(240) [contexte visuel : une scène comportant deux grands triangles]
I_1 ensuite prendre une grande barre horizontale
I_2 et la mettre ... il faut que ça touche les coins des deux triangles
I_3 en fait je sais pas peut être il aurait mieux valu
déplacer le grand triangle de droite (C11Egypte)

I_1 ensuite prendre une grande barre horizontale
I_2 et la mettre ... il faut que ça touche les coins <de id= "de1"> des deux triangles </de>
I_3 en fait je sais pas peut être il aurait mieux valu
déplacer <de id= "de2"> le grand triangle de droite </de>

```
<coref:link href="de2" type="extract">
    <coref:anchor href="de1"/>
</coref:link>
```

10.3.4 Mise en œuvre par des méta-schémas XML

Les sections précédentes ont été consacrées à la définition de certains principes pour l'annotation (co)référentielle de corpus : à partir de l'insuffisance des schémas existants, nous avons introduit et justifié trois types de liens, défini un ordre partiel sur ces liens et proposé de séparer formellement l'annotation coréférentielle de l'annotation référentielle, tout en donnant, en cas d'ambiguïté, la préférence à l'annotation du niveau coréférentiel. Cette section montrera comment ces principes peuvent être exprimés de façon cohérente et économique en utilisant le langage de description XML.

L'approche classique pour la définition de standards d'annotation linguistique consiste à définir des DTDs (*Document Type Definition*), pouvant être considérées comme des grammaires hors contexte décrivant la bonne formation de documents semi-structurés. Mais, étant donné que l'annotation linguistique dépend à la fois des phénomènes étudiés et des théories à valider, ces DTDs ont tendance à se multiplier. La proposition MATE en est un bon exemple : elle définit au moins une DTD (quelquefois plusieurs) pour chacun des quatre niveaux d'annotation (syntaxe, prosodie, dialogue, coréférence). Cette solution ne fait alors plus apparaître les liens qui peuvent exister entre différents niveaux d'annotation ou même entre des annotations du même niveau, destinées à valider différentes théories. Pourtant, la mise en lumière de ces relations contribue non seulement à des annotations plus cohérentes d'un point de vue théorique, mais aussi à une maintenance et mise à jour plus faciles des interfaces correspondantes.

Notre idée consiste alors à introduire des niveaux d'annotation abstraits : ceux-ci regroupent des caractéristiques communes à plusieurs schémas d'annotation, mais restent sous-spécifiés pour des propriétés spécifiques à un niveau d'annotation particulier. L'avantage de ces schémas génériques (ou méta-schémas) est de proposer une formulation à la fois plus économique et plus modulaire des principes d'annotation. En XML, la description des méta-schémas passe par l'utilisation d'éléments et de types abstraits. Par la suite, nous proposons de rendre cette idée plus concrète, en l'appliquant à l'expression des principes d'annotation (co)référentielle que nous avons définis dans les deux sections précédentes (Figure 70).

Le méta-schéma pour l'annotation (co)référentielle regroupe les informations suivantes : un document annoté est composé d'une tête optionnelle `<head>` – contenant en général des méta-données comme l'annotateur, la date et la version – et d'un corps `<body>`. A l'intérieur du corps sont admises des balises pour des entités (`<segment>`) et pour des liens entre entités (`<link>`). Le type de ces entités reste abstrait, dans la mesure où il varie en fonction du niveau d'annotation : il s'agira d'éléments discursifs (`<de>`) pour les niveaux coréférentiels strict et large et d'éléments discursifs (`<de>`) ou extradiscursifs (`<ue>`) pour le niveau référentiel. En ce qui concerne les liens, leur nature (`<coref:link>` ou `<ref:link>`), leur attribut type (`ident`, `codom`, `extract`) et la nature de l'élément sur lequel ils pointent par `href` varient également en fonction de l'instanciation. Par ailleurs, il s'agit d'éléments complexes, englobant au moins une ancre (`<anchor>`), dont la nature (`<de>` ou `<ue>`) est aussi tributaire du niveau d'annotation.

Ce schéma, qui résume tous les principes introduits dans les sections précédentes, peut être exprimé facilement en XML. A titre d'exemple, nous le montrerons sur la description de l'élément abstrait `<segment>` et son instanciation selon les niveaux d'annotation (Figure 71).

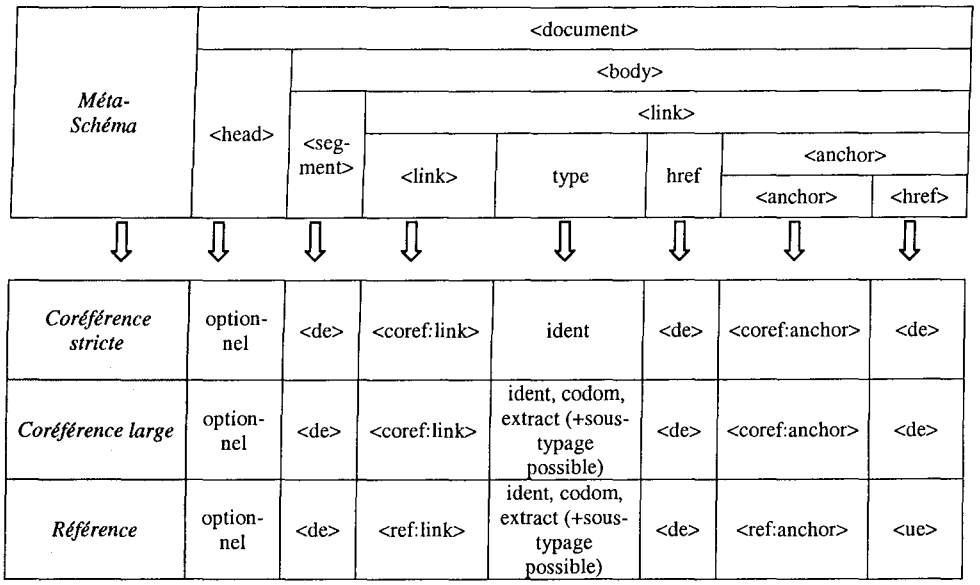


Figure 70 : Méta-schéma XML et instanciations en fonction du niveau d'annotation

Le méta-schéma introduit un élément `<body>` composé, entre autres, d'éléments abstraits de type Segment. Le type abstrait Segment, défini dans le même schéma, spécifie pour l'instant un seul attribut obligatoire qui est l'identificateur. En raison des éléments abstraits, ce méta-schéma ne peut pas être utilisé en tant que tel, mais il sert de base pour définir les segments spécifiques à chaque niveau d'annotation. Au niveau référentiel, ces segments peuvent être des entités discursives (`<de>`) ou des entités d'univers (`<ue>`). En fonction de leur nature, les éléments comportent des attributs différents : un attribut `det` portant la détermination du syntagme pour les entités discursives et un attribut `desc` de description pour les entités situationnelles. Ces caractéristiques sont ajoutées aux nouveaux types `De` et `Ue`, eux-même définis sur la base du type `Segment`, dont ils héritent les autres caractéristiques.

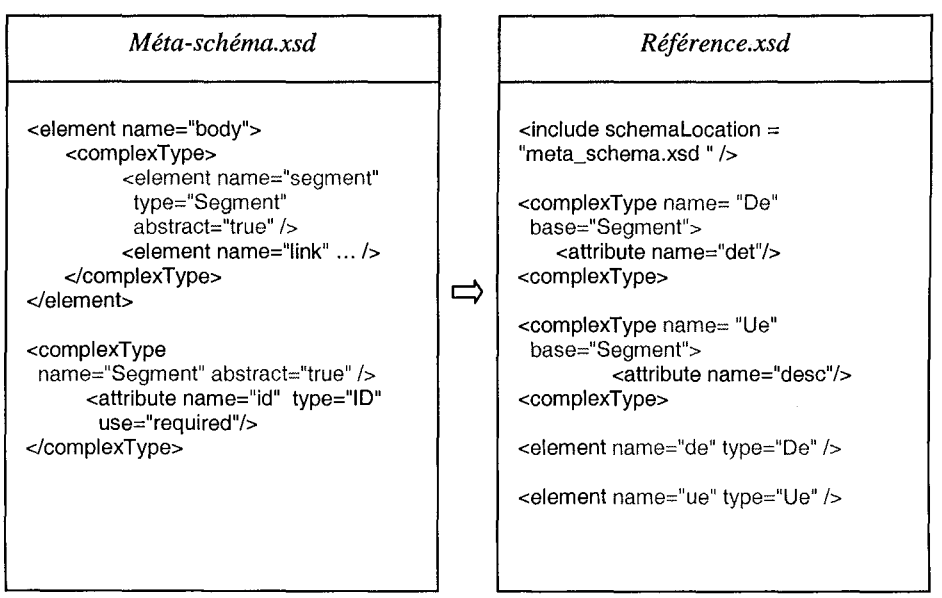


Figure 71 : Extrait du méta-schéma et instanciation par le schéma d'annotation référentielle

10.4 Validation du modèle d'interprétation référentielle de *autre*

10.4.1 De l'annotation du corpus à la construction d'un tableau synthétique

Ces réflexions sur l'annotation (co-)référentielle ainsi qu'une première interface implémentant les principes d'annotation sous forme de méta-schémas XML ont été mises à l'épreuve pour la validation de nos prédictions théoriques concernant l'interprétation de *autre*. Contrairement aux recommandations pré-existantes (MUC, MATE, Popescu-Belis (1999) pour le français), le nouveau schéma permet en effet d'annoter les expressions d'altérité en les reliant à leur objet-repère par un lien de codomanialité. L'avantage du schéma réside dans le fait qu'il n'est pas autant inapproprié pour l'annotation des liens plus « classiques », tels que définis dans les schémas MUC et MATE.

total des expressions référentielles du corpus : 1344	dont expressions d'altérité : 52			
	indéfinies : 33		définies : 19	
	non elliptiques : 22	ellipse nominale : 11	non elliptiques : 5	ellipse nominale : 14

Tableau 25 : Répartition des expressions d'altérité du corpus Ozkan

A l'issue d'une première étape de marquage des entités discursives, nous avons relevé 52 expressions d'altérité[115] sur un total de 1344 expressions référentielles du corpus (3,9 %). Le Tableau 25 donne des précisions sur leur détermination et leur forme (ellipse de la tête nominale ou non). Sur la base des études linguistiques (Berrendonner et Reichler-Béguelin, 1996 ; Schnedecker, 1998), postulant que l'interprétation de *autre* se fait relativement à un objet-repère, nous avons donc cherché à annoter ces expressions d'altérité systématiquement par un lien de codomanialité avec une autre entité discursive ou une entité extralinguistique, tout en nous tenant aux principes définis dans les sections précédentes. Les principaux problèmes d'annotation se sont posés lors du choix entre annotation coréférentielle (lien entre entités discursives) ou référentielle (lien entre une entité discursive et une entité extra-discursive).

Selon les principes définis ci-dessus, dans les 22 cas ambigus entre annotation coréférentielle ou référentielle, nous aurions dû préférer l'annotation coréférentielle. Or, cette solution ne nous a pas toujours semblé plausible : dans bon nombre d'exemples, l'objet-repère textuel se trouvait loin en amont du dialogue, alors que l'objet-repère situationnel était saillant grâce à une manipulation récente (241) :

(241) I_1 bon maintenant il faut faire <de id="de1"> les maisons </de> de chaque coté
I_2 donc y a <de id="de2"> une petite maison </de> qui est un petit carré avec un petit toit
[7 interventions discursives et gestes : construction par M de la première maison]
M_1 et il faut en faire <de id="de3"> une autre </de> alors (C7Route)

```
<ref:link href="de3" type="codom">
    <ref:anchor href="ue:maison1"/>
</ref:link>
```

Dans un exemple comme (241), où l'objet-repère textuel *une petite maison* (de2) ne se trouve pas dans le contexte discursif immédiat de l'expression d'altérité (*une autre*, de3), nous avons préféré annoter une relation de codomanialité avec une entité extra-discursive[116]. La décision sur la « proximité discursive immédiate » est particulièrement délicate : pour l'instant, nous l'avons fixée aux trois interventions discursives précédentes, mais nous estimons que le problème reste posé.

[115] Nous n'avons pas retenu dans cette étude les occurrences de *autre* dans des actes dialogiques de répétitions confirmatives. Dans des exemples comme « *Un autre triangle ? – Oui, un autre* », nous n'avons compté que la première occurrence, la seule intéressante par rapport à nos hypothèses.
[116] Nous avons néanmoins inséré une marque spécifique, de façon à pouvoir distinguer ces cas facilement.

Un deuxième problème concerne un conflit possible entre deux préférences : une préférence des relations IDENT et CODOM sur EXTRACT et une préférence du niveau coréférentiel sur le niveau référentiel. Dans certains cas, il est en effet possible d'annoter une relation d'extraction au niveau coréférentiel et une relation de codomanialité au niveau référentiel (242) :

(242) I_1 t'en mets <de id="de1"> deux </de> tu vois
M_1 [geste : pose la première des deux barres]
I_2 et tu prends <de id="de2"> l'autre </de> (C11Eglise)

```
<coref:link href="de2" type="extract">
    <coref:anchor href="de1"/>
</coref:link>
<ref:link href="de2" type="codom">
    <ref:anchor href="ue:barre1"/>
</ref:link>
```

La difficulté réside ici dans une relation codomaniale avec une entité perceptive qui « bloque » pour l'instant la déduction d'une relation d'extraction. Comme nous n'avons pas prévu d'annoter des relations entre une entité non discursive et d'autres entités (discursives ou non), nous n'aurions pas de moyen de restituer automatiquement la relation d'extraction entre de2 et de1, si nous n'annotions pas les deux relations dans de tels cas.

Enfin, un dernier problème, plus facile à résoudre, a consisté à choisir entre plusieurs objets-repère possibles (243) :

(243) I_1 Ensuite prendre <de id="de1"> un petit rond </de>enfin un petit cercle
[...]
I_2 et prendre <de id="de2"> un autre rond </de>
I_3 ensuite prendre <de id="de3"> un autre rond </de>
[...]
I_4 et je crois qu'il en faut <de id="de4"> quatre autres </de> (C11Lampe)

En (243), *un autre rond* (de3) et *quatre autres* (de4) peuvent en effet être considérés comme codomaniaux avec de2 ou de1 et de3, de2 ou de1, respectivement. Dans ces cas, classiques pour l'annotation des relations d'identité coréférentielle, il suffit de définir la relation de codomanialité comme une relation transitive, chose à laquelle nous ne voyons pas d'inconvénient.

En partant du corpus ainsi annoté, nous avons classé les occurrences des expressions d'altérité en fonction de plusieurs critères, susceptibles de valider ou d'invalider nos prédictions exposées dans la section 10.2.3. Conformément à nos hypothèses, les critères de classement sont les suivants :

la détermination de l'expression d'altérité (*un autre N, l'autre N, cet autre N*),

la détermination de l'objet-repère (défini ou non),

l'extension (connue ou non) du domaine dont l'expression d'altérité extrait son référent.

Le premier point n'a posé aucun problème. En ce qui concerne la détermination de l'objet-repère, il a fallu tenir compte des expressions à objet-repère extra-discursif, comme en (242) : nous les avons classées à part. Ce choix ne devait pas influer sur les résultats, puisque d'après nos hypothèses, la seule détermination marquée est la détermination par un article défini. Le troisième pointa été le point le plus délicat : classer les expressions d'altérité en fonction de l'extension de leur domaine d'extraction demande en effet de déterminer ce domaine, puis d'en déterminer l'extension.

Afin de déterminer le domaine d'extraction d'une expression d'altérité A, il faut d'abord identifier, à travers la relation de codomanialité, son objet-repère O_R. Suivant la règle R5, la codomanialité entre deux entités implique qu'elles soient extraites du même domaine. Il s'agit ensuite d'identifier le domaine d'extraction de l'objet-repère, en cherchant une relation EXTRACT(O_R,D). Or, cette relation ne peut être codée que pour les O_R de type discursif DE (traits pleins de la Figure 72), puisque nous n'avons annoté aucune relation à partir d'une entité extralinguistique (cf. l'exemple (242)). En ce qui concerne les O_R extradiscursifs de type UE, leur domaine d'extraction doit donc être déterminé manuellement (traits pointillés de la Figure 72), sauf dans les cas tels que l'exemple (242), où nous avons pu établir une relation d'extraction entre l'expression d'altérité et un domaine introduit discursivement (traits discontinus de la Figure 72). Dans tous les cas (O_R discursif ou non), le domaine lui-même correspond soit à une entité discursive (des syntagmes nominaux coordonnés ou au pluriel), soit à une entité saillante de la situation de communication (un ensemble de figures à l'écran, la pile virtuelle des figures disponibles sur la palette[117]). La dernière étape consiste à décider de l'extension (connue ou non) du domaine. Lorsqu'il s'agit d'un domaine discursif, nous avons déterminé son extension en fonction des informations données par le syntagme : syntagmes coordonnés (*N1 et N2*), déterminants cardinaux (*deux*) et quantificateurs cardinaux (*chaque*) ont été considérés comme indicateurs d'un domaine à extension connue ; déterminants indéfinis (*des*) et quantificateurs proportionnels (*beaucoup*) ont été considérés comme indicateurs d'un domaine à extension inconnue. Lorsqu'il s'agissait d'un domaine extradiscursif, l'ensemble des figures de la scène en construction a été considéré comme domaine à extension connue et la pile virtuelle des figures géométriques de la palette a été considérée comme domaine à extension inconnue.

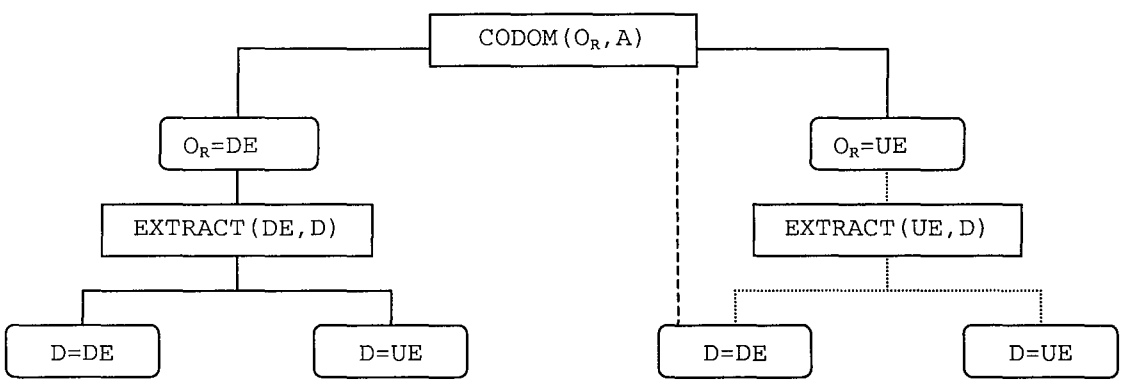


Figure 72 : Détermination et classement des domaines d'extraction pour les expressions d'altérité

En suivant ces principes, nous avons pu construire un tableau synthétique (Tableau 26), répertoriant les usages des expressions d'altérité dans le corpus Ozkan. Ceux-ci sont regroupés en fonction de la détermination de l'expression d'altérité, de la détermination de l'objet-repère (colonnes) et de l'extension connue ou non de leur domaine d'extraction (lignes). Ce tableau nous servira de base pour la validation des nos hypothèses.

[117] La palette ne contient effectivement qu'un exemplaire par figure disponible, mais cet exemplaire fonctionne en tant que type instanciable à l'infini. C'est pour cette raison que nous parlons d'une pile virtuelle de figures disponibles.

Expression d'altérité	*un autre (N)*			*l'autre (N)*			
Détermination de l'objet-repère	un	le	perceptif	un	le	ce	perceptif
Domaine à extension connue (mentionné)	2			1	1		6 (+1)
Domaine à extension connue (perceptif)				3	2	1	4
Domaine à extension non connue	6	2	23				
Total	33			19			

Tableau 26 : Classement et comptage des expressions d'altérité du corpus Ozkan

10.4.2 Interprétation et validation des prédictions

- *L'extension du domaine d'extraction doit être connue et limitée pour* « l'autre N ».

La prédiction P1, issue d'une intégration des principes de fonctionnement de *autre* dans notre modélisation de l'interprétation des descriptions définies, se trouve très clairement confirmée par les données du corpus Ozkan. On observe en effet qu'aucun emploi n'est répertorié dans la zone de croisement des expressions d'altérité définies et des domaines à extension inconnue. Un seul cas était difficile à classer (244) :

(244) I_1 donc *l'autre dessin* donc en fait c'est euh, c'est une personne assise donc sur une chaise et juste à côté d'une table (C7Homme)

Ici, I emploie l'expression *l'autre dessin* pour parler du deuxième dessin d'une série de six dessins à réaliser. Les instructions avant l'expérience renseignaient sur le nombre total de dessins à réaliser. On peut donc conclure que l'extension du domaine des dessins est fini et connu. En revanche, comme I n'emploie pas *l'autre dessin* pour écluser la partition des six dessins, cette expression peut être considérée comme problématique, dans la mesure où elle suggère qu'il n'y a que deux dessins à reproduire. On pourrait éventuellement supposer que I prend en compte uniquement le dessin réalisé préalablement et le dessin à réaliser dans l'immédiat, ce qui réduit la cardinalité du domaine effectif à deux et restituerait des conditions pour un emploi non marqué d'une expression d'altérité définie.

Mais, en plus de la confirmation de la prédiction, il est intéressant d'observer la distribution quasi-complémentaire des *autre* définis et indéfinis sur les domaines à extension connue et inconnue, respectivement. Bien qu'un indéfini ne demande pas obligatoirement un domaine à extension inconnue, les emplois sur des domaines à extension connue sont rares dans notre corpus (2 sur 33). Cela peut s'expliquer par le fait que l'indéfini *un autre (N)* est « en concurrence » avec le défini *l'autre (N)* pour désigner le dernier élément d'un domaine parcouru. Or, le défini *l'autre*, qui identifie le complément du domaine, est dans ces cas plus approprié, car il traduit explicitement la notion d'épuisement de la partition. Pour que l'indéfini *un autre (N)* puisse s'employer « hors concurrence » avec le défini *l'autre (N)*, il faudrait que le domaine possède au moins trois éléments, ce qui est plus rare dans notre corpus.

- *Les possibilités d'isolement de l'objet-repère par un défini sont limitées.*

Notre deuxième prédiction porte sur la détermination de l'objet-repère. Des principes référentiels propres à *autre* suggèrent effectivement une limitation des possibilités d'une détermination définie de l'objet-repère. Notre modélisation prédit que la seule possibilité d'employer *autre* relativement à un objet-repère *le N* est sa combinaison avec un nom tête compatible avec un sur-type de *N*. L'absence totale d'un nom tête a été considérée comme un facteur rendant l'interprétation encore plus difficile.

Cette prédiction correspond à ce que l'on peut observer dans le corpus Ozkan : sur 52 expressions d'altérité, seul cinq s'interprètent relativement à un objet-repère introduit par une expression définie. Sur ces cinq expressions, une relève de l'usage de *l'un - l'autre*, cas particulier pour lequel le statut défini ou indéfini de l'objet-repère peut être considéré comme sujet à discussion (cf. Schnedecker (1998) pour un examen détaillé). Deux expressions sont effectivement formées à partir d'un sur-type (*maison*) compatible avec le type de l'objet-repère (*église*), construction qui a pour effet une recatégorisation de l'objet-repère (245) :

(245) I_1 voilà donc l'église est faite là
I_2 et donc à sa gauche il y a *une autre maison* (C9Eglise)

Enfin, deux cas présentent une absence du nom tête. L'un des deux ressemble à l'exemple (218). L'autre est donné en (246) :

(246) I_1 [scène de construction, composée d'une maison avec un toit et une église (sans toit) à terminer]
I_2 maintenant tu prends un grand triangle
I_3 et tu fais le toit
I_4 comme *l'autre* (C11Eglise)

Un examen attentif de cet exemple fait apparaître que le statut défini de l'objet-repère (le toit) y est particulier : *le toit* est en effet défini par rapport au domaine fourni par l'église en construction, dont il est extrait. En revanche, le locuteur semble « élargir » son champ d'attention au cours de l'élaboration de l'instruction et constater la présence d'un premier toit (celui de la maison déjà construite), mettant ainsi en cause l'unicité du toit de l'église. *L'autre*, pour désigner ce premier toit, est alors employé comme si *le toit* en I_3 était isolé par un indéfini.

- *L'emploi du démonstratif* cet autre N *n'est justifié que s'il y reclassification.*

Nous avons fait l'hypothèse que l'usage d'une expression d'altérité démonstrative ne se justifie que lorsqu'elle reclassifie son référent. Or, les efforts cognitifs élevés, nécessaires à l'opération de reclassification, nous ont amenée à supposer que l'emploi de cette construction serait rare, en tout cas dans des dialogues oraux. Cette prédiction s'est entièrement vérifiée sur le corpus étudié : il n'y a aucune expression de cette catégorie.

10.5 Conclusion

La question initiale du « cercle vicieux » entre validation d'une théorie et annotation d'un corpus trouvera probablement sa réponse entre deux extrêmes. D'une part, l'annotation en l'absence totale d'une base théorique (en supposant que cela soit possible) mène non seulement à des incohérences, mais aussi à des ressources probablement non réutilisables. D'autre part, l'annotation dans le seul but de tester un système ou une théorie particulière aura le même effet – des ressources peu réutilisables, car lacunaires – et risque de fausser les phénomènes observés, car projetés dans l'annotation.

Si l'annotation de corpus peut être définie comme une « valeur ajoutée » consistant en un apport d'information aux données brutes de nature interprétative (Leech et al., 1997), l'atteinte d'un équilibre entre ces deux extrêmes passe par la prise de conscience et la maîtrise de la part interprétative projetée dans l'annotation : à partir d'un acquis théorique minimal, on crée une ressource dont l'exploitation systématique devra pouvoir faire émerger des phénomènes qui ne soient pas issus directement des principes sous-jacents à l'annotation.

En ce qui concerne l'annotation coréférentielle, il nous semble que les difficultés observées à propos des schémas existants proviennent essentiellement du fait que ces schémas ont tendance à s'inspirer des théories coréférentielles, alors que les phénomènes à traiter relèvent fondamentalement de la référence. Schématiquement, on peut caractériser une approche coréférentielle par le postulat de l'existence de relations entre unités discursives, avec un lien prototypique qui serait le lien exprimant l'identité des référents désignés. Nous pensons que ce postulat[118], qui limite le champ des annotations subséquentes à une perspective strictement intradiscursive, est précisément à l'origine d'un certain nombre de problèmes, relevés régulièrement à propos des annotations : rappelons ici les discussions autour de la transitivité des liens coréférentiels (Popescu-Belis, 1999), les difficultés liées à la nature des liens autres que purement coréférentiels (Poesio, 2000 ; Tutin et al., 2000), le problème des antécédents non linguistiques (Bruneseaux et Romary, 1997), ou plus radicalement, l'exclusion de toute une classe de marqueurs référentiels, considérés comme « trop peu consensuels » (Tutin et al., 2000).

En dépit de ce constat, nous sommes loin de préconiser pour autant une annotation strictement référentielle, c'est-à-dire créant systématiquement des liens entre des expressions référentielles et des représentations pour les objets du monde[119]. Cette solution, dont la faisabilité à grande échelle semble compromise d'avance, serait non seulement un excès dans le sens inverse, mais aussi parfaitement inadéquate par rapport aux préoccupations prédominantes de la communauté scientifique « coréférentielle », concentrée pour l'instant sur l'étude de phénomènes d'organisation textuelle à travers les anaphores. Mais si l'on admet que l'organisation textuelle d'un discours (à travers la coréférence, les anaphores ou les ellipses) n'est qu'une manifestation parmi d'autres des processus référentiels sous-jacents, il apparaît plus clairement que certains éléments d'une théorie référentielle peuvent contribuer à l'élaboration de schémas d'annotation plus adaptés. C'est ce point précis que nous avons tenté de mettre en évidence en proposant un schéma d'annotation reposant sur l'hypothèse (référentielle) selon laquelle la nature profonde d'un acte référentiel consiste en une opération d'extraction.

Il nous semble en effet que le schéma inspiré de cette hypothèse est plus à même de répondre aux problèmes courants des annotations coréférentielles, en particulier au problème de la définition des différents types de liens et de la précision d'un ordre préférentiel en cas de plusieurs possibilités d'annotation. Nous avons par ailleurs pu mettre en évidence l'existence d'un lien de codomanialité, conférant au schéma proposé la capacité de couvrir des problèmes d'ordre référentiels, tels que l'annotation des liens impliqués dans l'interprétation des expressions d'altérité. Mais au-delà des expressions d'altérité, l'intérêt de ce nouveau lien réside dans sa généricité : il a un champ d'application beaucoup plus vaste que les seules expressions d'altérité – les expressions ordinales, les ellipses nominales ou les syntagmes coordonnées – et devrait ainsi permettre d'annoter, dans un cadre cohérent, un certain nombre de phénomènes jusqu'alors traités par des conventions *ad hoc* (tels que le lien « description » défini par Tutin et al., 2000) ou tout simplement oubliés (Poesio, 2000). Enfin, la mise en œuvre du schéma d'annotation sur un corpus de dialogues nous a d'ores et déjà permis de

[118] cf. Reboul et Moeschler (1998) pour des arguments théoriques contre une « linguistique discursive » reposant sur ces postulats.
[119] ce à quoi aboutirait une mise en œuvre radicale des propositions de F. Bruneseaux et L. Romary (1997).

valider des prédictions – allant au-delà de la part interprétative projetée dans l'annotation – sur le fonctionnement de *autre*.

11 Un modèle opérationnel : implémentation et algorithmes

11.1 Introduction

Ce chapitre est consacré à la présentation d'une plate-forme dialogique dont le module de calcul référentiel s'appuie sur la modélisation proposée dans cette thèse. L'implémentation de la plate-forme a été effectuée par O. Grisvard, durant un stage post-doctoral au Laboratoire Central de Recherche de la société Thalès (ex-Thomson). Ce stage a été pour nous l'occasion d'une collaboration autour des aspects concernant le calcul référentiel : à travers cette collaboration, nous avons pu mettre notre modélisation à l'épreuve, non seulement dans une réelle plate-forme dialogique, mais aussi sur un domaine différent de celui qui nous a servi d'appui tout au long de la phase de modélisation (le monde des figures géométrique du corpus Ozkan). Notre apport à cette implémentation a consisté à participer à la spécification et à fournir le modèle de données ainsi que les algorithmes pour la gestion du contexte et le calcul référentiel proprement dit.

Le système dialogique implémenté est caractérisé par le fait que toutes ses composantes d'interprétation sémantique et pragmatique reposent sur un format de représentation commun qui sont les représentation mentales : selon O. Grisvard (2000), ce format de représentation répond en effet aux exigences d'une modélisation de la gestion du dialogue et ceci à plusieurs niveaux : au niveau dialogique (représentation des énoncés et de la structure du dialogue, par exemple des couples question/réponse), au niveau thématique (représentation des entités auxquelles il est fait référence ainsi que des contraintes d'accessibilité) et au niveau applicatif (représentation des données propres à la tâche, structurées par le modèle de la tâche ou par planification). L'implémentation de ces composantes est donc à la fois une mise en œuvre de ces principes de gestion dialogique et confirme l'adaptation du modèle des représentations mentales à une implémentation informatique.

Dans la suite, nous commenterons d'abord l'architecture générale de la plate-forme, pour nous arrêter plus longuement sur l'agent d'interprétation, responsable entre autres du calcul référentiel. L'étude du fonctionnement interne de cet agent nous donnera l'occasion de présenter la structure de données et certaines méthodes de la classe *RM.java* (l'implémentation des représentations mentales), la gestion des concepts (représentations mentales génériques) et l'interaction des modules d'interprétation pragmatique. Dans la section suivante, nous présenterons les algorithmes intervenant dans le calcul référentiel. Cette présentation suivra de très près notre modélisation du fonctionnement des expressions référentielles (cf. chapitre 8) : nous commenterons successivement les algorithmes dédiés au calcul des domaines de référence sous-spécifiés, aux procédures d'unification avec un domaine contextuel et à l'opération de restructuration, menant à l'identification du référent.

11.2 Architecture générale

11.2.1 Les agents

THOMSPEAKER est une plate-forme de dialogue oral homme-machine, destinée à effectuer différents types de tâches, dont la commande vocale d'applications et l'interrogation de bases de données. L'architecture de THOMSPEAKER repose sur un système multi-agents (Open Agent Architecture – OAA). Le principe en est le suivant : un agent offre, à partir de messages reçus, un certain nombre de services et envoie les résultats du traitement effectué également sous forme de messages. Le choix de cette architecture est justifié pour plusieurs raisons : d'abord, cela permet aux agents d'échanger, à toutes les étapes du traitement, des messages avec l'application. Ensuite, il est

prévu à terme de changer dynamiquement les grammaires de reconnaissances et d'analyse grammaticale en fonction du contexte d'interprétation. Dans la même perspective, la plate-forme étant bilingue (anglais – français), son utilisation pourra s'étendre de la commande vocale en anglais ou en français à la traduction automatique. Enfin, une architecture multi-agents permet de normaliser les interfaces entre modules de traitement et de remplacer à moindre coût un agent par un autre pour tester d'autres approches.

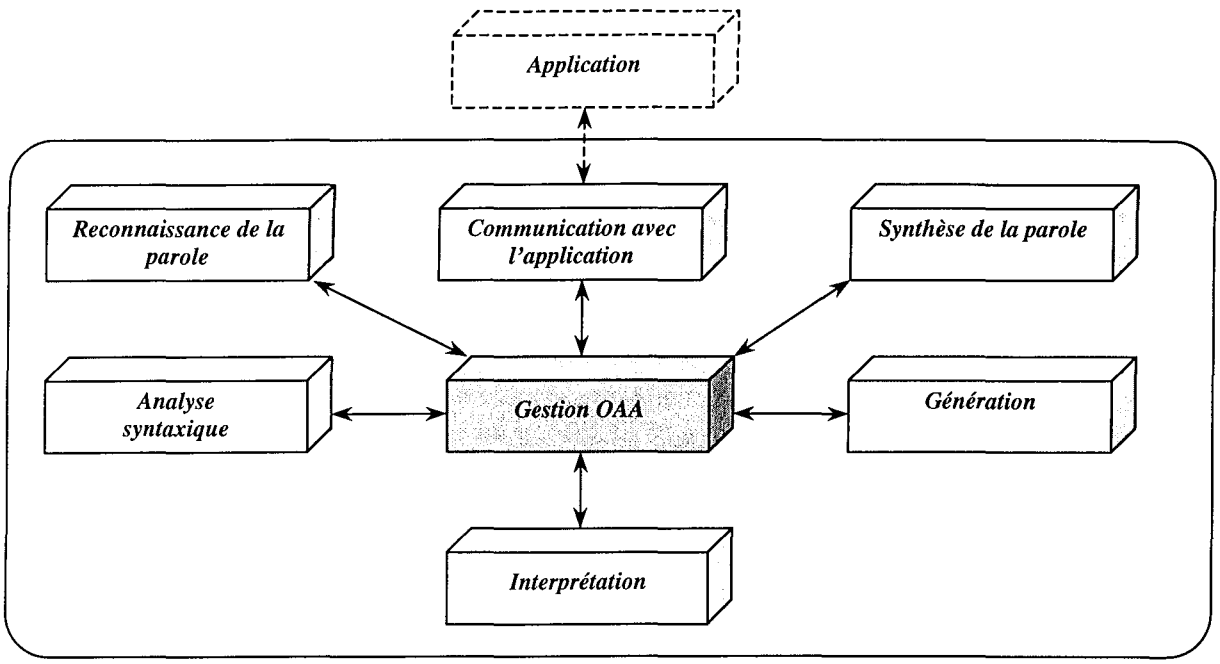


Figure 73 : Architecture de THOMSPEAKER

Actuellement, les agents qui composent le système sont les suivants (Figure 73) :

un agent de reconnaissance de la parole ;

un agent d'analyse syntaxique qui prend en entrée la chaîne de caractères reconnue pour énoncé et renvoie une analyse syntaxico-sémantique sous une forme proche d'un arbre de dépendance ;

un agent d'interprétation qui effectue l'analyse sémantique et pragmatique de l'énoncé. Cet agent à qui incombe l'interprétation des références aux objets sera décrit plus en détail ci-dessous ;

un agent de génération, qui élabore les interventions linguistiques du système. Cela concerne aussi bien le *quoi*-dire (détermination de la force illocutoire et celle du contenu du message à produire) que le *comment*-dire (élaboration de la structure morpho-syntaxique du message). Pour effectuer ce service, l'agent de génération communique en particulier avec le module responsable de la gestion du contexte, intégré à l'agent d'interprétation ;

un agent de synthèse de la parole qui produit un message oral à partir de la structure morpho-syntaxique fournie par l'agent de génération ;

un agent de communication avec l'application qui se charge de transmettre les informations entre l'application à piloter et la plate-forme de dialogue.

11.2.2 L'agent d'interprétation

Dans la suite, nous nous concentrerons sur l'agent d'interprétation, car c'est au sein de cet agent que sont effectués les traitements relatifs au calcul référentiel. Cet agent comporte lui- même plusieurs modules (analyse sémantique, interprétation pragmatique, gestion du modèle contextuel) qui ont chacun des fonctionnalités précises. Pour des raisons techniques, liées au gestionnaire multi-agents OAA ainsi qu'au langage de programmation JAVA, il n'était pas possible de les implémenter en tant qu'agents. Ils sont néanmoins séparés dans des packages JAVA bien distincts. A terme, il est prévu de les intégrer en tant qu'agents dans la plate-forme.

La particularité des différents modules de l'agent d'interprétation consiste dans le fait qu'ils échangent toutes les informations sous forme de représentations mentales (RM). Si nous nous sommes jusqu'ici limitée à une présentation des représentations mentales en tant que format de représentation pour des objets ou des ensembles d'objets (cf. le chapitre 7), les ambitions du projet CERVICAL (Reboul et al., 1998) allaient en effet au-delà : les représentation mentales sont également conçues pour pouvoir représenter des événements et des états. A partir de là, O. Grisvard a montré dans sa thèse que ce format de représentation répond aux principaux besoins d'une modélisation de la gestion du dialogue (Grisvard, 2000). C'est une telle gestion du dialogue, reposant sur un format d'échange sous forme de représentations mentales, qui est mise en œuvre pour l'agent d'interprétation.

La structuration interne de l'agent d'interprétation est celle de la Figure 74. Les différentes composantes de cet agent seront commentées ci-dessous. Dans un premier temps, nous détaillerons la gestion des représentations mentales, dans la limite de ce qui sera indispensable à la compréhension des algorithmes de calcul référentiel. Dans un deuxième temps, nous présenterons brièvement la gestion des concepts. Enfin, nous décrirons l'interaction des trois modules centraux pour l'interprétation – analyse sémantique, interprétation pragmatique et gestion du contexte.

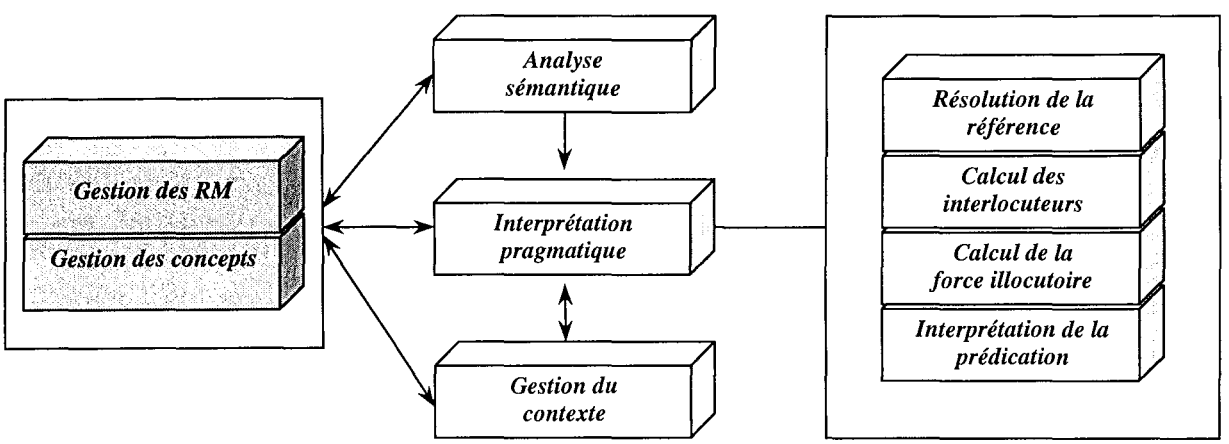


Figure 74 : Organisation de l'agent d'interprétation

11.2.3 La gestion des RM

Les trois modules centraux de l'agent interprétation – analyse sémantique, interprétation pragmatique et gestion du contexte – sont dépendants d'un module de gestion des RM qui gère la classe des représentations mentales (*RM.java*). Il s'agit d'une classe abstraite qui définit les caractéristiques communes à toutes les représentations mentales, en particulier l'entrée conceptuelle (type, cardinalité, propriétés, implication dans des événements pour les RMs objet) et l'entrée logique (partitions). Par ailleurs, une représentation mentale comporte une champ lexical qui sauvegarde les

désignations successives de l'entité ainsi qu'un champ qui pointe sur l'objet de l'application représenté. Les principales méthodes définies pour la classe des RM sont la création, la comparaison, la fusion, le groupement et l'extraction. Des classes plus spécifiques pour la représentation des éventualités (événements et états) et des objets héritent de cette classe. Par la suite, nous nous limitons à la présentation de la classe *RMObjects.java*, car c'est sur elle que s'appuient les algorithmes du calcul de la référence aux objets.

- *La classe RMObjects.java : structure des données*

Conformément aux caractéristiques des représentations mentales pour des objets (cf. le chapitre 7), les membres de la classe *RMObjecs.java* sont ceux du Tableau 27 :

Classe RMObjects.java		
Membre	*Type des données*	*Commentaire*
`id`	`string`	identificateur unique
`cat`	`GenericRM`	pointeur sur une RM générique, représentant le type de la RM spécifique (absent pour les RMs génériques)
`card`	`int`	cardinalité de l'entité représentée
`not`	`Hash-Table`	propriétés (et implication dans des éventualités)
`log`	`List(Partition)`	partitions
`lex`	`string`	désignations lexicales
`ref`	`Object`	pointeur sur l'objet de l'application

Tableau 27 : Structure des données de la classe RMObjects.java

Une RM objet comporte un identificateur unique, des indications sur son type et sa cardinalité une entrée notationnelle qui répertorie les propriétés de l'entité représentée, une entrée logique pour les partitions, une entrée lexicale et un champ qui pointe sur l'objet de l'application représenté par cette RM. Un exemple de représentation graphique pour un objet de type *PLAYER* (ayant le numéro 11, appartenant à l'équipe des « bleus » et étant impliqué dans deux éventualités – possession de ballon et mission) est donné dans la Figure 75.

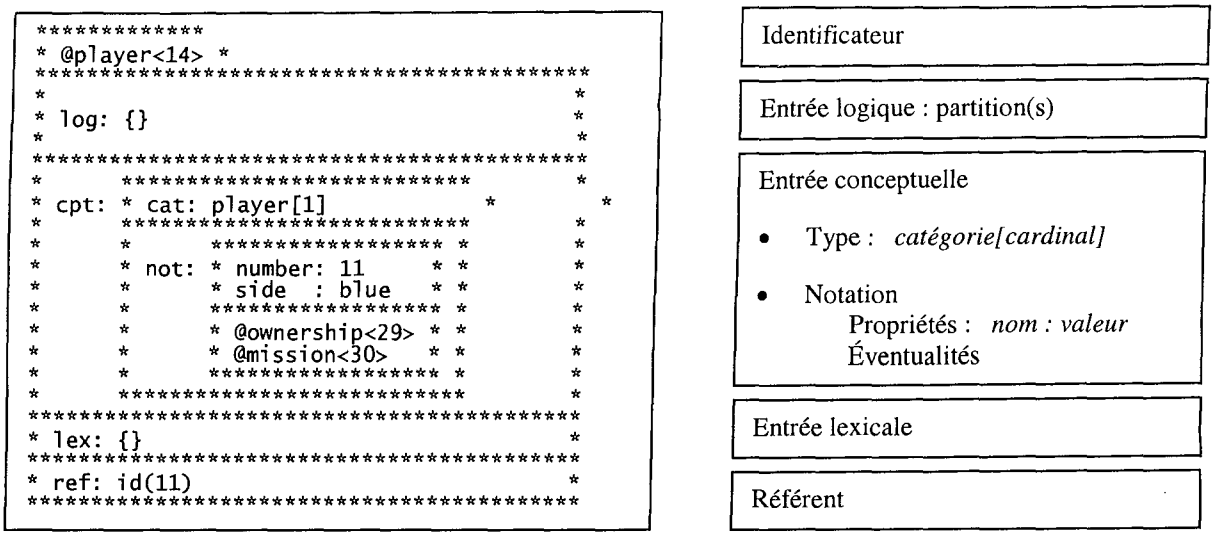


Figure 75 : RM pour un objet de type PLAYER

L'entrée logique d'une représentation mentale fournit la liste de ses partitions. Une partition est elle-même une instance de la classe *Partition.java*, pour laquelle la structure de données est détaillée dans le Tableau 28. Conformément à ce que nous avons expliqué au chapitre 7 , les partitions représentent des décompositions possibles de l'entité représentée. Chaque partition est identifiée de façon unique et comporte des indications concernant l'ordonnancement éventuel de ses éléments et sa valeur de focalisation. Par ailleurs, une partition est caractérisée par un critère de différenciation : ce critère est un attribut commun à tous les éléments de la partition et permet de distinguer un par un en fonction de la valeur prise pour cet attribut. Il peut s'agir soit de la catégorie des éléments de la partition (un point et un triangle se distingueront par leur catégorie), soit d'une propriété (deux points se distingueront par exemple par la coordonnée de leur abscisse). Dans l'implémentation présentée ici, la manipulation du critère de différenciation se fait donc par une méthode qui permet d'accéder soit à la catégorie (`getCategory()`), soit à la propriété différenciante (`getPropertyValue(Property)`) d'une RM.

Classe Partition.java		
Membre	*Type des données*	*Commentaire*
`p_id`	`string`	identificateur unique
`ord`	`Ord`	critère d'ordonnancement {0 = pas d'ordre, D = ordre discursif, T = ordre temporel, S = ordre spatial}
`focused`	`boolean`	indice de focalisation : indique si la partition contient un élément focalisé ou non {0 = pas de focus, 1 = focus }
`cd`	`Cd`	nature du critère de différenciation : {catégorie des éléments, propriété des éléments}
`items`	`List(P_Item)`	éléments d'une partition

Tableau 28 : Structure des données : la classe Partition

La Figure 76 donne un exemple de représentation mentale partitionnée pour un objet de type *HOUSE* : En supposant qu'une maison se compose, entre autres, d'une porte et d'un toit, les éléments de la partition (`@door`, `@roof`) sont distingués par un critère de différenciation qui est leur catégorie (`getCategory()`). Chaque partie de la partition est représentée par la suite « *valeur du critère de*

différenciation : [pointeur sur la RM partie, activation de la partie] ». Par ailleurs, pour cet exemple, les éléments de la partition ne sont pas ordonnés et aucun n'est focalisé.

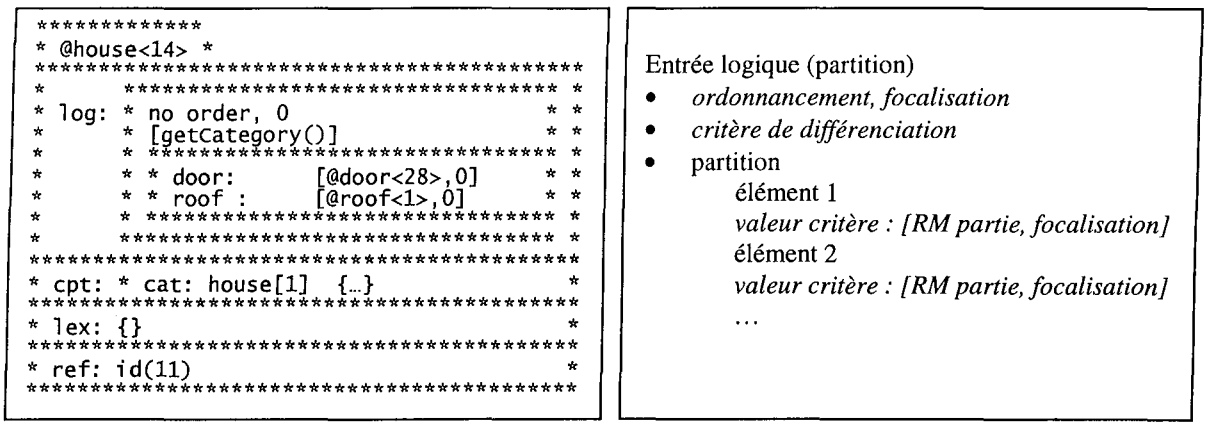


Figure 76 : RM pour un objet de type MAISON, se composant d'une porte et d'un toit

Cet exemple montre également qu'un élément d'une partition est caractérisé par sa valeur spécifique pour le critère de différenciation, par un pointeur sur la représentation de l'entité correspondante à la partie et par une valeur relative à sa focalisation. Dans les algorithmes de la suite de ce chapitre, nous utiliserons les données du Tableau 29 pour faire référence à ces propriétés.

Éléments d'une partition (P_Item)		
Propriété	*Type des données*	*Commentaire*
`cdvalue`	`string`	valeur du critère de différenciation pour cet élément de la partition
`rm`	`rm`	pointeur sur la RM représentant cet élément de la partition
`focus`	`boolean`	élément focalisé ou non : focus ∈ {0 = pas de focus, 1 = focus }

Tableau 29 : Structure des données : éléments d'une partition

- *La classe RMObjects.java : méthodes*

Sur les classes *RMObjects.java* et *Partition.java* sont définies un certain nombre de méthodes dont nous ne présentons que celles qui ont un intérêt direct pour la résolution référentielle et dont nous faisons usage dans les algorithmes présentés dans la section 11.3.

`create_RM(category)` : crée une nouvelle RM de catégorie `category` ;

`RM.card_compatible(RM)` : teste la compatibilité de la cardinalité de deux RMs et retourne un booléen ;

`RM.nb_log()` : retourne le nombre de partitions d'une RM ;

`RM.create_log(cd, order, activation)` : crée une nouvelle partition à l'intérieur d'une RM. Cette partition aura comme critère de différenciation `cd`, comme indice d'ordonnancement `order` et comme indice de focalisation `activation` ;

`RM.getProperty(RM)` : retourne une propriété qui permet de distinguer une RM d'une autre RM ;

`Partition.nb_items()` : retourne le nombre d'items d'une partition ;

`Partition.create_item(cdvalue)` : crée une nouvelle RM dans la partition . Celle-ci sera distinguée des autres éléments de la partition par la valeur de son critère de différenciation ;

`Partition.create_item(RM)` : insère une RM existante en tant que nouvel élément dans la partition. Elle sera distinguée des autres par sa focalisation ;

`Partition.focus()` : retourne l'élément focalisé d'une partition.

- *La classe RMObjects.java : algorithmes de groupement*

Parmi les méthodes de la classe *RMObjects.java*, l'opération de groupement est particulièrement importante pour le calcul référentiel. Comme nous l'avons expliqué lors de la modélisation (cf. le chapitre 7), c'est cette opération qui permet de structurer le contexte dynamiquement en domaines de référence. Conformément à ce que nous avons présenté au chapitre 7, l'opération de groupement peut être déclenchée par des facteurs linguistiques (coordination, prépositions spatiales) et perceptifs (proximité, similitude). Elle est appelée pour une RM donnée ($@RM_1$), prend en argument une deuxième RM[120] ($@RM_2$) et retourne une nouvelle RM ($@RM_3$) avec une partition (P_{new}), résultat du groupement. Pour cette nouvelle RM, l'opération groupement implique :

la création d'une nouvelle RM dont il faut calculer le type : $@RM_3$`.cat` ;

la création d'une nouvelle partition avec le calcul d'un critère de différenciation : $@RM_3$`.`P_{new}`.cd` ;

le calcul d'un ordre éventuel des éléments de la partition : $@RM_3$`.`P_{new}`.ord` ;

le calcul de la focalisation d'une partition : $@RM_3$`.`P_{new}`.focused` ;

le calcul du valeur du critère de différenciation pour les items de la partition (c'est-à-dire les RM à grouper) : $@RM_3$`.`P_{new}`.items[1,2].cdvalue` ;

le calcul éventuel de la focalisation d'un des éléments : $@RM_3$`.`P_{new}`.items[1,2].focus`.

Les algorithmes pour le groupement sont détaillés dans le Tableau 30 : les lignes spécifient le type du déclencheur de l'opération de groupement (coordination/énumération, préposition, perception) ; les colonnes détaillent le calcul des paramètres à passer aux actions de création des RM, partitions et items de partition. Les actions elles-mêmes sont répertoriées dans la dernière ligne du tableau.

Concernant le calcul du type commun des éléments à grouper, il faut distinguer quatre cas. Si les éléments sont du même type, celui-ci sera le type des éléments regroupés. Si un des éléments à grouper est un sur-type de l'autre, le type commun sera vide : ceci évite de prédire, dans des exemples comme *des sapins et des arbres* (exemple du corpus Ozkan) la possibilité d'une reprise de l'ensemble par ce sur-type, ici *les arbres*. Lorsqu'il existe un sur-type commun pour les éléments à grouper, celui-

[120] L'opération de groupement n'est pas limitée à deux RM., mais pour des raisons de simplicité, nous présentons les algorithmes seulement pour ce cas.

ci sera le type de la RM résultante. Enfin, lorsqu'il est impossible de calculer un sur-type commun (sauf un type très général comme *entité*), nous ne spécifions pas de type commun.

Le calcul du critère de différenciation et des valeurs de celui-ci pour les items à grouper dépend du calcul précédent. Dans le premier cas, le critère de différenciation sera une propriété des objets regroupés. Dans tous les autres cas, le critère de différenciation est pour l'instant la catégorie des entités groupées. Ici, l'algorithme reste à affiner pour les cas de groupement sans catégorie commune : dans le cas d'un groupement de *sapins* et d'*arbres*, retenir les catégories comme valeurs du critère de différenciation va en effet à l'encontre de la contrainte d'exclusivité mutuelle de ses valeurs à l'intérieur d'une même partition. Une solution pourrait être de travailler dans ces cas, non pas avec des catégories, mais avec des lexicalisations. Cela permettrait en effet d'éviter une reprise des *sapins* par *arbres*.

En ce qui concerne l'ordonnancement de la partition, il est dépendant du déclencheur de l'opération de groupement. Le groupement d'entités coordonnées ou énumérées donne lieu à un ordonnancement discursif correspondant à l'ordre des mentions. Un groupement perceptif peut donner lieu à un ordonnancement spatial (par exemple, en cas d'alignement des éléments) ou temporel (lors de groupements successifs d'entités introduites dans le contexte visuel), mais ici, le calcul d'ordonnancement doit s'appuyer sur des algorithmes perceptifs externes au module des RMs.

Enfin, conformément à notre souhait de rester compatible avec différents algorithmes de focalisation (Grosz et al., 1995 ; Hajicova, 1993), le calcul de la focalisation s'appuie également sur des procédures externes. Nous supposons que de telles procédures (`Focus(@RM1,@RM2,...)`) prennent un ensemble de RMs en paramètres et en retournent celle qui est focalisée. En fonction du résultat de ce calcul, un des items de la partition nouvellement créée dans $@RM_3$ sera alors focalisé.

Déclencheur	Type $@RM_3.cat$	Critère de différenciation $@RM_3.P_{new}.cd$	Ordonnancement $@RM_3.P_{new}.ord$	Focalisation $@RM_3.P_{new}.focused$	Valeurs du CD $@RM_3.P_{new}.items[1,2].cd\_value$	Focalisation des items de la partition $@RM_3.P_{new}.items[1,2].focus$

Linguistique (coordination, énumération)	if $@RM_1.cat = @RM_2.cat$) then
	$cat <= @RM_1.cat$[121]
	else if
	($@RM_1.cat = super\_cat(@RM_2.cat)$) ∨
	($@RM_2.cat = super\_cat(@RM_1.cat)$) then
	$cat <= \varnothing$
	else if
	$\exists x$ (($x = super\_cat(@RM_1.cat)$ ∧
	$(x = super\_cat(@RM_2.cat))$ then
	$cat <= x$
	else
	$cat <= \varnothing$
	end if
	end if
		if $@RM_3.cat = @RM_1.cat$ then
		$cd <= Property$
		else
		$cd <= Category$
		end if
			$ord <= D$[122]	$focused <= 0$	if $@RM_3.P_{new}.cd = Property$ then
					$cd\_value_1 <= @RM_1.getProperty(@RM_2)$
					$cd\_value_2 <= @RM_2.getProperty(@RM_1)$
					else
					$cd\_value_1 <= @RM_1.cat$
					$cd\_value_2 <= @RM_2.cat$
					end if
						$Item_f <= Focus(@RM_1, @RM_2)$

Linguistique (préposition)			$ord <= D$	$focused <= 1$		
Perceptif (proximité, similitude)			$ord <= S/0$	$focused <= 1/0$		
Action	$@RM_3 <= create\_RM(cat)$	$P_{new.} <= @RM_3.create\_log(cd, ord, focused)$			$I_{new1} <= @RM_3.P_{new}.create\_item(cd\_value_1)$ $I_{new2} <= @RM_3.P_{new}.create\_item(cd\_value_2)$	$@RM_3.P_{new.}\ Item_f.focus <= 1$

Tableau 30 : Algorithmes pour l'opération de groupement

[121] Dans les algorithmes, nous utiliserons « = » comme opérateur de comparaison et « <= » comme opérateur d'affectation.
[122] pour « D », « S », « 0 » cf. le Tableau 28, p.231

11.2.4 La gestion des concepts

En plus des outils nécessaires à la gestion des propriétés et méthodes des représentations mentales, l'agent d'interprétation s'appuie sur une description des concepts de l'application. Cette description fournit la hiérarchie des représentations mentales génériques. Par conséquent, le format de présentation des données conceptuelles suit la structure des représentations mentales : sont fournies toutes les informations conceptuelles sur les entités et les événements, en particulier leur catégorie, leur super-catégorie, leurs propriétés ainsi que leurs partitions possibles. La représentation de ces données se fait au format XML. Pour la phase de test de la plate-forme, cette ressource a été établie manuellement, mais son acquisition semi-automatique à partir du modèle de l'application est à l'étude.

La description conceptuelle de l'application remplit essentiellement trois fonctions. D'abord, elle fournit le point de départ pour l'initialisation d'une partie du modèle contextuel : le gestionnaire du contexte (Figure 74) comprend en effet les représentations des objets et éventualités effectifs de l'application à piloter. Ces représentations sont des instanciations des représentations conceptuelles ou génériques. Elles sont construites au début de chaque dialogue, à partir d'une base de données statique de l'application et des informations contenues dans la description conceptuelle. Ensuite, elle fournit un certain nombre de méthodes nécessaires au calcul référentiel. Par la suite, nous nous servirons en particulier des méthodes suivantes :

`cat(`*N*`)` : calcule à partir d'un nom *N*, le type de l'entité en question (exemple : `cat(`*triangle*`) = TRIANGLE`)[123] ;

`super_cat(`*N*`)` : calcule, à partir d'un nom *N*, le sur-type de l'entité en question (exemple : `super_cat(`*triangle*`) = FIGURE`) ;

`whole_of(`*N*`)` : calcule, à partir d'un nom *N*, le type des entités dont l'entité désignée par *N* fait partie (exemple : `whole_of(`*toit*`) = MAISON`) ;

`prop(`*P*`)` : calcule, à partir d'un modifieur *P*, la propriété correspondante (exemple : `prop(`*gauche*`) = POSITION_HORIZONTALE`) ;

`Ordering_Attributes()` : retourne l'ensemble des modifieurs pouvant exprimer des ordonnancements (`Ordering_Attributes() =` *premier, deuxième, ...*).

Enfin, la description conceptuelle fournit toutes les informations nécessaires à la création d'une nouvelle représentation lors de l'introduction d'un nouvel objet dans la tâche.

11.2.5 L'interprétation sémantique et pragmatique

En s'appuyant sur les structures de données et les méthodes gérées par les deux outils présentés précédemment, le module d'interprétation prend un charge l'interprétation sémantique et pragmatique des énoncés. En information d'entrée, il reçoit le résultat de l'analyse syntaxico-sémantique fourni par l'agent d'analyse syntaxique. En sortie, il retourne au gestionnaire OAA l'interprétation pragmatique de l'énoncé : celle-ci permet ensuite de déclencher des réponses ou actions appropriées de la part du système. Le traitement mis en œuvre pour l'interprétation pragmatique repose sur l'hypothèse que l'interprétation d'un énoncé consiste à reconnaître un certain nombre d'informations liées à l'acte de langage correspondant. Dans la modélisation de la gestion du dialogue que propose O. Grisvard dans

[123] Étant donné le nombre réduit de concepts de l'application actuellement pilotée, il n'y pas de résultats multiples pour ces fonctions. En perspective, il faudra réfléchir à la façon dont on gère les priorités dans de tels cas.

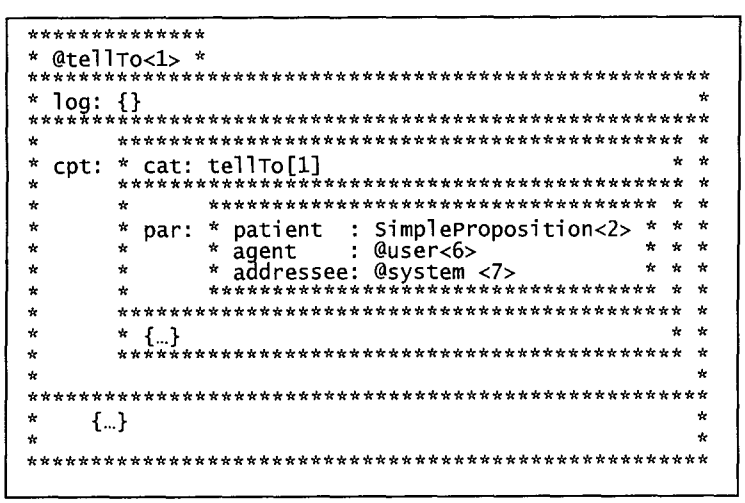
sa thèse (Grisvard, 2000), un acte de langage (*AdL*) est caractérisé par ses participants (*L* = Locuteur, *I* = Interlocuteur) et un contenu propositionnel (*CP*) sous le champ d'une force illocutoire (*FI*) :

AdL = (L, I, FI (CP)).

Pour un énoncé donné, l'interprétation pragmatique consiste donc à calculer les interlocuteurs, la force illocutoire et le contenu propositionnel. Le calcul du contenu propositionnel implique à son tour de calculer les référents des objets et éventualités mentionnés. Le résultat de ce traitement est la création d'une représentation mentale pour l'événement d'énonciation : la catégorie de cet événement correspond à la force illocutoire de l'énoncé et ses participants sont les locuteurs et le contenu propositionnel. Le contenu propositionnel renvoie lui-même à une représentation de l'éventualité correspondant au prédicat et celle-ci pointe sur les représentations des objets impliqués en tant que référents.

(247) Déplacez le triangle rouge au point alpha.

Pour un exemple tel que (247), le résultat de l'interprétation pragmatique est donc une représentation mentale de l'événement d'énonciation *@tellTo<1>* (Figure 77) aux caractéristiques suivantes: sa catégorie correspond à la force illocutoire de l'énoncé (ordre ou *tellTo*) et ses participants sont les interlocuteurs (locuteur = *@user<6>*, interlocuteur = *@system<7>*) et le contenu propositionnel (*SimpleProposition*). Le contenu propositionnel renvoie à son tour à la représentation de l'éventualité correspondant au prédicat *déplacer* : *@move<3>* (Figure 78). Celle-ci donne alors accès aux représentations des objets impliqués (triangle rouge = *@triangle<4>*, point alpha = *@point<5>*).



RM pour l'événement d'énonciation

- Catégorie (force illocutoire) : *tellTo*
- Participants :
  contenu propositionnel (*SimpleProposition*)
  locuteur
  interlocuteur

Figure 77 : Interprétation pragmatique de l'exemple (247) : représentation pour l'événement d'énonciation

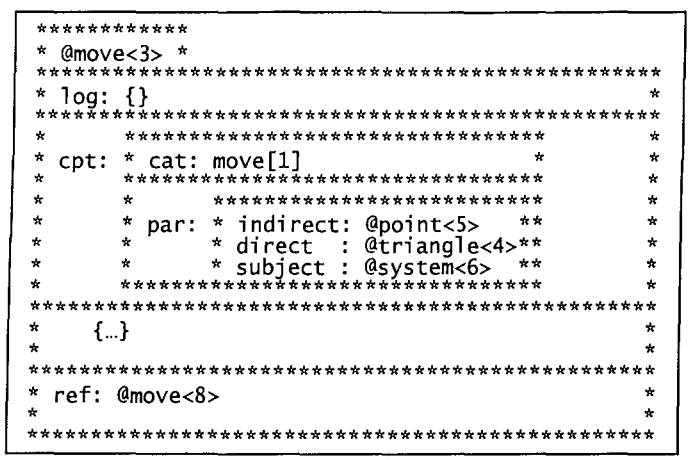


Figure 78 : Interprétation pragmatique de l'exemple (247) : représentation de l'événement correspondant au prédicat

Par la suite, nous présentons brièvement les modules impliqués dans l'interprétation pragmatique. Comme le montre la Figure 74 (page 229), il s'agit d'un module d'analyse sémantique, d'un module d'analyse pragmatique et d'un module qui gère le modèle du contexte :

Le **module d'analyse sémantique** construit, à partir de l'analyse syntaxico-sémantique fournie par l'agent d'analyse syntaxique, une représentation sous-spécifiée de l'acte de langage correspondant à l'événement d'énonciation. Cette représentation est sous-spécifiée dans la mesure où la catégorie de l'événement (la force illocutoire de l'énoncé) est une catégorie générale (@say), les locuteurs ne sont pas encore spécifiés et le contenu propositionnel renvoie à une représentation pour l'événement du prédicat, sans que le référent de cet événement ainsi que les référents des participants de cet événement soient calculés.

Le **module d'interprétation pragmatique** effectue les différents calculs nécessaires au passage de cette représentation sous-spécifiée à l'interprétation pragmatique : le calcul des interlocuteurs, le calcul de la force illocutoire, le calcul de la référence aux objets et le calcul de la référence aux événements. Le calcul des interlocuteurs se fait par recours à des règles par défaut[124], selon que le système est en phase d'interprétation ou de génération. Le calcul de la force illocutoire s'appuie sur les marqueurs prosodiques et grammaticaux du mode, sur le calcul de certains implicites et sur des données du contexte (en échange avec le module *Gestion du contexte*, présenté ci-dessous). Ce calcul aboutit à la sélection d'une des quatre catégories suivantes : affirmation, ordre, question fermée ou question ouverte. Le calcul de la référence aux objets se fait essentiellement en interaction avec le module de gestion du contexte. Il passe par la mise en œuvre de mécanismes issus de la modélisation faisant l'objet de cette thèse. Nous y reviendrons dans la deuxième partie de ce chapitre, consacrée spécifiquement aux algorithmes pour le calcul de la référence aux objets. Ensuite, le calcul de la référence à l'événement passe par l'instanciation d'un prototype d'événement correspondant au prédicat. Enfin, la dernière étape d'interprétation consiste à calculer la forme à envoyer à l'application pour que celle-ci déclenche le traitement approprié pour un énoncé donné : en fonction de la force illocutoire et du contenu propositionnel, il s'agit par exemple d'exécuter un ordre ou de répondre à une question.

[124] Ces règles prévoient des traitements adaptés en cas de vocatif ainsi que des vérifications sur la cohérence par rapport aux schémas d'action.

A tous les stades de l'interprétation pragmatique – calcul de la force illocutoire, calcul de la référence et traitement effectif des énoncés – le module d'interprétation pragmatique interagit étroitement avec **le module de gestion du contexte**. O. Grisvard a montré (Grisvard, 2000) que la tâche de ce module consiste à modéliser et à mettre à jour un modèle du contexte et ceci à trois niveaux différents : à un niveau dialogique, le gestionnaire du contexte maintient l'historique du dialogue. Cet historique est modélisé par les représentations des événements d'énonciation, structurées selon une grammaire de dialogue. A un niveau thématique, le gestionnaire du contexte maintient une représentation des contenus propositionnels des énoncés, modélisés par des représentations mentales pour les objets et les événements mentionnés. La division entre niveaux dialogique et thématique répond à la nécessité de représenter séparément les énoncés proprement dits et la structure thématique du dialogue. Enfin, à un niveau applicatif, le gestionnaire du contexte maintient des représentations pour les objets et actions effectifs de l'application. Cette deuxième division se justifie premièrement par la non-correspondance entre niveau dialogique et niveau applicatif : à titre d'exemple, si l'on se situe dans une application pour laquelle une fenêtre est grise par défaut, un énoncé unique tel que « *Crée une fenêtre rouge* » se traduit par deux actions au niveau applicatif (création et coloration). Deuxièmement, elle se justifie par la non-correspondance entre niveau thématique et niveau applicatif : la DRT (Kamp et Reyle, 1993) montre en effet que certaines constructions[125] peuvent introduire des référents de discours qui n'ont pas d'existence en dehors de l'univers discursif. Cela se traduit dans la gestion dialogique présentée ici par l'introduction, au niveau thématique, d'une représentation qui ne possède pas d'équivalent au niveau applicatif. Le calcul de la référence aux objets, qui va nous intéresser par la suite, s'effectuera donc essentiellement en interaction avec le niveau contextuel thématique.

11.3 Les algorithmes du calcul référentiel

11.3.1 Schéma algorithmique général

Suivant les principes de la modélisation présentée au chapitre 8, le schéma algorithmique général pour la résolution d'une expression référentielle suit les étapes de la Figure 79 :

1. A partir des caractéristiques propres à l'expression à interpréter (type de détermination et sémantique des constituants), construire un domaine de référence sous-spécifié DR_{ER} qui exprime les contraintes sur un domaine de référence contextuel compatible.

2. Parcourir les entités du modèle contextuel, fourni par le niveau thématique du gestionnaire contextuel de dialogue et unifier le domaine sous-spécifié DR_{ER} avec un domaine de référence contextuel DR_C approprié.

3. Appliquer l'opération de restructuration propre au type de l'expression (DR_{RS}) sur le domaine contextuel DR_C retenu. Le domaine résultant est le domaine DR_C' : à l'intérieur de celui-ci, le référent correspond à l'élément focalisé.

[125] dont les fameuses « donkey sentences » tel que « Si Pedro possède *un âne*, il *le* bat. **Il* (l'âne) a de grandes oreilles.... ».

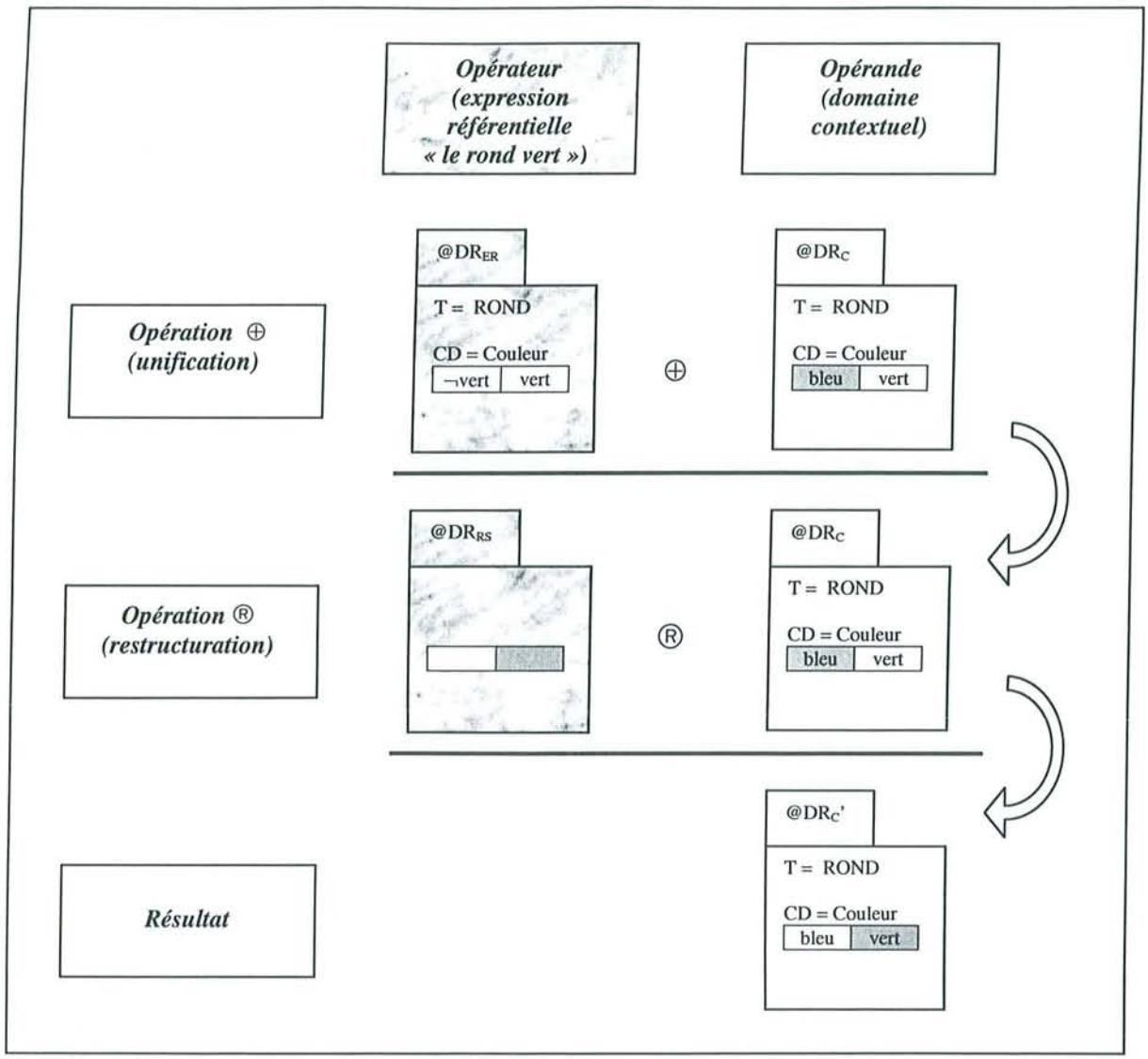


Figure 79 : L'interprétation référentielle : unification et restructuration

La suite de ce chapitre est consacrée à la présentation des algorithmes spécifiques à chacun de ces trois points. En 11.3.2, nous présenterons les algorithmes pour le calcul des domaines de référence sous-spécifiés. La section 11.3.3 porte sur les algorithmes d'unification entre le domaine sous-spécifié et un domaine contextuel. La section 11.3.4 détaille les opérations impliquées dans la restructuration du domaine contextuel sélectionné.

11.3.2 Algorithmes pour le calcul des domaines sous-spécifiés

Le calcul d'un domaine de référence sous-spécifié DR_{ER} pour une expression donnée implique le calcul de la catégorie (`@DRER.cat`), des propriétés (`@DRER.not`) et de la cardinalité (`@DRER.card`) des éléments du domaine, ainsi que le calcul des caractéristiques de la partition : son ordonnancement (`@DRER.log[#].ord`)[126], l'existence d'un élément focalisé (`@DRER.log[#].focused`), la valeur du critère de différenciation distinguant le référent des autres éléments de la partition (`@DRER.log[#].items[#].cdvalue`) et la focalisation du référent (`@DRER.log[[#].items[#].focus`). Lorsqu'une de ces données n'est pas calculée, cela signifie qu'il n'y a pas de contraintes sur sa valeur.

[126] Nous utilisons le symbole « # » en tant qu'identifiant d'un élément quelconque de la liste des partitions/ des items de la partition.

Par la suite, nous présentons les algorithmes permettant de calculer ces données pour les expressions indéfinies, définies, démonstratives et les pronoms personnels de 3ième personne, respectivement. Ces algorithmes suivent les principes introduits au chapitre 8. Par souci de clarté de la présentation, nous nous limitons à la prise en compte des pronoms et des expressions de type *un/le/ce N* (*un/le/ce triangle*), *un/le/ce N P* (*un/le/ce triangle bleu*) et *un/le P* (*un/le bleu*). Par ailleurs, nous ne présentons pas ici le calcul des contraintes de cardinalité et de genre, dans la mesure où ce n'est pas sur ces points que notre modélisation apporte du nouveau par rapport à d'autres propositions (cf. les chapitres 4 et 5). Enfin, nous laissons de côté dans un premier temps le calcul des contraintes sur l'ordonnancement de la partition et l'existence d'un élément focalisé. Ces calculs étant liés à des propriétés d'ordonnancement ou d'altérité (*deuxième, autre*), orthogonales aux différents types d'expressions (un *deuxième, le deuxième*), nous y reviendrons à la fin de cette section.

- *Expressions indéfinies*

$@DR_{ER}$	Expression indéfinie

Catégorie
$@DR_{ER}$.cat
	if ER[127] = *un N* ∨ *un N P* then
	$@DR_{ER}$.cat <= cat(*N*)
	end if

Propriétés $@DR_{ER}$.not	if ER = (*un N P* ∨ *un P*) ∧ ((prop(*P*) ≠ Ordering_attributes()) ∧ (*P* ≠ *autre*)) $@DR_{ER}$.not <= prop(*P*) end if

Tableau 31 : Algorithmes pour le calcul des DR sous-spécifiés : expressions indéfinies

Le Tableau 31 montre que le calcul du domaine sous-spécifié pour une expression indéfinie passe par le calcul de contraintes sur la catégorie et/ou les propriétés des éléments du domaine de référence. Notons au passage que ce calcul fait usage de certaines des fonctions mises à disposition par le gestionnaire des concepts (`cat(N)`, `prop(N)`), présentées dans la section 11.2.4 ci-dessus. En fonction de sa forme, une expression indéfinie s'interprète en effet dans un domaine d'éléments à catégorie déterminée (*un triangle* s'interprète dans un domaine d'éléments de type *TRIANGLE*), à catégorie et propriétés déterminées (*un triangle bleu* s'interprète dans un domaine d'éléments de type *TRIANGLE* ayant la propriété d'être bleu) ou à propriétés déterminées (*un bleu* s'interprète dans un domaine d'éléments ayant la propriété d'être bleus, sans contraintes sur la catégorie). On remarquera également qu'aucune contrainte n'est imposée sur une éventuelle partition : en cela, l'algorithme suit les principes mis en lumière par la modélisation du fonctionnement des expressions indéfinies.

- *Expressions définies*

Il apparaît à travers le Tableau 32 qu'une description définie impose obligatoirement des contraintes sur la partition de son domaine d'interprétation. Cela est conforme aux principes de modélisation propres aux expressions définies : dans tous les cas de figure (*le N, le NP, le P*), on calcule un critère de différenciation permettant d'isoler le référent des autres éléments du domaine : pour une expression telle que *le triangle*, le critère de différenciation est *CATÉGORIE* et sa valeur est *TRIANGLE*. Pour des expressions telles que *le triangle bleu* ou *le bleu*, le critère de différenciation est la propriété *COULEUR*, prenant la valeur *BLEU*. Par ailleurs, il peut y avoir une contrainte sur la catégorie

[127] L'abréviation « ER » est utilisée pour « expression référentielle ».

des éléments du domaine (*TRIANGLE* pour *le triangle bleu*). Enfin, comme nous l'avons souligné dans la modélisation du fonctionnement des définis, le futur référent d'une expression définie est de préférence un élément non focalisé de la partition.

$@DR_{ER}$	Description définie

Catégorie
$@DR_{ER}$.cat
	if ER = *le N* then
	$@DR_{ER}$.cat <= super_cat(*N*) ∨ whole_of(*N*)
	else if ER = *le N P*
	$@DR_{ER}$.cat <= cat(*N*)
	end if

Critère de différenciation
$@DR_{ER}$.log[#].cd
	if ER = *le N* then
	$@DR_{ER}$.log[#].cd <= Category
	else if ER = (*le N P* ∨ *le P*) ∧ ((prop(*P*) ≠ Ordering_attributes)) ∧ (*P* ≠ *autre*))
	$@DR_{ER}$.log[#].cd = Property
	end if

Valeur du critère de différenciation
$@DR_{ER}$.log[#].items[#].cdvalue
	if ER = *le N* then
	$@DR_{ER}$.log[#].items[#].cdvalue <= cat(*N*)
	else if ER = (*le N P* ∨ *le P*) ∧ ((prop(*P*) ≠ Ordering_attributes()) ∧ (*P* ≠ *autre*))
	$@DR_{ER}$.log[#].items[#].cdvalue <= prop(*P*)
	end if

Valeur de focalisation des items de la partition $@DR_{ER}$.log[#].items[#].focus	$@DR_{ER}$.log[#].items[#].focus <= 0 // préférence

Tableau 32 : Algorithmes pour le calcul des DR sous-spécifiés : descriptions définies

- *Pronoms personnels de 3^{ième} personne et expressions démonstratives*

$@DR_{ER}$	Pronom personnel / Expression démonstrative
Critère de différenciation $@DR_{ER}$.log[#].cd	$@DR_{ER}$.log[#].cd <= Property ∨ Category
Focalisation de la partition $@DR_{ER}$.log[#].focused	$@DR_{ER}$.log[#].focused <= 1
Valeur de focalisation des items de la partition $@DR_{ER}$.log[#].items[#].focus	$@DR_{ER}$.log[#].items[#].focus <= 1

Tableau 33 : Algorithmes pour le calcul des DR sous-spécifiés : pronom personnel et expression démonstrative

Conformément aux principes de notre modélisation, les contraintes sur le domaine d'interprétation sont les mêmes pour les pronoms personnels et les expressions démonstratives[128] : ces expressions demandent à s'interpréter dans un domaine à l'intérieur duquel existe une partition focalisée. Cela se traduit dans le Tableau 33 par les algorithmes qui calculent des contraintes sur le critère de différenciation et la focalisation. Il faut noter ici que cet algorithme implémente la version « extrémiste » du démonstratif[129] : celui-ci est en effet considéré comme reclassifiant sans contrainte un élément identifié sans ambiguïté par sa focalisation.

- *Contraintes orthogonales aux types d'expressions : l'altérité et l'ordonnancement*

Certains modifieurs – les expressions d'ordonnancement (*premier, deuxième, ...*) et les expressions d'altérité (*autre*) – n'imposent pas de valeur à un critère de différenciation, comme cela est par exemple le cas pour *bleu* qui impose une valeur particulière à un critère de différenciation *COULEUR*. Elles imposent plutôt des contraintes structurelles sur la partition. Ainsi, les expressions d'ordonnancement présupposent un domaine avec une partition dont les éléments puissent être ordonnés : *un/le/ce deuxième N*, par exemple, s'interprètent tous dans un domaine à l'intérieur duquel deux éléments entretiennent une relation d'ordre (temporel, spatial ou discursif). De la même façon, les expressions d'altérité *un/l'/cet autre N* présupposent un domaine avec un partition amorcée, c'est-à-dire une partition non épuisée contenant un élément focalisé. Ces contraintes, imposées par la sémantique des modifieurs et indépendantes du type de l'expression, se superposent aux contraintes spécifiques aux indéfinis, définis et démonstratifs. Elles se traduisent par des algorithmes concernant l'ordonnancement et la focalisation d'une partition existante. Le Tableau 34 présente l'algorithme spécifique aux expressions d'altérité : celui-ci précise qu'une expression en *autre* cherche à s'interpréter dans un domaine à cardinalité > 1 et possédant une partition focalisée. En superposant ces contraintes à celles du défini, on obtient, pour un exemple tel que *l'autre triangle* le domaine sous-spécifié de la Figure 80 : *l'autre triangle* nécessite un domaine de référence à cardinalité > 1, dont les éléments sont de catégorie *TRIANGLE* et contenant une partition avec un élément focalisé.

$@DR_{ER}$	**Expressions d'altérité**

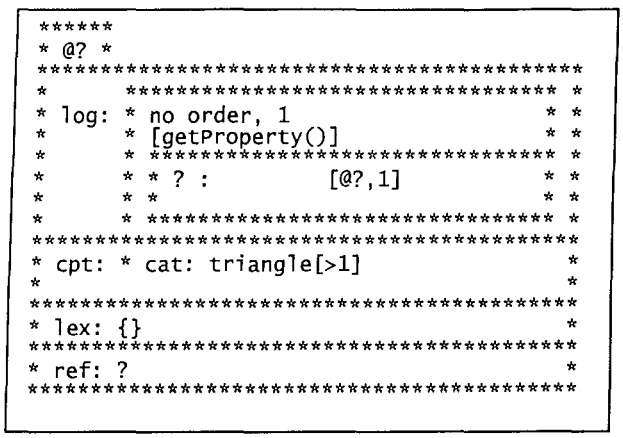
Cardinalité du domaine
$@DR_{ER}$.card
	if *P* = « *autre* » then
	$@DR_{ER}$.card <= (>1)
	end if

Focalisation de la partition
$@DR_{ER}$.log[#].focused
	if *P* = « *autre* » then
	$@DR_{ER}$.log[#].focused <= 1
	end if

Valeur de focalisation des items de la partition $@DR_{ER}$.log[#].items[#].focus	$@DR_{ER}$.log[#].items[#].focus <= 0

Tableau 34 : Algorithme pour le calcul des DR sous-spécifiés : les expressions d'altérité

[128] Dans le modèle proposé, les deux types d'expressions se distinguent seulement par leur opération de restructuration.
[129] voir à ce sujet la discussion présentée dans la section 9.4.1.



Domaine de référence sous-spécifié

- Entrée logique
  pas de contrainte d'ordonnancement
  partition focalisée (1)
  partition contenant au moins :
  un *item de type triangle, focalisé*

- Entrée conceptuelle
  catégorie : triangle
  cardinalité > 1

Figure 80 : Domaine de référence sous-spécifié pour « l'autre triangle »

11.3.3 Algorithmes de parcours et d'unification

Une fois le domaine de référence sous-spécifié calculé, la deuxième étape du schéma algorithmique général consiste à parcourir les entités du modèle contextuel afin d'unifier le domaine sous-spécifié DR_{ER} avec un domaine de référence DR_C approprié. Le parcours des domaines du contexte est effectué par un algorithme de parcours qui, lui, teste pour chaque élément parcouru la possibilité d'unification.

- *Algorithmes de parcours*

L'objectif des algorithmes de parcours est de fixer l'ordre dans lequel les domaines de référence du modèle contextuel seront parcourus.

Une première possibilité consiste à s'appuyer sur l'ordre de création des domaines. L'historique du dialogue consiste, dans ce cas, en une pile de domaines, ordonnés selon la récence de leur création, sachant qu'un domaine peut être créé lors de l'introduction discursive ou perceptive d'un nouveau référent ou lors de la présence d'un des déclencheurs de l'opération de groupement (cf. 11.2.3). C'est cette solution qui est actuellement implémentée dans THOMSPEAKER : les domaines créés correspondent aux groupements des participants d'un même événement. La pile de l'historique du dialogue est donc une liste de RM événementielles, donnant accès aux objets impliqués dans les événements représentés. Cette pile est parcourue en commençant par le dernier élément, c'est-à-dire le domaine le plus récent ;

Pour aller plus loin, on pourrait réfléchir à une structuration des domaines de référence autrement que sous forme de pile. Une possibilité serait de se reporter aux travaux présentés au chapitre 4, s'appuyant sur la structure discursive pour former des espaces de recherche sous forme arborescente (cf. en particulier la S-DRT, Asher, 1993). En faisant l'hypothèse que les domaines de référence puissent représenter non seulement des objets, mais aussi des énoncés ou complexes d'énoncés (Reboul et al., 1997 ; Grisvard, 2000), l'historique serait alors un arbre, dont les nœuds seraient des domaines de référence. Cependant, comme nous l'avons noté dans la synthèse consacrée à ces travaux (cf. la section 4.4), cela demande d'abord de disposer d'algorithmes supplémentaires pour le calcul des relations entre segments. Par ailleurs, il faut adapter les approches discursives au dialogue, ce qui implique en particulier de tenir compte des relations spécifiquement dialogique, comme des couples question–réponse. Enfin, le parcours de ces arbres selon le principe de la frontière droite (Webber, 1991) comporte le risque de rendre inaccessible des informations nécessaires au calcul référentiel (cf. 4.4.2) ;

Enfin, un prolongement de notre travail consisterait à se demander si la structure interne des domaines de référence ne peut pas être mise au profit de l'organisation du modèle contextuel. Un domaine de référence donne en effet accès, à travers ses partitions, à d'autres domaines. L'organisation contextuelle des domaines se ferait alors sous forme arborescente, mais contrairement à la proposition précédente, ces arbres ne seraient pas conçus uniquement en fonction de relations discursives ou dialogiques, mais plus généralement en fonction de critères liés à leur (dé-)composition(s) possible(s). L'avantage d'une telle conception est de rester compatible avec les propositions de structuration sur critères discursifs (en les considérant comme un cas particulier de groupement)[130], tout en maintenant un cadre unifié de représentation sous forme de domaines. Un exemple d'une telle structure est reproduite dans la Figure 81. Cette figure, reprise du chapitre 9, représente un historique du dialogue composé d'un domaine de formes géométriques, lui-même partitionné selon le type des formes et ainsi de suite. Un des domaines, le dernier ayant servi à une extraction (@gt)[131], est activé. Or, prendre cette structure comme base pour un algorithme de parcours demande d'instaurer un ordre entre les domaines. Dans une perspective mémorielle, cet ordre devrait s'appuyer sur l'activation propre à chaque domaine. Le calcul du taux d'activation, réitéré pour chaque domaine après chaque mise à jour contextuelle, demande alors la prise en compte de facteurs multiples, dont l'activation du domaine en question par une opération d'extraction, son lien avec des domaines précédemment activé et le facteur d'oubli. Ces questions doivent encore faire l'étude d'un examen détaillé et d'une confirmation expérimentale des heuristiques adoptées.

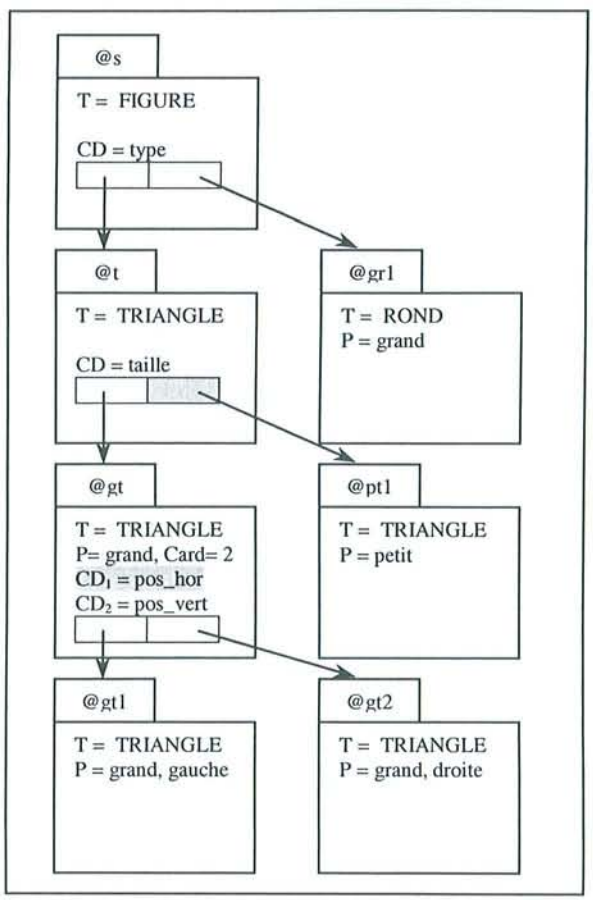


Figure 81 : Exemple de structuration contextuelle selon la composition des domaines

[130] Cette proposition rejoint un des arguments en faveur des domaines de référence. (cf. la section 6.3.5 – « Arguments discursifs : des domaines de référence aux structures du discours ? »)

[131] Le dernier domaine ayant servi à une extraction est reconnaissable, dans notre représentation graphique, par le sur-lignage du critère de différenciation utilisé.

- *Algorithme d'unification*

$@DR_{ER}$ / $@DR_C$	Contraintes d'unification

Compatibilité des catégories
`@DR.cat`
	if ($@DR_{ER}$.cat = $@DR_C$.cat) ∨ ($@DR_{ER}$.cat = *)[132] then
	cat_compatible
	end if

Compatibilité des cardinalités
`@DR.not`
	if ($@DR_{ER}$.card = *) ∨ ($@DR_{ER}$.card_compatible_with($@DR_C$.) then
	card_compatible
	end if

Compatibilité des propriétés
`@DR.not`
	if ($@DR_{ER}$.not ⊂ $@DR_C$.not) ∨ ($@DR_{ER}$.not = *) then
	not_compatible
	end if

Compatibilité des partitions,
`@DR.log`
du critère de différenciation,
`@DR.log[?].cd`
des valeurs du critère de différenciation,
`@DR.log[?].items[?].cdvalue`
de l'ordonnancement,
`@DR.log[?].cd`
de la focalisation
`@DR.log[?].cd`
	if (¬∃ $@DR_{ER}$.log) then
	log_compatible(_ , _)
	else if (∃ $@DR_{ER}$.log ∧ ∃ $@DR_C$.log) then
	n <= DR_C.nb_log()
	i <= 1
	until ((i > n) ∨ (log_compatible (i , _)))
	log <= $@DR_C$.log[i]
	if $@DR_{ER}$.log[#].cd = * then
	if (($@DR_{ER}$.log[#].ord = log.ord) ∨ ($@DR_{ER}$.log[#].ord = *)) ∧
	(($@DR_{ER}$.log[#].focused = log.focused) ∨ ($@DR_{ER}$.log[#].focused = *))
	then
	log_compatible(i, _)
	end if
	else if ($@DR_{ER}$.log[#].cd = log.cd) then
	if (($@DR_{ER}$.log[#].ord = log.ord) ∨ ($@DR_{ER}$.log[#].ord = *)) ∧
	(($@DR_{ER}$.log[#].focused = log.focused) ∨ ($@DR_{ER}$.log[#].focused = *))
	then
	m <= log.nb_items()
	j <= 1
	until (j >m) ∨ (log_compatible(i,j)))
	item <= log[i].items[m]
	if $@DR_{ER}$.log[#].items[#].cdvalue = item.cdvalue then
	log_compatible(i,j)
	end if
	next j
	end if
	end if
	next i
	end if

Tableau 35 : Critères de compatibilité entre un DR sous-spécifié et un DR contextuel

[132] L'étoile « * » est utilisée lorsque le champ correspondant n'est pas soumis à des contraintes.

Pour chaque domaine parcouru, l'algorithme d'unification est appelé afin de tester si le domaine contextuel est compatible avec les contraintes imposées par le domaine sous-spécifié. Le test s'effectue sur chacun des champs des deux domaines à unifier – la catégorie, la cardinalité, les propriétés et les partitions – selon les modalités détaillées dans le Tableau 35. En règle générale, un domaine sous-spécifié est compatible avec un domaine contextuel lorsque les informations qu'il contient sont égales à ou moins précises que celles contenues dans le domaine contextuel. Cela se traduit dans les algorithmes par le fait qu'un domaine sous-spécifié sans précision sur la catégorie, les propriétés, la cardinalité ou la partition reste compatible avec n'importe quel domaine contextuel. En cas d'existence d'une partition à l'intérieur du domaine sous-spécifié, l'algorithme parcourt successivement les partitions du domaine contextuel, jusqu'à ce qu'il en trouve une pour laquelle l'ordonnancement, la structure focale de la partition, le critère de différenciation ainsi que la valeur du critère de différenciation conviennent. Pour l'instant, l'ordre de parcours des partitions se fait selon l'ordre de leur création. Néanmoins, dans une version plus sophistiquée, l'ordre de parcours devrait suivre l'ordre décroissant de l'activation des partitions, afin de tester la compatibilité d'abord sur la partition la plus active du domaine contextuel. En cas de compatibilité des partitions, l'algorithme retourne non seulement un indice positif sur la compatibilité, mais aussi les indices de la partition (`i`) et de l'élément de la partition (`j`) pour laquelle la compatibilité a été calculée. Ces indices servent en effet dans la suite pour l'opération de restructuration. Si le test de compatibilité réussit pour tous les champs domaniaux, les deux domaines DR_{ER} et DR_C sont unifiés et le domaine contextuel DR_C (accompagné des indices `i` et `j`) est retenu comme domaine de référence pour l'expression à interpréter. En cas d'échec, le parcours du contexte reprend la main.

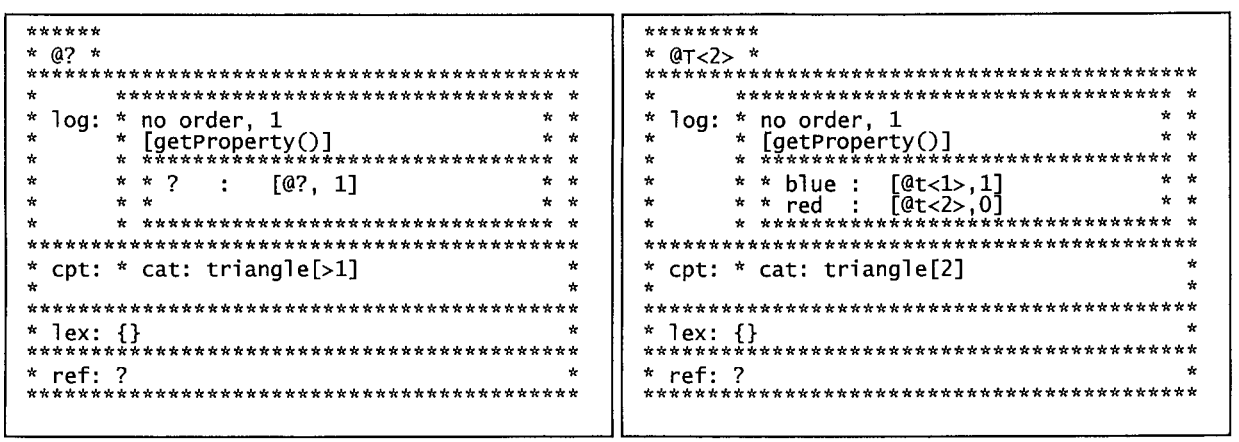


Figure 82 : Domaine sous-spécifié pour « l'autre triangle » et domaine contextuel compatible

La Figure 82 donne un exemple de domaines compatibles : elle reprend le domaine sous-spécifié pour l'expression *l'autre triangle* et le compare à un domaine contextuel `@T<2>` pour un ensemble de deux triangles, se distinguant par leur couleur et leur valeur de focalisation. Selon les algorithmes de compatibilité, les caractéristiques de ce domaine sont compatibles avec celui du domaine sous-spécifié. Par conséquent, les deux domaines sont unifiés et le domaine contextuel `@T<2>` est retenu comme domaine de référence pour l'expression à interpréter. Les indices `i` et `j` de la partition et de l'élément compatible désignent respectivement la seule partition disponible (`log`) et aucun élément particulier (*autre* n'identifie pas son référent à travers une valeur de différenciation particulière).

11.3.4 Algorithmes pour la restructuration du domaine de référence

Après la construction d'un domaine sous-spécifié et l'unification de celui-ci avec un domaine contextuel, la troisième phase du schéma algorithmique prévoit la restructuration du domaine

contextuel retenu : cette restructuration remet à jour le modèle contextuel, ce qui a pour « effet de bord » l'identification du référent de l'expression à interpréter. Conformément à notre modélisation, l'opération de restructuration se fait en fonction du type de la détermination et de la sémantique de l'expression à interpréter : comme pour la construction des domaines sous-spécifiées, nous laissons de côté, dans un premier temps, les spécificités liées à la sémantique des expressions ordinales et d'altérité pour présenter d'abord les algorithmes de restructuration en fonction du type de détermination.

	Restructuration du $@DR_c$ en $@DR_{c'}$
Indéfini :	$@DR_{c'}$ <= $@DR_c$, avec :
Création d'une nouvelle partition	P_{new} <= $@DR_{C'}$.create_log(Property,D,1)
Extraction d'un nouvel élément de cette partition	I_{new} <= $@DR_{C'}$.log[P_{new}].create_item(prop(*predicate*))
Focalisation de l'élément extrait	$@DR_{C'}$.log[P_{new}].items[I_{new}].focus <= 1
Identification du référent	Referent <= $@DR_{C'}$.log[P_{new}].focus()
Défini :	$@DR_{c'}$ c <= $@DR_c$, avec :
Extraction et focalisation d'un élément de cette partition	if ER = (*le N P* ∨ *le P*) ∧ ((prop(*P*) ≠ Ordering_attributes ()) ∧ (*P* ≠ *autre*)) $@DR_{C'}$.log[i].item [j].focus <= 1 end if
Identification du référent	Referent <= $@DR_{C'}$.log[i].focus()
Pronom :	$@DR_{c'}$ <= $@DR_c$, avec :
Aucune restructuration	
Identification du référent	Referent <= $@DR_{C'}$.log[i].focus()
Démonstratif :	
Création d'un nouveau domaine	$DR_{c'}$ <= create_dr(cat(*N*)) P_{new} <= $@DR_{C'}$.create_log(Property, D, 1)
Insertion de l'ancien élément focalisé de $@DR_c$ et focalisation	I_{new} <= $@DR_{C'}$.log[P_{new}].create_item($@DR_C$.log[i].focus()) $@DR_{C'}$.log[P_{new}].items[I_{new}].focus <= 1
Identification du référent	Referent <= $@DR_{C'}$.log[P_{new}].focus()

Tableau 36 : Algorithmes pour l'identification du référent et restructuration du DR

Les algorithmes du Tableau 36 reprennent les grands principes dégagés dans notre modélisation : un indéfini crée une nouvelle partition pour en extraire et focaliser un nouvel élément qui sera le référent de l'expression à interpréter. Un défini extrait et focalise un élément d'une partition existante. Notons que l'algorithme présenté ici n'intègre pas encore le test sur l'existence d'une partition dans la RM générique, en cas d'absence d'une partition dans DR_C. Pour un pronom, aucune opération de restructuration n'a lieu : le référent correspond à l'élément focalisé de la partition retenue (`i`) de DR_C.

Enfin, le démonstratif sort l'élément focalisé de la partition (i) de DR_C pour l'insérer dans un nouveau domaine.

En ce qui concerne les expressions d'ordre et d'altérité, elles imposent des contraintes supplémentaires sur l'opération de restructuration. Les expressions ordinales déterminent le lieu d'extraction du référent en fonction de l'ordonnancement de la partition. Les expressions d'altérité imposent que la focalisation du référent ne doit pas se faire sur l'ancien élément focalisé de DR_c (Tableau 37).

	Restructuration du @DR$_c$ en @DR$_{c'}$
« Autre » **Incompatibilité entre ancien et nouveau focus**	Extraire et focaliser un référent en tenant compte de la contrainte suivante: @DR$_C$.log[i].focus ≠ @DR$_{C'}$.log[i].focus

Tableau 37 : Contraintes de restructuration pour « autre »

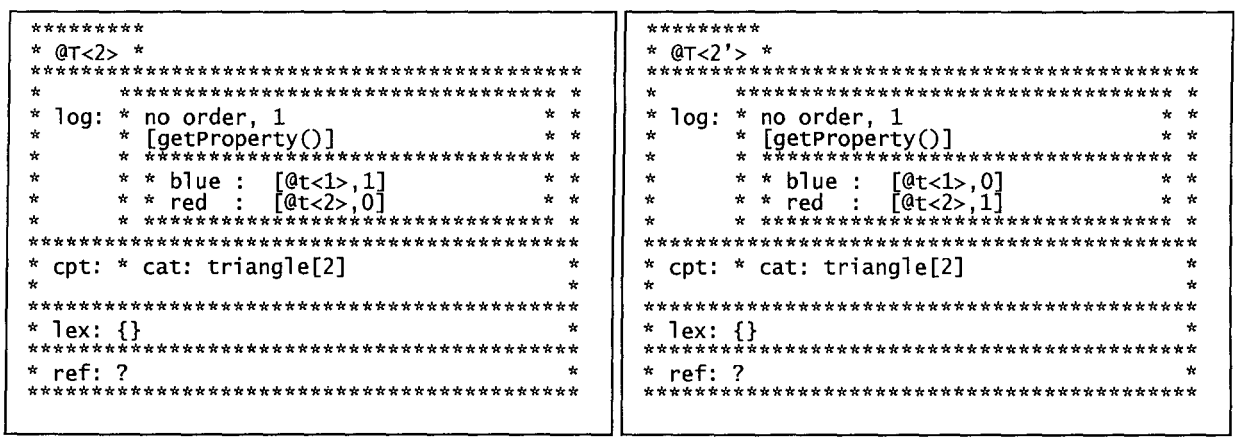


Figure 83 : Restructuration du domaine @T<2> en @T<2'>

La Figure 83 illustre l'opération de restructuration sur l'exemple de l'interprétation de *l'autre triangle* dans le domaine contextuel @T<2>. Selon les algorithmes présentés ci-dessus, la restructuration consiste en une extraction et focalisation de @t<2> comme référent de l'expression. La seule modification structurelle de la partition de @T<2'> consiste donc en un changement de la structure focale.

11.4 Bilan

L'intégration de notre modèle dans une réelle plate-forme de dialogue nous a permis de mettre en évidence les points forts de cette modélisation tout en ouvrant des pistes pour des améliorations possibles.

En commençant par ces dernières, faisons d'abord un bilan de l'état actuel du calcul référentiel implémenté jusqu'ici : le système traite actuellement les descriptions définies anaphoriques, non anaphoriques et associatives, accompagnées ou non d'un modifieur. Il traite également les pronoms de 3^{ième} personne et les démonstratifs anaphoriques et non anaphoriques (accompagnés d'un geste de désignation), sans ou avec un modifieur. Enfin, il traite les possessifs sans ou avec modifieur (*ton*

chapeau) et les constructions *le N1 de N2*. Notons que ces deux dernières constructions, n'ayant pas fait l'objet de la modélisation proposée dans notre thèse, sont traitées pour l'instant par analogie aux constructions *le N à (pronom)* pour les possessifs et *le N1 dans le domaine fourni par N2* pour le second. Si une étude plus approfondie doit décider de la pérennité de ces solutions, le traitement adopté est d'ores et déjà un indice de la productivité de notre modèle.

Parmi les expressions non traitées, on trouve d'abord un certain nombre d'expressions intégrées dans notre modélisation, mais dont le traitement n'est par prioritaire pour les besoins de l'application pilotée : cela concerne les indéfinis, les groupes nominaux sans noms et les expressions d'ordre et d'altérité. Par ailleurs, ne sont pas non plus traitées certaines expressions dont notre modélisation ne tient actuellement pas compte : en particulier les noms propres et les constructions à modifieurs enchâssés dans une relative (*le N que P, le N qui P*) ou à plusieurs modifieurs. D'un point de vue théorique, ces constructions semblent pouvoir être intégrées dans la modélisation proposée, mais pour ce qui est des relatives, cela demande une extension de la modélisation aux éventualités pour le calcul des propriétés.

Parmi les points forts, nous retenons tout d'abord la preuve de l'opérationnalité de notre modèle. Ceci est d'autant plus encourageant que nous sommes partie d'une modélisation qui se voulait le plus proche possible des modèles linguistiques. Or, la préoccupation première de ceux-ci n'est évidemment pas la mise en œuvre informatique, mais l'adéquation aux données observées.

Un deuxième point positif est le caractère unifié des procédures de calcul. En dehors du fait que nous faisons l'hypothèse que ce trait correspond à notre fonctionnement cognitif, cela a facilité la mise en œuvre et le maintien de l'implémentation. Par ailleurs, nous avons déjà mentionné que cette caractéristique permet d'intégrer le traitements d'expressions non encore étudiées en détail, mais pour lesquelles il est possible de faire des hypothèses de fonctionnement qui s'intègrent facilement dans le cadre théorique du modèle proposé.

Enfin, un troisième point positif est la facilité avec laquelle notre approche, basée initialement sur des données linguistiques obtenues dans un monde restreint (le corpus Ozkan), a pu être transposée vers des applications réelles. Ce transfert a probablement profité du fait que le monde des applications pilotées par le système dialogique est également limité, de par le nombre des entités et actions disponibles. Les conséquences pour la mise en œuvre d'un modèle comme le nôtre, nécessitant des connaissances extensives sur les objets et leur composition, s'en ressentent : la description conceptuelle de l'application est plus facile à établir que pour des systèmes traitant des textes très divers et la semi-automatisation de cette tâche, lourde pour un être humain, est par ailleurs à l'étude. Or, le bénéfice d'un modélisation du monde de l'application *in extenso* est un calcul référentiel plus fin et surtout unifié pour tout type d'expression. Si la mise en œuvre pratique de notre approche pour d'autres applications, comme la recherche d'information ou le résumé automatique sur des textes tout-venant, reste à étudier, il nous semble que nous sommes sur la bonne voie en ce qui concerne le dialogue homme-machine finalisé, qui restera restreint, même à moyen terme, à des applications bien circonscrites.

12 Conclusion et perspectives

12.1 Résultats obtenus : un modèle pour l'interprétation référentielle

- *Rappel de la problématique et des objectifs*

L'objectif de notre thèse était de proposer un modèle pour la résolution de la référence aux objets en dialogue homme-machine. Nous avons défini la résolution référentielle comme le processus d'identification, à partir des expressions référentielles employées par un utilisateur, des représentations identifiantes pour les entités du monde de l'application.

Les approches traditionnelles pour l'interprétation automatique d'une expression référentielle consistent essentiellement à filtrer un ensemble d'entités contextuelles en fonction des informations sémantiques véhiculées par l'expression. Tout le problème est alors de déterminer la nature de cet espace de recherche : son extension, sa structuration interne et les mécanismes de sa mise à jour dynamique. Un autre problème est posé par le fait que toutes les expressions référentielles ne semblent pas recruter leur référent en fonction de leur description : c'est en particulier le cas des expressions démonstratives et des pronoms. Une étude des théories et systèmes existants montre d'ailleurs que les modélisations varient considérablement en fonction des expressions référentielles traitées (pronoms vs. descriptions définis), de l'application envisagée (systèmes de compréhension vs. dialogues de commande), de l'environnement d'interaction (mono- vs. multimodal) et de la réutilisabilité ou non du système pour la génération d'expressions référentielles. Or, un dialogue naturel réunit toutes ces conditions à la fois. Il contient tous les types d'expressions référentielles, évolue dans un environnement multimodal et les processus d'analyse et de génération y ont lieu en parallèle.

Partant de l'hypothèse qu'une trop grande diversité des mécanismes n'est ni plausible d'un point de vue cognitif, ni souhaitable d'un point de vue informatique, nous nous sommes fixé comme objectifs :

> de développer un modèle qui permette d'intégrer différents aspects du calcul référentiel dans un **cadre théorique unifié** ;
>
> de veiller à une **adéquation cognitive** de ce modèle, qui se traduit par sa prédictivité et sa réversibilité pour la génération ;
>
> de **valider** la portée théorique de notre modèle **sur un corpus** de dialogues homme-homme (Ozkan, 1994) ;
>
> et de montrer que le modèle proposé est suffisamment formel pour pouvoir faire l'objet d'une **implémentation** dans une plate-forme dialogique réelle.

- *Résultats théoriques*

Un certain nombre de travaux – issus du domaine du traitement automatique de la langue (Romary, 1991 ; Gaiffe, 1992 ; Romary, 1993 ; Dale et Reiter 1995), mais aussi de la linguistique (Corblin 1987) et des sciences cognitives (Olson, 1971 ; Langacker, 1991) – ont mis en évidence l'importance de l'ancrage de toute acte de référence dans un contexte d'interprétation assis sur la notion de contraste ou de différenciation. Il s'en dégage l'hypothèse que l'interprétation d'une expression

référentielle demande l'isolement de l'objet désigné dans un contexte local – ou domaine de référence – sur la base de critères distinctifs le long d'une axiologie.

La conséquence de ce point de vue est que l'interprétation d'une expression est plus que l'identification d'un référent : c'est aussi l'identification de son domaine de référence et donc d'un ensemble de référents alternatifs exclus. Ce point de vue a été intégré depuis une dizaine d'années dans les travaux de génération automatique d'expressions référentielles, en particulier par Dale (1992) et Dale et Reiter (1995). Or, dans les systèmes de compréhension, l'approche la plus courante pour identifier le référent d'une expression référentielle définie (*le triangle*) consiste à filtrer successivement les entités de l'application jusqu'à n'en retenir qu'une seule, compatible avec la description (Kievit et Piwek, 2000). En principe, ce processus est réitéré pour chaque expression, mises à part les heuristiques pour traiter les expressions d'altérité comme *l'autre triangle* (Vivier et al., 1998), les ellipses comme *le rouge* et les *one*-anaphora de l'anglais (SHRDLU, Winograd, 1972), où il est nécessaire de revenir sur des alternatives précédentes.

Contrairement à ces approches, nous pensons que l'identification systématique des « alternatives exclues » peut être mise au profit de processus de compréhension à la fois plus efficaces d'un point de vue informatique et plus proches du fonctionnement cognitif. Le fait qu'elles soient considérées comme fournissant le domaine d'interprétation pour les expressions elliptiques, ambiguës ou ayant trait à l'altérité traduit, selon nous, un principe d'interprétation plus fondamental : ces domaines forment l'espace d'ancrage contextuel préférentiel pour toutes les expressions à interpréter. Nous proposons, par conséquent, **une modélisation du contexte par des domaines de référence**. Il s'agit de sous-ensembles d'entités discursives ou perceptives, créés et modifiés dynamiquement au cours du dialogue et structurés selon des contrastes. Ensuite, nous faisons l'hypothèse que **différents types d'expressions référentielles imposent des contraintes spécifiques sur l'extraction de leur référent** dans ce contexte et qu'elles le restructurent différemment. A partir de là, nous sommes capable de formuler des prédictions portant à la fois sur l'acceptabilité de différents types d'expressions dans un contexte donné et sur des interprétations préférentielles en cas d'ambiguïté. Par ailleurs, nous nous sommes intéressée à la modélisation de deux phénomènes moins étudiés, mais observés fréquemment dans notre corpus : la coréférence entre expressions indéfinies (Salmon-Alt, Gaiffe et Romary, 2000) et le fonctionnement des expressions d'altérité. Nous avons également montré comment nos propositions pourraient s'adapter à la génération d'expressions référentielles dans des contextes multimodaux (Salmon-Alt et Romary, 2000).

- *Résultats pratiques*

La validation systématique du modèle sur un corpus de dialogues a demandé au préalable une réflexion sur des formats de codage de ressources linguistiques. En partant d'une proposition de standard international pour l'annotation coréférentielle (Poesio, 2000), nous avons proposé de rendre ce schéma plus générique par l'utilisation de méta-schémas XML (Salmon-Alt, Romary et Schaaff, 2000). Le nouveau schéma peut ainsi être instancié différemment selon la nature des données à annoter (discursives vs. multimodales) ou la théorie à valider (coréférentielle vs. référentielle).

L'implémentation de notre modèle a été effectuée en collaboration avec le laboratoire central de recherche de la société *Thalès* (ex-*Thomson*) à Orsay. Elle s'intègre dans un cadre plus vaste d'implémentation d'un système de gestion de dialogue. Nous y avons contribué en fournissant la spécification du modèle des données et les algorithmes pour le traitement de la référence aux objets et la gestion du contexte.

12.2 Perspectives de recherche : validation et extension du modèle

Nos perspectives de recherche s'articulent autour de deux axes principaux :

des validations supplémentaires de la notion de domaine de référence ;

l'étude de son extension à de nouveaux champs de recherche.

La validation de la notion de domaine de référence est envisagée sous différents aspects complémentaires : validation en aval par des séries d'expériences psycholinguistiques ciblées permettant de confirmer des prédictions relatives à la stabilité des domaines de référence dans le dialogue ; validation par une confrontation à des données réelles, couplée à une réflexion sur l'acquisition de ressources linguistiques réutilisables ; validation par l'implémentation dans le cadre d'un projet européen sur le dialogue multimodal. Parmi les possibilités d'extension du modèle développé, nous proposons d'explorer trois domaines : une gestion dialogique plus efficace permettant d'éviter ou de traiter des « accidents référentiels », la génération automatique d'expressions référentielles par une réversion du modèle et le développement de systèmes de compréhension multilingues.

- *Validation psycholinguistique des domaines de référence*

Un de nos objectifs sera de tester le pouvoir prédictif de la modélisation référentielle proposée dans un contexte combinant des données linguistiques et visuelles. Bien que des expériences psycholinguistiques, portant sur l'identification d'objets dans un contexte visuel, soient complémentaires à l'étude de corpus, il s'agit d'un point très peu abordé dans la littérature (Kessler et al., 1996). Ce type de validation présente pourtant l'avantage d'un contrôle explicite des facteurs de structuration contextuelle, d'une part par la disposition des objets sur un support visuel et d'autre part par le contrôle des instructions langagières associées. Comme les opérations cognitives de base pour la structuration du contexte – le groupement et l'extraction – ne changent pas selon qu'il s'agit d'un support linguistique ou visuel, cette expérience nous permettra d'extrapoler les résultats portant sur des contextes visuels à des contextes linguistiques.

En collaboration avec F. Landragin (doctorant en informatique au LORIA) et M. Fossard (doctorante en psycholinguistique au Laboratoire J. Lordat de Neuropsycholinguistique, Toulouse), nous avons commencé à réfléchir à un protocole expérimental visant à valider, dans un premier temps, le fonctionnement des expressions définies. Plus précisément, notre modèle prédit qu'une expression définie suppose, pour s'interpréter, un domaine de référence à l'intérieur duquel elle puisse extraire son référent en vertu d'un contraste véhiculé par sa sémantique. Cette opération d'extraction peut donner lieu à la création d'une nouvelle partition du domaine de référence (« création de partition ») ou se faire sur la base d'une partition déjà existante (« parcours de partition »). Or, des raisons d'économie cognitive conduisent à supposer que l'existence de plusieurs partitions au sein d'un même DR est plus difficile à gérer qu'une seule partition. Nous en déduisons alors que le mode de fonctionnement préféré du défini est le parcours domanial le long d'une même partition. Les différentes séries expérimentales envisagées serviront à valider cette hypothèse.

Une première expérience (Figure 84a) est destinée à opposer, dans des configurations minimales, les deux types de parcours : la condition 1 demande au sujet d'identifier le référent de l'expression *le triangle noir* après l'identification du *triangle blanc*, c'est-à-dire dans le prolongement du parcours d'un même domaine, structuré selon la couleur des objets. La condition 2 introduit une rupture de domaine en changeant de critère d'identification (*couleur* vs. *position horizontale*) ce qui doit amener

le sujet à créer une deuxième partition du même domaine. Il s'agit alors d'observer les temps de réaction des sujets et d'en déduire des conclusions sur l'effort cognitif de l'interprétation référentielle en fonction du type d'extraction (parcours ou création). Un temps de réaction plus long en cas de création d'une nouvelle partition (condition 2) confirmerait nos hypothèses.

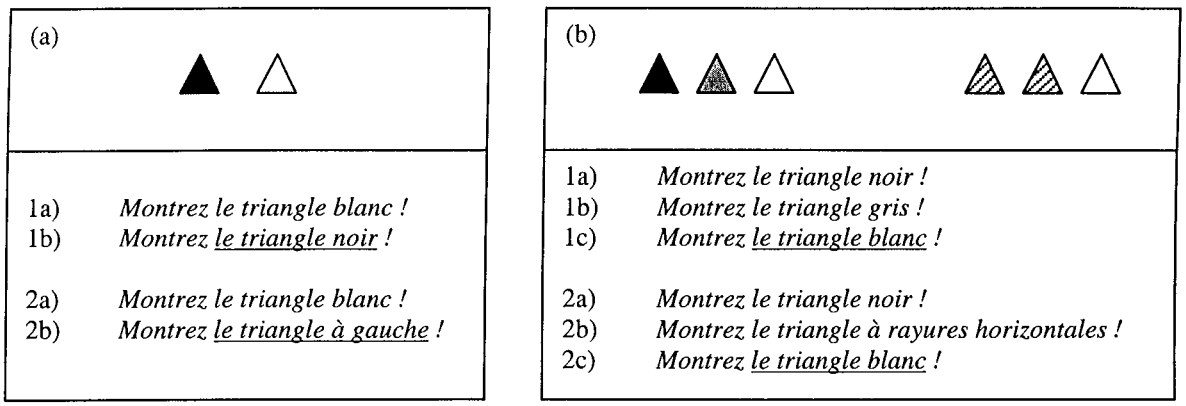


Figure 84 : Exemples de scénario pour la validation expérimentale de la notion de domaine de référence

Une autre expérience (Figure 84b) explore la relation entre le parcours stabilisé d'un domaine et l'interprétation d'expressions ambiguës. Notre hypothèse, selon laquelle l'interprétation préférentielle du défini est le parcours d'une même partition, permet effectivement de faire des prédictions sur l'interprétation préférentielle d'expressions ambiguës. Nous aimerions donc montrer qu'une expression comme *le triangle blanc*, ambiguë par rapport à la totalité du contexte visuel, peut être désambiguisée s'il est possible de lui trouver un référent dans le prolongement du parcours du domaine actif le long de la partition activée (condition 1). Un nombre de réactions référentielles plus élevées et plus rapides en (1) confirmerait cette hypothèse.

Si le protocole expérimental s'avère adapté et les résultats encourageants, ces expériences pourraient se poursuivre en élargissant le type des expressions examinées aux pronoms, démonstratifs et indéfinis et les modes de structuration contextuelle aux gestes de désignation.

- *L'acquisition et la gestion de ressources linguistiques*

Si nous nous intéressons à l'élargissement des champs théoriques qui peuvent être couverts par le modèle des domaines de référence, nous ne voudrions pas perdre de vue pour autant la nécessité de valider les résultats théoriques sur des données linguistiques. Or, il est difficile de disposer de véritables ressources linguistiques réutilisables dans le domaine du traitement de la référence. Un recensement des ressources existants dans ce domaine fait apparaître deux problèmes : d'une part, les formats d'annotation varient considérablement d'une ressource à l'autre, ce qui rend la réutilisation de corpus annotés difficile. D'autre part, les phénomènes annotés ont souvent tendance à relever d'une approche théorique particulière (par exemple, d'une approche anaphorique, réduisant le problème de la référence à une relation entre deux expressions référentielles).

Notre objectif est alors de proposer un schéma d'annotation référentielle qui soit exploitable par rapport à nos préoccupations théoriques, tout en garantissant la réutilisabilité de ces ressources linguistiques : cela demande de veiller à une compatibilité maximale du format d'annotation avec des standards existants (tels qu'issus du projet MATE). L'intégration de phénomènes jusqu'alors non pris en compte devra se faire par une approche modulaire, permettant de maintenir un schéma d'annotation suffisamment générique pour être compatible avec des approches théoriques différentes de la nôtre.

A travers une première expérience dans le codage de dialogues oraux multimodaux, nous avons pu nous rendre compte qu'un schéma d'annotation coréférentielle tel que proposé par MATE (*Multilevel Annotation Tools Engineering*) peut être considéré, non pas directement comme un schéma d'annotation, mais plutôt comme un méta-modèle : les balises d'annotation proposées peuvent effectivement s'intégrer dans un schéma d'annotation abstrait, qui s'instancie différemment selon les propriétés effectives du corpus à annoter ou la théorie à valider. Ce processus d'instanciation repose sur une sélection d'un ensemble de catégories pertinentes pour le typage des entités à annoter et des liens à établir entre ces entités.

Nous avons commencé à concrétiser cette idée par la conception d'un prototype pour une interface permettant d'effectuer le choix entre différents niveaux d'annotation référentielle. Cette réalisation s'appuie largement sur les possibilités offertes par XML : par rapport à un langage de représentation comme SGML, l'utilisation de XML a l'avantage d'augmenter la modularité dans la description des schémas, en permettant de décrire un méta-schéma abstrait dont les catégories de données sont paramétrables en fonction des choix de l'annotateur. Le contrôle des types de données, l'unification de la syntaxe des schémas (en remplacement des DTDs) avec celle des documents codés et la modularisation de la description facilitent la maintenance et la conception de nouvelles interfaces. Enfin, la présentation et la mise en ligne des données annotées sont facilités largement par l'utilisation des feuilles de styles.

- *Validation informatique : le projet MIAMM*

MIAMM (*Multidimensional Information Access using Multiple Modalities*) est un projet européen sur deux ans et demi (début prévu : septembre 2001), réunissant des partenaires académiques et industriels de quatre pays autour d'un objectif commun : le développement de nouveaux concepts et techniques dans le domaine du dialogue multimodal. Ce projet est motivé par le souhait de permettre à un utilisateur non expert d'accéder rapidement et de façon naturelle – c'est-à-dire en utilisant la langue orale ou écrite et des gestes de désignation – à des bases de données multimédia. Un des aspects novateurs du projet consiste en la conception d'une composante de gestion dialogique devant intégrer l'approche multimodale.

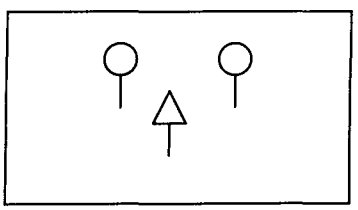
Cette tâche, qui a été confiée à l'équipe *Langue et Dialogue* du Loria, comprend entre autres la modélisation du dialogue et des processus d'interprétation de références multimodales. Dans la mesure où ce projet nous offre un contexte stimulant pour mettre en œuvre notre modèle de l'analyse référentielle, nous aimerions nous y impliquer, en particulier en ce qui concerne la spécification et l'implémentation des modules de calcul référentiel. L'entrée de ce module viendra des composantes d'analyse linguistique et d'analyse des gestes. Il s'agira ensuite de combiner deux approches complémentaires pour la résolution de la référence dans un cadre multimodal : la DRT (Kamp et Reyle, 1993), bien adaptée à la résolution d'anaphores et de présuppositions dans un cadre vériconditionnel et notre modèle des domaines de référence, conçu pour gérer des phénomènes référentiels dans un contexte multimodal, spécifique à un domaine particulier et bien adapté pour traiter par exemple des expressions référentielles accompagnées par des gestes. La complémentarité de ces deux approches se manifestera par ailleurs lors du traitement d'expressions référentielles non interprétables dans un cadre vériconditionnel (*le précédent*, *le premier*,...), mais fréquentes dans des échanges langagiers naturels.

Le cadre fourni par ce projet nous permettra non seulement de valider, sur une autre plate-forme de dialogue homme-machine, notre modèle d'analyse de la référence. Il favorisera également des ouvertures vers d'autres aspects de recherche qui nous intéressent et que nous présenterons ci-dessus : la génération automatique d'expressions référentielles et la gestion des « malentendus dialogiques ».

- *Perspectives dialogiques : prévenir et guérir les « accidents référentiels » ?*

Un autre prolongement possible de notre travail consiste à appliquer notre modèle à l'étude des interactions dialogiques et en particulier à des sous-dialogues de clarification ou de précision (interventions « incidentes », selon D. Luzzati, 1995). Dans le corpus que nous avons étudié, une grande partie de ces interventions incidentes concerne en effet la gestion d' « accidents référentiels », c'est-à-dire de situations dans lesquelles les interlocuteurs re-négocient l'interprétation d'une expression référentielle. Ainsi, en (248), l'interprétation par M de l'expression *le deuxième* dans le contexte visuel associé demande une clarification, car elle est en conflit avec l'indication spatiale *tout à droite*. I a en effet voulu désigner *le deuxième des arbres ronds* et sa formulation s'explique par le fait qu'il prolonge une vision de la scène en construction, à l'intérieur de laquelle il avait précédemment activé le domaine des deux arbres ronds.

(248)
a.	I	là tu prends un triangle
b.	M	mm
c.	I	tu mets euh vers *le deuxième* tout à droite
d.	I	de l'arbre rond
e.	M	à droite de l'arbre euh du deuxième à droite
f.	I	voilà



Notre objectif est alors de relever systématiquement ces « accidents référentiels » et de les comparer par rapport aux prédictions que ferait notre modèle. En particulier, nous aimerions répondre à deux questions. Comment un système de compréhension peut-il « guérir » de tels accidents, c'est-à-dire déployer des stratégies pour comprendre même des expressions ambiguës ou déviantes (bien que courantes dans des dialogues naturels) ? Et comment un système de génération peut-il « prévenir » ces accidents, c'est-à-dire éviter des expressions réellement ambiguës ou déviantes, sans retomber dans des stratégies simplistes (désignation systématiquement non ambiguë dans la totalité du contexte et en utilisant des termes de désignation de base) qui lasseraient rapidement un interlocuteur humain ? Répondre à ces deux questions permettrait en effet de réduire les interventions dialogiques « incidentes » et d'augmenter l'efficacité des systèmes de dialogue.

Une piste de recherche consiste à étudier ces « accidents » en termes de non-concordance entre les « états mentaux » (Traum et Dillenburg, 1996). Une étude préliminaire des conflits référentiels du corpus Ozkan a en effet montré que ceux-ci s'expliquent très souvent par une différence entre les représentations contextuelles (c'est-à-dire des domaines de référence) activées respectivement par le locuteur A et l'interlocuteur B. Par conséquent, B doit, pour interpréter une expression problématique, opérer un changement de domaine actif. Les extraits étudiés de façon ponctuelle montrent que cette opération dépend de plusieurs facteurs susceptibles de la favoriser ou de la freiner : historique du dialogue, nature des tâches effectuées précédemment, contexte visuel et connaissances encyclopédiques. Si une étude plus détaillée doit encore mettre à jour l'influence exacte de ces facteurs, nous faisons l'hypothèse que la réponse aux deux questions initiales sera la suivante : afin de « guérir » des « accidents référentiels », c'est-à-dire interpréter des expressions problématiques, le système doit opérer des changements de domaine, dans la limite de contraintes à étudier plus précisément. Afin de « prévenir » des « accidents référentiels », c'est-à-dire utiliser des expressions non problématiques, le système doit générer des expressions qui demandent, pour être interprétées, un changement minimal par rapport au domaine activé de l'interlocuteur humain.

- *Perspectives pour la génération automatique d'expressions référentielles*

Une de nos motivations majeures était de proposer un modèle pour le traitement automatique des langues qui soit à la fois implémentable et plausible d'un point de vue cognitif. Comme nous faisons

l'hypothèse que les processus cognitifs impliqués dans l'analyse et la génération ne sont pas fondamentalement différents l'un de l'autre, il s'ensuit que la preuve de la réversibilité de notre modèle pour la génération automatique d'expressions référentielles serait un bon indicateur pour la plausibilité cognitive de notre proposition.

L'enjeu fondamental de la génération de textes écrits ou oraux consiste à passer d'une représentation non linguistique du contenu à une expression de ce contenu sous forme linguistique. Un des problèmes est la génération d'expressions référentielles. Ce problème a fait l'objet de travaux importants de la part de R. Dale et E. Reiter (1992, 1995, 1996). L'algorithme « standard » est leur algorithme incrémental dont le but est de générer des descriptions « distinctives » (*distinguishing descriptions*). Une description distinctive a la propriété de caractériser une entité *R* (le référent), mais aucune autre entité alternative (ou *distractor*) d'un ensemble contextuel *C*. Étant donné un tel ensemble *C*, l'algorithme parcourt une liste ordonnée des attributs caractérisant les entités dans *C* et retient la valeur d'un attribut lorsque celui-ci permet d'exclure au moins une des entités alternatives de l'ensemble. Il s'arrête lorsque toutes les alternatives ont été exclues et génère une description contenant toutes les valeurs retenues.

Un problème non pris en compte par cet algorithme est la construction et la mise à jour de l'ensemble contextuel. Même s'il est théoriquement associé à la notion d'espace focal de B. Grosz et C. Sidner (1986), il semble correspondre, en pratique, à l'ensemble global des entités contextuelles (*global working set* ; Dale, 1992). Or, nombre d'exemples de notre corpus montrent que cela ne traduit pas le fonctionnement référentiel tel que l'on observe à travers des données réelles . Ces exemples suggèrent plutôt que le caractère approprié ou non d'un attribut dans une description change dynamiquement au cours du dialogue et ceci en fonction du ou des ensemble(s) contextuel(s) – ou domaine(s) de référence – activé(s). Il semble donc que la nécessité de gérer des ensembles contextuels locaux se pose aussi bien pour l'analyse que pour la génération automatique d'expressions référentielles. Nous retrouvons donc ici un aspect qui a motivé une grande partie de notre travail de thèse.

Comme pour la modélisation des processus d'interprétation, la structuration des domaines de référence se fait par ailleurs selon des facteurs multimodaux. Ainsi, en (249), les deux expressions elliptiques *l'autre* et *la première*, bien que parfaitement interprétables (et interprétées) ne peuvent pas être générées par les algorithmes courants. Générer de telles expressions demande effectivement d'intégrer, dans la modélisation de l'évolution contextuelle, des informations visuelles et gestuelles. Or, le modèle des domaines de référence est précisément conçu pour représenter de façon unifiée des informations contextuelles provenant de différentes sources. L'utiliser pour la génération automatique d'expressions référentielles permettrait donc non seulement de générer des expressions plus variées, mais aussi des expressions plus proches de la maxime de quantité de Grice (1979), préconisant de dire tout ce qui est nécessaire, mais sans dire plus que ce qui est nécessaire.

(249) A1 alors il va falloir que tu prennes deux grandes barres
B1 [geste : B prend une grande barre et la pose sur l'écran]
A2 bon voilà...et *l'autre* tu vas la mettre parallèle à *la première*

Enfin, une étude attentive des données linguistiques montre que le choix des attributs composant la description se fait selon des principes dont les algorithmes classiques ne tiennent pas toujours compte. Lorsque plusieurs attributs permettent de distinguer de façon unique une entité de l'ensemble contextuel, le choix des attributs est effectivement lié au point de vue projeté par le locuteur sur l'ensemble contextuel pertinent. Ainsi, un rond sur l'écran sera successivement désigné par *le rond* et *ce soleil*, selon qu'il est conçu en tant qu'objet géométrique ou en tant que partie intégrante d'un dessin cible.

(250) I donc tu prends le gros rond et tu le mets à droite de l'église
I et en dessous de *ce soleil* on va faire la petite fillette

- *Perspectives multilingues*

Enfin, il serait intéressant de valider le modèle des domaines de référence sur le fonctionnement référentiel d'autres langues. En plus d'une validation psycholinguistique, cette validation apporterait des arguments supplémentaires en faveur de la plausibilité cognitive de ce que nous proposons. Nous pensons d'abord à une étude comparative *français-anglais-allemand* de certains marqueurs référentiels. Il semble en particulier que le démonstratif ne fonctionne pas tout à fait de la même façon, même dans ces langues relativement proches. Afin d'élargir le cercle des langues traitées, nous avons pris contact avec une collègue brésilienne, Renata Vieira, dont les thèmes de recherche sont proches des nôtres (Poesio et Vieira, 1998 ; 2000). Dans le cadre d'un appel d'offre INRIA-CNPq, nous avons soumis un projet de collaboration qui devrait mener à l'implémentation d'un système de compréhension des descriptions définies et démonstratives du français et du portugais. Mais nous aimerions également explorer l'adaptation de notre modèle pour des langues dont le système des déterminants est très différent de celui du français, comme le russe ou le japonais. En ce qui concerne le japonais, nous avons pris contact, dans le cadre d'une collaboration qui se met en place entre le CRLOA et le LORIA, avec un spécialiste du japonais au CNRS et nous comptons réaliser une étude exploratoire à travers l'encadrement d'une stagiaire linguiste, parlant le japonais.

Bibliographie

Almor A. (1999). Noun Phrase Anaphora and Focus: The Informational Load Hypothesis. *Psychological Review*, 106/4, 748-765.

Apothéloz D., Reichler-Béguelin M.-J. (1995). Construction de la référence et stratégies de désignation. *TRANEL*, 23, 27-271.

Ariel M.(1990). *Accessing Noun-Phrase Antecedents*. Routledge, London and New York.

Asher N. (1993). *Reference to Abstract Objects in Discourse*. Kluwer Academic Publishers. Dordrecht, Boston, London.

Asher N., Lascarides A. (1998). Questions in Dialogue. *Linguistics and Philosophy*, 21, 237-309.

Beaver D. (2000). *Centering in OT*. Communication "Ingénierie des Langues", LORIA, 6th October 2000, Nancy, France.

Berrendonner A., Reichler-Béguelin M.-J. (1996). De quelques adjectifs à rendement anaphorique : premier, dernier, autre. *Studi italiani di Linguistica Teoretica e Applicata*, XXV(3), 475-502.

Blanche-Benveniste C. (1990). *Le français parlé : études grammaticales*. Éditions du CNRS, Paris.

Blanche-Benveniste C. et Chervel A. (1966). Recherches sur la syntagme substantif. *Cahiers de Lexicologie*, IX/2, 3-37.

Bobrow D., Kaplan R., Kay M., Norman D., Thompson H., Winograd T. (1977). GUS : A Frame-Driven Dialogue System. *Artificial Intelligence*, 8, 155-173.

Bordeaux F. (1993). Introduction aux théories de Langacker : Notions de grammaire cognitive. *Internal Rapport INRIA*, LIMSI-93-12.

Bos, J., Mineur, A-M., Buitelaar, P. (1995). Bridging as Coercive Accommodation. *Proceedings of CLNLP '95*, Edinburgh.

Bosch P., Geurts B. (1990). Processing definite NPs. *Rivista di Linguistica*, 2/1, 178-199.

Bourguet, M.-L. (1992). *Conception et réalisation d'une interface de dialogue personne-machine multimodale*. Ph.D., Institut National Polytechnique, Grenoble.

Brennan, S. E., Friedman, M. W., Pollard, C. J. (1987). A centering approach to pronouns. *Proceedings, 25th Annual Meeting of the ACL*. Stanford, CA, 155-162.

Bruneseaux F. et Romary L., (1997). Codage des références et coréférences dans le DHM. *Proceedings ACH-ALLC '97*, Kingston Ont, Canada.

Caelen J. (1991). Interaction multimodale dans ICP-Draw. Expérience et Perspective. *Actes IHM'91. GRECO-PRC Commmunication Homme-Machine*, 1991.

Carberry S. (1989). A pragmatics-based approach to ellipsis resolution. *Computational Linguistics*, 15-2, pp. 75-96.

Charolles M. and Schnedecker C. (1993). Coréférence et identité. Le problème des référents évolutifs. *Langages*, 112, 106-126.

Chomsky N. (1964). *Aspects de la théorie syntaxique*. Éditions du Seuil, Paris.

Clark H.H., Wilkes-Gibbs D. (1986). Referring as a collaborative process. *Cognition*, 22, 1-39.

Colineau N. (1997). *Etude des marqueurs discursifs dans le dialogue finalisé*. Ph.D., Université Joseph Fourier, Grenoble 1.

Corblin F. (1983). Défini et démonstratif dans la reprise immédiate. *Le français moderne*, 51/2, 118-133.

Corblin F. (1987). *Indéfini, Défini et Démonstratif*. Droz, Genève.

Corblin F. (1990). Les groupes nominaux sans nom du français. In : *L'anaphore et ses domaines*. Kleiber G. et Tyvaert J.E. (eds.), Klinksieck, Paris, 63-80.

Corblin F. (1994). La condition de nouveauté comme défaut. *Faits de Langues*, 4, 147-160.

Corblin F. (1995). *Les formes de reprise dans le discours*. Presses Universitaires de Rennes, France.

Corblin F. (1997). Les indéfinis : variables et quantificateurs. *Langue Française*, 116, 8-32.

Corblin F. (1998). Celui-ci anaphorique : un mentionnel. *Langue française*, 120, 33-43.

Corblin F. (1999). Les références mentionnelles : le premier, le dernier, celui-ci. In : *La référence (2). Statut et processus.* Mettouch A., Quinyin H. (1999) (eds.), Travaux linguistiques du CERLICO, PUF Rennes.

Creissels D. (1995). *Eléments de syntaxe générale.* PUF Grenoble.

Cristea D., Ide N., Romary L. (1998). Veins Theory : A model for global Discourse Cohesion and Coherence, *Proceedings Coling-ACL 1998*, Nantes.

Dale R. (1992) *Generating Referring Expressions. Constructing Descriptions in a Domain of Objects and Processes.* The MIT Press, Cambridge, London.

Dale R., Reiter E. (1995). Computational Interpretations of the Gricean Maxims in Generating Referring Expressions. *Cognitive Science,* 18, 233-263.

Dale R., Reiter E. (1996). The Role of Gricean Maxims in the Generation of Referring Expressions. *Proc. of the 1996 AAAI Spring Symposium on Computational Models of Conversational Implicature,* Stanford University, California.

Dale R., Haddock N. (1991). Content determination in the generation of referring expressions. *Computational Intelligence*, 7/4, 252-265.

Danlos L. (1992). Contraintes syntaxiques de pronominalisation en generation de textes. *Langages,* 106, 36-62.

Danlos L. (1999). Event Coreference between two sentences. *Proceedings of International Workshop on Computational Semantics.* Tildburg, Sweden.

Danlos L., Gaiffe B. (2000). Coréférence événementielle et relations de discours. *TALN2000*, Lausanne, Switzerland.

David J., Kleiber G. (1986). *Déterminants : syntaxe et sémantique.* Klincksieck, Paris.

De Mulder W. (1998). Du sens des démonstratifs à la construction d'univers. *Langue française,* 120, 21-32.

De Mulder W. et Tasmowski-De Ryck L. (1997). Référents évolutifs, syntagmes nominaux et pronoms. *VERBUM*, XIX/1-2., 121-137.

Dekker P. (1998). Speaker's Reference, Descriptions, and Information Structure. *Journal of Semantics*, 15-4.

Dubois J. (1965). *Grammaire structurale du français. Nom et Pronom.* Larousse, Paris.

Duermael F. (1994). *Référence aux actions dans des dialogues de commande homme-machine.* Ph.D., Institut National Polytechnique de Lorraine.

Fillmore C.J. (1968). The case for case. In : *Universals in Linguistic Theory.* Bach E., Harms R. (eds.), Holt, Rinehart and Winston, New York, 1-88.

Fossard M. (1999). Traitement anaphorique et structure du discours : Etude psycholinguistique des effets du « Focus de discours » sur la spécificité de deux marqueurs référentiels : le pronom anaphorique « il » et le nom propre répété. *Troisième colloque des jeunes chercheurs en sciences cognitives.* Soulac, France.

Frege G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik.* Berlin, 25-50.

Frey C. (1989). L'anaphore et l'ellipse. In : Morel M. A. (1989), 169-214.

Gaiffe B. (1992) *Référence et Dialogue Homme-Machine : Vers un modèle adapté au multi-modal.* Ph.D., Université H. Poincaré, Nancy.

Gaiffe B., Grisvard O., Reboul A. (1999). La représentation des actes de langage comme événements pour traiter le dialogue. *TALN 1999*, Cargèse.

Gaiffe B., Reboul A., Romary L. (1997) Les SN définis : anaphore, anaphore associative et cohérence. In : *Relations anaphorique et (in)cohérence.* De Mulder W., Tasmowski-Ryck L., Vetters C. (eds.), Rodopi, Amsterdam-Atlanta, 69-97.

Gaiffe B., Reboul A., Pierrel J-M. (2000) Le traitement des déictiques dans le DOHM. *TALN 2000*, 409-418.

Galmiche M. (1986). Référence indéfinie, événements, propriétés et pertinence. In : *Déterminants : Syntaxe et sémantique*. David J., Kleiber G. (eds.), Klincksieck, Paris, 41-71.

Grice P. (1979). Logique et Conversation. *Communication*, 30, 57-72.

Grisvard O. (2000) *Modélisation et gestion du dialogue oral homme-machine de commande.* Ph.D., Université H.Poincaré, Nancy 1.

Grisvard O., Gaiffe B. (1999). An Event Based Dialogue Model and its Implementation in MultiDial2. *Eurospeech 1999*, vol.3, 1155-1158. Budapest, Hungary.

Groenendijk J., Stokhof M., Veltman F. (1995). Coreference and Contextually Restricted Quantification. *Proceedings of the Fourth Conference on Semantics and Linguistic Theory*, Ithaca, NY.

Groenendijk J., Stokhof M., Veltman F. (1996). Changez le contexte ! *Langages,* 123, 8-29.

Grosz B.J. (1981). Focusing and Description in Natural Dialogues. In : *Elements of Discourse Understanding.* Joshi A., Webber B., Sag I. (eds.), Cambridge University Press, p.48-105.

Grosz B.J., Sidner, C. (1986). Attention, Intention and the Structure of Discourse. *Computational Linguistics,* 12, 175-204.

Grosz B.J., Joshi A.K., Weinstein S. (1995) Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics,* 12(2), 203-225.

Guarino N. (1998). Some Ontological Principles for Designing Upper Lexical Resources. *Proceedings 1st International Conference on Language Resources and Evaluation*, Granada, Spain.

Guillaume G. (1919). *Le problème de l'article et sa solution dans la langue française.* Paris, Hachette. [Réédition avec préface de R. Valin, Paris et Québec, Nizet et Presses de l'Université Laval].

Gundel J., Hedberg N.A., Zacharski R. (1993). Giveness, Implicature and the form of referring expressions in discourse. *Language*, 69/2, 275-307.

+ DML RYi (,VVXHV RI 6HQWHQFH 6WUXFWXUH DQG ' LVFRXUVH 3DWWHUQV *Theoretical and Computational Linguistics*, 2, Charles University, Prague.

Hawkins J. A., (1978). *Definiteness and Indefiniteness – A Study in Reference and Grammaticality Prediction.* Croom Helm, London.

Heim I. (1982). *The Semantics of Definite and Indefinite Noun Phrases.* Ph.D., University of Massachusetts-Amherst.

Heim I. (1992). Presupposition Projection and the Semantics of Attitude Verbs. *Journal of Semantics*, 9, 183-221.

Hirschman L., Chinchor M. (1997). MUC-7 Coreference Task Definition, Version 3.0. *Proceedings MUC-7.*

Hobbs J. (1978). Resolving Pronoun Anaphora. *Lingua*, 44, 311-338.

Hovy E. (1993). From interclausal relations to discourse structure : a long way behind, a long way ahead. In : *New concepts in natural language generation : planning, realization and systems.* Horacek, H., Zock, M. (eds.), Printer Publishers, New York, 57-68.

Johnson-Laird P.-N. (1983). *Mental Models.* Havard University Press, Cambridge, Mass.

Johnson-Laird P.-N. (1988). La représentation mentale de la signification. *RISS*, 115, 53-69.

Kamp H. (1981). A Theory of Truth and Semantic Representation. In : *Formal Methods in the Study of Language.* Groenendijk J., Janssen T. , Stokhof M. (eds.), Amsterdam. 277-322.

Kamp H. and Reyle U. (1993). *From Discourse to Logic.* Kluwer Academic Publishers. Dordrecht, Boston, London. .

Karttunen L. (1976). Discourse referents. In : *Syntax and Semantics 7 : Notes from the Linguistic Underground.* McCawley J.D. (ed.), New-York, Academic Press, 363-385.

Karttunen, L. (1974). Presupposition and linguistic context. *Theoretical Linguistics,* 1,181-94.

Kempson, R. (1986). Definite NPs and Context-Dependence: A Unified Theory of Anaphora. In : *Meaning and Interpretation.* Travis (ed.) , Oxford Blackwell, 209-239.

Kessler K., Duwe I., Strohner H. (1996). Sprachliche Objektidentifikation in ambigen Situationen : Empirische Befunde. *SFB 360 Situierte künstliche Kommunikation, Report 96/1*, Universität Bielefeld.

Kievit, L., P. Piwek (2000). Multimodal Cooperative Resolution of Referential Expressions in the DenK System. In : *Proceedings of CMC'98.* Bunt, H. and R.J. Beun (eds), Lecture Notes in AI, Springer Verlag, Berlin.

Kintsch W. (1974). *The representation of meaning in memory.* Lawrence Erlbaum Assciates, Hillsdale, N.Y.

Kleiber G. (1986) Pour une explication du paradoxe de la reprise immédiate Un N – Le N / Un N – Ce N. *Langue française,* 72, 54-79.

Kleiber G. (1988). Reprise immédiate et théorie des contrastes. *Studia Romanica Posnaniensa* 13, 67-83.

Kleiber G. (1989). *Reprise(s). Travaux sur les processus récérentiels anaphoriques.* Publications du groupe Anaphore et Déixis, 1, Université des Sciences Humaines de Strasbourg. Strasbourg.

Kleiber G. (1990). Quand il n'a pas d'antécédent. *Langages*, 97, 24-50.

Kleiber G. (1994). *Anaphores et Pronoms.* Editions Duculot, Louvain-la-Neuve.

Kleiber G. (1997). Massif/comptable et partie/tout, *VERBUM* XIX/3, 321-338

Kleiber G., Patry R., Menard N. (1992). L'anaphore associative : Dans quel sens roule-t-elle ? *Recherches Linguistiques,* XIX, 129-152.

Kleiber G., Schnedecker C., Ujma L. (1994). L'anaphore associative, d'une conception à l'autre. *L'anaphore associative, aspects linguistiques, psycholinguistiques et automatiques*, 5-66.

Krahmer E., Theune M. (1999). Efficient Generation of Descriptions in Context. In : *Proceedings of the ESSLLI workshop on the generation of nominals*, August 9-14, 1999, Utrecht, The Netherlands.

Laborde M.-C. (1999). *Anaphore nominale et référence mentionnelle.* Mémoire de DEA, Université Paris 7.

Landragin F. (1998). *Interaction multimodale dans un environnement virtuel.* Rapport de stage de fin d'études de l'Institut d'Informatique d'Entreprise (IIE), Thomson-CSF, LCR, septembre 1998.

Langacker R. (1969). Pronominalization and the chain of command. In : *Modern Studies in English.* Reibel D, Schane S. (eds.), Englewood Cliffs, Prentice Hall.

Langacker R.W. (1987). Foundations of Cognitive Grammar. Stanford University Press. Stanford.

Langacker R.W. (1991). *Concept, image, and symbol : the cognitive basis of grammar.* Mouton de Gruyter, 395 p.

Lappin, S., Leass H. (1994). An algorithm for pronominal anaphora resolution. *Computational Linguistics*, 20(4), 535-561.

Lascarides A., Asher N. (1993). Temporal interpretation, discourse relations and common sense entailment. *Linguistics and Philosophy*, 16/5, 473-493.

Linsky L. (1974). *Le problème de la référence.* Editions du Seuil, Paris, France.

Luzzati D. (1995). *Le Dialogue verbal Homme-Machine.* Masson, Paris.

Mann W.C., Thompson S.A. (1988). Rhetorical structure theory : A theory of text organization. *Text*, 8/3, 243-281.

Manuélian H. (1999). *Les G-RM : Proposition d'adaptation des RM à la génération automatique d'expression référant à des objets.* Mémoire de DEA, Université Nancy 2.

McCoy K.F. and Strube M. (1999). Taking Time to Structure Discourse : Pronoun Generation Beyond Accessibility. *Proc. of the 21th Annual Conference of the Cognitive Science Society.* Vancouver, Canada.

Mengel, A., Dybkjaer, L., Garrido, J.M., Heid, U., Klein, M., Pirrelli, V., Poesio, M., Quazza, S., Schiffrin, A., and Soria, C. (2000). *MATE Dialogue Annotation Guidelines*, January 2000. (http://www.ims.uni stuttgart.de/projekte/mate/mdag/).

Milner C. (1978). *De la syntaxe à l'interprétation. Quantités, Insultes, Exclamation.* Éditions du Seuil, Paris.

Milner C. (1982). *Ordres et raisons de langue.* Edition du Seuil, Paris, France.

Mitkov R. (1998). Robust pronoun resolution with limited knowledge. *Proceedings of the 18.th International Conference on Computational Linguistics (COLING'98)/ACL'98 Conference*, Montreal, Canada.

Moeschler J., Béguelin M.-J. (1996). *Référence temporelle et nominale.* Actes du 3e cycle romand de Sciences du Langage, Cluny. Peter Lang, Bern.

Morel M. A. (1989). *Analyse linguistique d'un corpus.* Publications de la Sorbonne Nouvelle, Paris.

Moser M., Moore J. (1996). Toward a synthesis of two accounts of discourse structure. *Computational Linguistics* 22/3, 409-419.

Moulton J. (1994) An AI module for Reference Based on Perception. *AAAI Workshop on Integration of Natural Language and Vision Processing*, Seattle.

Olson D. (1970). Language and Thought : Aspects of a Cognitive Theory of Semantics. *Psychological Review*, 77/4, 257-273.

Ozkan N. (1994). *Vers un modèle dynamique du dialogue : analyse de dialogues finalisés dans une perspective communicationnelle.* Ph.D., INP Grenoble.

Ozkan N., Caelen J. (1994). Design Issues for adaptative multimodal interfaces. *ERCIM Workshop Reports*, INRIA, 89-98.

Morin P., Pierrel. J.-M. (1987). PARTNER : un système de dialogue oral homme-machine. *Actes du congrès COGNITIVA*, 354-367.

3 DGX HYD ($ QDSKRULF UHODWLRQV DQG WKHLU UHSUHVHQWDWLRQV LQ WKH GHHS VWUXFWXUH RI D WH[W ,Q *Progress in Liguistics*. Bierwisch M. , Heidolph K.E. (eds.), Den Haag, Mouton, 224-232.

Partee B. (1984). Nominal and Temporal Anaphora, *Linguistics and Philosophy*, 7, 243-286.

Pierrel J.-M., Romary L. (1997). *Quelles références dans les dialogues homme-machine ?.* In : Sabah G. et al. (1997), 183-248.

Poesio M., Vieira R. (1998). A Corpus-Based Investigation of Definite Description Use. *Computational Linguistics*, 24/2, pp. 183-216.

Poesio M. (1998). Cross-speaker anaphora and dialogue acts. *Proc. of the Workshop on Mutual Knowledge, Common Ground and Public Information, ESSLLI 1998.*

Poesio M. (2000) *Coreference. MATE Dialogue Annotation Guidelines-Deliverable D2.1,* January 2000, 126-182. (http://www.ims.uni-stuttgart.de/projekte/mate/mdag/)

Popescu-Belis, A. (1999). *Modélisation multi-agent des échanges langagiers : application au problème de la référence et à son évaluation.* Ph.D., Université Paris-XI.

Popescu-Belis, A. et Robba, I. (1997). Cooperation between Pronoun and Reference Resolution for Unrestricted Texts. *ACL'97/EACL'97 Workshop on Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts*, Madrid, Spain, 94-99.

Popescu-Belis, A. et Robba, I. (1998). Three New Methods for Evaluating Reference Resolution. *Workshop on Linguistic Coreference, LREC'98,* Grenada, Spain.

Popescu-Belis, A., Robba, I. et Sabah, G. (1998). Reference Resolution Beyond Coreference: a Conceptual Frame and its Application. *Coling-ACL'98*, Montréal, Canada, 1046-1052.

Prince A., Smolensky P. (1993). *Optimality Theory: Constraints Interaction in Generative Grammar.* Technical Report, Rutgers University, Center for Cognitive Science.

Prince E. (1981). Toward a taxonomy of given-new information. In : *Radical Pragmatics*, Cole P.(ed.), Academic Press, New York, 223-255.

Pustejovski J. (1995). *The generative lexicon.* MIT, Massachusetts.

Reboul A. (1990). Pragmatique de l'anaphore pronominale. *Sigma*, 12/14, 197-231.

Reboul A. (1998). A relevance theoretic approach to reference. *Relevance Theory Workshop*, Luton, UK.

Reboul A. (2000). La représentation des éventualités dans la théorie des représentations mentales. *Cahiers de Linguistique Française*, 22, 13-55.

Reboul A., Moeschler J. (1994). *Dictionnaire encyclopédique de pragmatique*. Editions du Seuil, Paris.

Reboul A., Balkanski C., Briffault X., Gaiffe B., Popescu-Belis A., Robba I., Romary L., Sabah G. (1997). *Le projet CERVICAL : Représentations mentales, référence aux objets et aux événements.* Research Report, Loria-CNRS/Limsi, France.

Reboul A., Moeschler J. (1998). *Pragmatique du discours. De l'interprétation de l'énoncé à l'interprétation du discours*. Armand Colin, Paris.

Reinhart, T. (1983). Coreference and bound anaphora: A restatement of the anaphora questions. *Linguistics and Philosophy* 6/1.

Reiter E., Dale R. (1992). A Fast Algorithm for the Generation of Referring Expression. *COLING 1992*, 232-238.

Rich E., Luperfoy S. (1988). An architecture for anaphora resolution. *ACL Conference on Applied Natural Language Processing*, 18-24.

Rock I., Palmer S. (1991). L'héritage du gestaltisme. *Pour la Science*, 160, 64-70.

Romary L. (1989). *Vers la définition d'un modèle cognitif autour de la représentation du temps dans un système de dialogue homme-machine*. Ph.D., Université H. Poincaré, Nancy.

Romary L. (1993). L'interprétation de "ici" dans des énoncés de positionnement. In : *Le dialogue homme-robot en langage naturel : problèmes psychologiques*. Vivier J. (ed.), Presses Universitaires de Caen.

Rosch E. (1978). Principles of categorization. In : *Cognition and Categorization*. Rosch E., Lloyd B. (eds.), L.Erlbaum, Hillsdale, N.J., 27-48.

Roulet E. et al. (1985). *L'articulation du discours en français contemporain*. Lang, Berne

Russell B. (1905). On denoting. In : *Logic and Knowledge. Essays by Bertrand Russell from 1901 to 1950*. Allen & Unwin, 41-56.

Sabah G. (1988). *L'intelligence artificielle et le langage*. Hermès, Paris.

Sabah G. (1997). *La langue et la communication homme-machine, état et avenir*. In : Sabah et al. (1997), 23-74.

Sabah G., Vivier J., Vilnat A., Pierrel J.-M., Romary L., Nicolle A. (1997). *Machine, Langage et Dialogue*. L'Harmattan, Paris.

Sacks H., Schlegloff E. A., Jefferson, G. (1974). A simplest systematic for the organization of turn-taking in conversation. *Language*, 50, 696-735.

Salmon-Alt S. (2000). Interpreting referring expressions by restructuring context. *ESSLLI 2000*, Student Session, Birmingham, UK.

Salmon-Alt S. (2000). La référence aux objets dans le dialogue homme-machine finalisé. *TALN2000/Récital*, Lausanne, Switzerland.

Salmon-Alt S. (2001a). Reference Resolution within the Framework of Cognitive Grammar. *International Colloquium on Cognitive Science*, San Sebastian, Spain, May 2001.

Salmon-Alt S. (2001c). Entre corpus et théorie : l'annotation (co)référentielle. Submitted to *Traitement Automatique de la Langue*, Special Issue on Corpus Linguistics and Natural Language Processing.

Salmon-Alt S., Gaiffe B., Romary L. (2000). Questioning Indefinites in Human Machine Dialogues. *DAARC'2000*, Lancaster, UK.

Salmon-Alt S., Romary L. (2000). Generating Referring Expressions in Multimodal Contexts. *Workshop on Coherence in Generated Multimedia. INLG 2000*, Mitzpe Ramon, Israel.

Salmon-Alt S., Romary L., Schaaff A. (2000) Increasing the Genericity of the Mate Annotation Framework : The Case of Reference. *First EAGLES/ISLE Workshop on Meta-Descriptions and Annotation Schemas for Multimodal/Multimedia Language Resources. LREC 2000*, Athens, Greece, 29/30 May 2000.

Sanford S.C., Garrod A.J. (1982). The mental representation of discourse in a focussed memory system : implication for the interpretation of anaphoric noun phrases. *Journal of Semantics*, 1/1, 21-41.

Sauvage C. (1993). *Parallélisme et traitement automatique des langues : Application à l'analyse des énoncés elliptiques*.Ph.D., LIMSI, Orsay.

Scha R., Polanyi L. (1988). An augmented context free grammar for discourse. *Coling 1988*, 573-577.

Schang D. (1997). *Représentation et interprétation de connaissances spatiales dans un système de dialogue homme-machine.* Ph.D., Université H. Poincaré, Nancy.

Schank R. (1975). The structure of episodes in memory. In : *Representation and Understanding : Studies in Cognitive Science*, Academic Press, N.Y., 237-272.

Schnedecker C. (1998). L'un et l'autre ou quelques aspects d'une union libre. *Revue de sémantique et de pragmatique*, 3, 177-195.

Schnedecker C., Bianco M. (1995). Antécédents dispersés et référents conjoints. *Sémiotiques*, 8, 72-108.

Sidner C. (1979*). Towards a computational theory of definite anaphora comprehension in English discourse*. Ph.D., M.I.T., Massachusetts. Technical Report 537.

Sidner C. (1983). Focusing in the Comprehension of Definite Anaphora. In : *Computational Models of Discourse*. Brady M., Berwick R. (eds.), MIT Press, Cambridge, 267-333.

Simon P. (1987). *Parts and wholes.* Oxford, Clarendon.

Sperber D., Wilson D. (1986) *Relevance, Communication and Cognition.* Basil Blackwell, Oxford.

Sperber D., Wilson D. (1989). *Pertinence, Communication et Cognition.* Minuit, Paris.

Stalnaker R. (1978). Assertion. In : *Syntax and semantics IX. Pragmatics.* Cole P. (ed.), New York Academic Press, 315-22.

Strawson P.F. (1977). *Etudes de logique et de linguistique.* Editions du Seuil. Paris.

Strube M. (1998). Never look back : An alternative to centering. *Coling/ACL 1998*, Montréal, Canada.

Thorisson K.R. (1994) Simulated Perceptual Grouping : An Application to Human Computer Interaction. *16th Annual Conference of Cognitive Science Society,* Atlanta.

Traum D.R., Dillenbourg P. (1996). Miscommunication in multi-modal collaboration. *AAAI Workshop on Detecting, Repairing and Preventing Human-Machine Miscommunication.* August 1996.

van der Sandt R. (1992). Presupposition Projection as Anaphora Resolution. *Journal of Semantics*, 9, 333-377.

van Hoek K. (1995) Conceptual Reference Points: A cognitive grammar account of pronominal anaphora constraints. *Language,* 71/2, 310-340.

Vandeloise C. (1986). *L'espace en français.* Éditions du Seuil, Paris.

Vivier J., Nicolle A. (1997*). Questions de méthode en dialogue homme-machine : l'expérience Compèrobot.* In Sabah et. al. (1997), 249-305.

Webber B.L. (1978). Description Formation and Discourse Model Synthesis. In : *Theoretical issues in natural language processing.* Waltz D. (ed.), Association for Computing Machinery, N.Y.

Webber B.L. (1991). Structure and ostension in the interpretation of discourse deixis. *Language and Cognitive Processes*, 6/2, 107-135.

Wertheimer M. (1923). Untersuchungen zur Lehre von der Gestalt II. *Psychologische Forschung,* 4, 301-350.

Westerstahl D. (1984). Determiners and Context Sets. In : *Generalized Quantifiers in Natural Language.* van Benthem and ter Meulen (eds.), Foris Publications, Dordrecht, Holloand.

Winograd T. (1972). *Understanding Natural Language.* Edinburgh University Press.

Winston M.E., Chaffin R., Herrmann D. (1987). Taxonomy of Part-Whole Relation. *Cognitive Science*, 11, 417-444.

Wolff F. (1998). *Analyse contextuelle des gestes de désignation en dialogue homme-machine.* Ph.D., Université H. Poincaré, Nancy.

Wolff F., De Angeli A., Romary L. (1998). Acting on a visual world : the role of perception in multimodal HCI. *Proceedings of AAAI '98, Workshop on Representations for Multi-modal Human-Computer Interaction.* Madison Wisconsin, USA.

Table des matières

Université HENRI POINCARÉ
NANCY 1

Madame **ALT – SALMON Susanne**

DOCTORAT de l'UNIVERSITE HENRI POINCARE, NANCY-I

en INFORMATIQUE

VU, APPROUVÉ ET PERMIS D'IMPRIMER

Nancy, le 11 juin 2001 n° 526

Le Président de l'Université

UNIVERSITÉ HENRI POINCARÉ, NANCY
LE PRÉSIDENT
C. BURLET

Université Henri Poincaré, Nancy-I
24-30 rue Lionnois - B.P. 3069 - 54013 NANCY CEDEX
Tél. : 03 83 85 48 00 - Fax : 03 83 85 48 48