

**Le système conceptuel, à l'interface entre
le langage, le raisonnement, et l'espace qualitatif :
vers un modèle de représentations éphémères**

thèse présentée pour obtenir le diplôme de

Docteur de l'École Polytechnique

Domaine : **Sciences Humaines**
Spécialité : **Sciences Cognitives**

par

Laleh Ghadakpour

soutenue le 28 novembre 2002, devant le jury composé de

Daniel Andler (École Normale Supérieure, Paris)
Jean-Louis Dessalles (École Nationale Supérieure des Télécommunications, Paris)
Peter Gärdenfors (Lunds Universitet)
Hans Kamp (Universität Stuttgart)
Dan Sperber (Centre Nationale de la Recherche Scientifique, Paris)
Bernard Victorri (Centre Nationale de la Recherche Scientifique, Paris)

*L'École Polytechnique n'entend donner aucune approbation, ni improbation,
aux opinions émises dans les thèses.
Ces opinions doivent être considérées comme propres à leurs auteurs.*

Ce document reprend l'essentiel des idées et des arguments que j'ai développés pendant les années de ma thèse. Je veux ici remercier tous ceux qui m'ont offert la possibilité de mener à bien ce travail et de soutenir son résultat.

Le professeur Peter Gärdenfors et le professeur Hans Kamp ont accepté spontanément d'examiner ma thèse et de participer au jury de ma soutenance à Paris. Je leur en suis très reconnaissante, et je les remercie pour la bienveillance qu'ils ont bien voulu montrer à mon égard.

Dan Sperber et Bernard Victorri ont accepté de rapporter ma thèse. Ils ont fait l'effort d'entrer dans les détails des raisonnements de ce document et d'émettre des critiques constructives qui m'aideront dans la suite de ma recherche. Je les en remercie, et j'espère pouvoir leur démontrer par la suite que je suis digne de la confiance qu'ils m'ont manifestée.

Daniel Andler a bien voulu assumer la direction de ma thèse. Je le remercie de m'avoir permis de m'engager dans cette recherche qui, au départ, pouvait sembler très ambitieuse par l'ampleur de sa problématique. Ses encouragements, lors des différentes phases de mon travail, ont été indispensables pour l'aboutissement de mon projet.

Jean-Louis Dessalles a encadré mon travail de thèse. Je lui adresse toute ma gratitude pour le temps qu'il a consacré à examiner les questions que mon travail a soulevées, et à s'engager dans les nombreuses discussions qui s'en sont suivies. Je lui sais gré de son aide précieuse pour la mise en forme du présent document, tant sur le plan de l'organisation des arguments que sur la correction de la langue.

Ma thèse a été réalisée dans le cadre de l'école doctorale de l'École Polytechnique. Je remercie Jean Petitot, directeur du Centre de Recherches en Épistémologie Appliquée, de m'avoir accueillie comme étudiante doctorante et de s'être intéressé à mon travail.

L'essentiel de mon travail a été effectué au Département Informatique et Réseaux de l'École Nationale Supérieure des Télécommunications. Je suis redevable à cette institution pour m'avoir permis de mener à bien ma recherche dans ses murs. Je garderai un très bon souvenir de l'ambiance amicale qui règne dans ce laboratoire. Je remercie chaleureusement mes collègues Talel Abdessalem, Irène Charon, Alain Grumbach, Olivier Hudry, Antoine Lobstein, et Nadine Richard pour leur accueil, leur gentillesse et leur sollicitude à mon égard. Je remercie tout particulièrement François Yvon dont l'attention et l'aide m'ont été fort précieuses.

Le bon déroulement de mon travail est également dû à la bienveillance de l'équipe de Logique de l'Université Panthéon-Sorbonne. Je remercie Susana Berestovoy et François Rivenc, qui m'ont eue jadis comme élève, et m'ont fait ensuite l'honneur de m'accepter comme collègue. Je suis heureuse de leur montrer que j'ai réussi à mener mon projet à son terme.

Je veux ici saluer mes amis, Pierre Grialou et Sarah Peyronnet, qui ont vécu avec moi les hauts et aussi les bas de mes années de thèse. Leur écoute et leur présence affectueuse ont beaucoup compté pour moi.

Je veux dire ici le soutien que m'a procuré ma famille en France, Noëlle et Mehdi Danehpajouh. Durant les années de ma thèse, ils m'ont offert protection et réconfort, et leur confiance inconditionnelle ne m'a jamais fait défaut.

La fin d'une thèse constitue une étape dans la vie de son auteur. Il est naturel que ma pensée la plus forte aille vers mes parents, Zibandeh Fadaii et Hormoz Ghadakpour, à qui je dois la vie et une part essentielle de ma personnalité. Qu'ils sachent que l'amour qu'ils me donnent continue à m'animer et me permet d'envisager l'avenir comme un défi.

Laleh Ghadakpour

Le système conceptuel, à l'interface entre le langage, le raisonnement, et l'espace qualitatif : vers un modèle de représentations éphémères

Introduction	11
Partie 1 : Le cas du temps	
Introduction	15
Chapitre 1 : Expression du temps dans le langage	
Introduction	19
1.1. La temporalité dans les langues humaines	19
1.2. La séquence de temps	22
1.3. Les schémas temporels	24
1.4. L'interprétation modale du temps	27
1.5. Les instants et les intervalles	32
1.6. Les états et les événements	46
1.7. La topologie des situations	50
1.8. Le temps, l'analogie et la métaphore	56
Conclusion	58
Chapitre 2 : La leçon de Zénon	
Introduction	61
2.1. Le dilemme de la granularité	61
2.2. Procédure de localisation temporelle	65
Conclusion	82
Conclusion	85
Partie 2 : Questions concernant la nature des concepts	
Introduction	89
Chapitre 3 : Ancrage des concepts	
Introduction	93
3.1. Les données du problème	93
3.2. Percepts intégrés	104
3.3. Symboles	107
3.4. Intentionnalité et communication	108
3.5. Intentionnalité et information	112
3.6. Le paradoxe de la connexion	116
Conclusion	118
Chapitre 4 : Rôle inférentiel des concepts	
Introduction	123
4.1. Concepts et nécessité	124
4.2. Concepts et connaissance	130
4.3. Concepts et typicité	135
Conclusion	140
Chapitre 5 : Caractère compositionnel des concepts	
Introduction	143
5.1. L'apport de la syntaxe	144
5.2. La représentation formelle du sens	149
5.3. Concepts et structures	161
5.4. Concepts et procédures	166
5.5. La représentation schématique du sens	171
Conclusion	175
Conclusion	177

Partie 3 : Le dilemme du mentalais	
Introduction	181
Chapitre 6 : Le calcul du sens	
Introduction	185
6.1. Le code du langage	185
6.2. Le jugement sémantique	186
6.3. La combinatoire du langage	190
6.4. Le langage mental	191
Conclusion	194
Chapitre 7 : Concepts moléculaires	
Introduction	199
7.1. Motifs pour structurer les concepts lexicaux	199
7.2. Méthodes pour structurer les concepts lexicaux	202
7.3. Les atouts de l'approche moléculariste	214
7.4. Les limites de l'approche moléculariste	216
Conclusion	230
Chapitre 8 : Atomisme conceptuel	
Introduction	237
8.1. Le refus des descriptions	237
8.2. Le problème de l'acquisition	239
8.3. Les représentations et les règles	242
8.4. La monotonie de la composition	245
Conclusion	246
Conclusion	249
Partie 4 : Procédure de construction du sens	
Introduction	255
Chapitre 9 : Vers un modèle de représentations conceptuelles éphémères	
Introduction	259
9.1. Concepts <i>versus</i> percepts	259
9.2. Espace qualitatif	260
9.3. Interface symbolique	263
9.4. Des structures géométriques aux formes logiques	273
9.5. Extensions du modèle	281
9.6. Bilan	286
Conclusion	289
Conclusion	291
Conclusion	293
Annexe : Contraintes d'un modèle d'apprentissage de concepts	
Les paradigmes d'apprentissage	297
De la nature des biais d'apprentissage	299
Bibliographie	303
Index	309

**Le système conceptuel,
à l'interface entre
le langage, le raisonnement,
et l'espace qualitatif :
vers un modèle de
représentations éphémères**

Introduction

Le travail présenté dans ce document s'inscrit dans le cadre de la modélisation cognitive. Le problème que nous abordons est celui de la nature des représentations sémantiques. Les individus parviennent à donner un sens aux énoncés langagiers qu'ils perçoivent ; ce sens constitue une représentation intermédiaire à partir de laquelle ils reconnaissent ou imaginent des scènes, forment des raisonnements, produisent des actions et conçoivent des réponses verbales. La modélisation cognitive se donne pour tâche d'expliquer ces performances en faisant des hypothèses minimales sur les compétences des sujets et en évitant de postuler des procédures ou des structures qui apparaîtraient comme non plausibles au regard des autres modèles de la cognition.

La question de la sémantique est une question ancienne, qui a été abordée dans une quantité de travaux, notamment en philosophie, en psychologie et en linguistique. Nous nous inspirerons d'un certain nombre de ces travaux. Nous serons amenée à critiquer certains des modèles proposés pour leur manque de plausibilité cognitive, en gardant présent à l'esprit que notre critique s'adresse non aux travaux eux-mêmes, mais à leur transposition directe à la modélisation cognitive. Prenons un exemple. Lorsque nous aborderons la question de la représentation du temps, nous mentionnerons des travaux qui offrent de représenter le temps à l'aide d'une structure dense. De tels modèles sont indispensables pour la représentation mathématique des relations temporelles véhiculées par le langage. Ils sont de même incontournables, compte tenu des techniques actuelles, lorsqu'il s'agit d'amener des systèmes d'intelligence artificielle à traiter ces mêmes relations temporelles. Cependant, nous insisterons sur le fait que de tels modèles ne sont pas utilisables tels quels en modélisation cognitive, car leur transposition directe supposerait que des structures denses soient matériellement représentées dans le cerveau humain, ce qui est strictement exclu. Notre proposition consistera à remplacer la densité structurelle par une procédure récursive qui, elle, est en principe accessible à un esprit humain.

Dans notre travail, nous présenterons la question du sens des énoncés langagiers comme un problème d'interface. Les représentations sémantiques sont supposées entretenir des rapports avec le langage, avec le raisonnement et avec la perception. Elles ont la propriété remarquable de pouvoir être évoquées par des expressions langagières, de pouvoir être le résultat d'inférences, et de pouvoir être suggérées par des situations perçues. Nous tenterons de montrer que satisfaire ces trois interfaces est une tâche difficile, au point que la plupart des modèles se limitent à expliquer l'une ou l'autre sans chercher à rendre compte de l'ensemble des potentialités des représentations sémantiques. Ainsi, certains modèles vont se concentrer sur l'émergence des représentations sémantiques à partir des données de la perception ; d'autres vont utiliser les représentations sémantiques pour expliquer les nuances exprimées par le langage ; d'autres, encore, vont tâcher de prédire les inférences que les représentations sémantiques occasionnent. Dans le premier cas, on expliquera par exemple par quel mécanisme un objet perçu sera catégorisé en tant que pomme ; dans le deuxième cas, on tentera de représenter la différence sémantique entre les mots dire et parler ; dans le troisième cas, on essaiera d'expliquer comment l'emploi du verbe tomber présuppose que le sujet de ce verbe se retrouve à un niveau inférieur.

La question fondamentale concerne la nature des représentations sémantiques. Pour la plupart des auteurs, ces représentations sont en partie indépendantes des mots, ne serait-ce que parce que les humains peuvent former beaucoup plus de pensées que les quelques dizaines de milliers de mots du lexique d'une langue. La tradition empiriste ne pose pas de différence qualitative entre les représentations sémantiques et les représentations

perceptuelles. Les autres traditions, soucieuses d'expliquer la manière dont les représentations sémantiques se combinent par le langage et par le raisonnement, renoncent à cette fusion. Elles supposent donc l'existence d'un système conceptuel autonome, dont la raison d'être est d'héberger des représentations sémantiques de nature symbolique. Ce système conceptuel entretient ainsi trois interfaces, l'une avec la perception, l'autre avec le raisonnement, et la troisième avec le langage.

Dans la majeure partie de cet exposé, nous nous placerons dans l'hypothèse d'un système conceptuel autonome, de manière à en explorer les conséquences. Si un tel système existe, il doit exister des "atomes de sens", les plus petites représentations conceptuelles qui entretiennent un rapport d'évocation avec la perception, qui soient capables de déclencher des inférences, et qui puissent se combiner en fonction des relations grammaticales présentes dans les énoncés langagiers. En particulier, il s'agira de déterminer le lien qui peut exister entre ces représentations élémentaires qu'un individu donné est supposé utiliser pour interpréter le sens des énoncés de sa langue, et les mots du lexique de cette même langue. Cette problématique nous amènera à avoir des doutes sur l'existence des représentations conceptuelles permanentes responsables du traitement sémantique des expressions langagières.

Pour commencer cette réflexion, nous nous intéresserons en premier à un problème particulier, l'expression langagière de la temporalité. Dans l'historique de notre recherche, c'est par un travail sur le temps que nous en sommes venue à développer notre critique des structures conceptuelles permanentes. Notre point de départ consiste à rappeler l'omniprésence de l'expression du temps. Nous montrerons ensuite que la majeure partie des modèles proposés, d'une manière ou d'une autre, sont réalistes par rapport au temps, c'est-à-dire qu'ils postulent la représentation mentale d'une structure temporelle globale et fixe sur laquelle agissent les différents opérateurs de construction des relations temporelles dans le langage. Notre suggestion consistera à abandonner l'idée d'une structure temporelle globale et fixe. Nous insisterons sur le fait que l'expression des relations temporelles, comme l'inclusion ou la précédence, par le langage, doit et peut être réalisée par une procédure récursive agissant sur des représentations éphémères. Elle doit être réalisée ainsi, car l'hypothèse d'une structure temporelle globale n'est pas plausible. Elle peut être réalisée ainsi, car l'aspect récursif d'une telle procédure permet de reproduire la richesse de l'expression langagière de la temporalité.

Contrairement à une attente légitime, l'extrapolation de cette idée à l'ensemble des champs sémantiques n'est pas aisée et nécessite un complément de réflexion théorique. Dans le cas du temps, il est facile de poser le problème des atomes de la structure de représentation. Dans le cas général, il faut commencer par montrer que l'existence d'atomes de sens est incontournable dès que l'on fait l'hypothèse d'un système conceptuel autonome doté de représentations permanentes symboliques. Nous allons donc commencer par cerner cette notion d'atomes de sens, avant de nous demander si leur existence est à son tour plausible. Comme dans le cas du temps, nous en viendrons à questionner le bien-fondé de l'hypothèse de l'existence de représentations conceptuelles organisées en une structure permanente et globale. Nous proposerons ensuite un modèle reposant sur la mise en œuvre d'une procédure de construction du sens dont le rôle est d'interfacer, à l'aide de représentations éphémères, l'espace qualitatif, où se structurent les données de la perception, avec le langage et le raisonnement.

Partie 1 :
Le cas du temps

Introduction

La notion du temps est constamment présente dans le langage et dans le raisonnement. La temporalité est à la base de la plupart des descriptions effectuées spontanément par les individus, et intervient dans de nombreuses argumentations. Or, l'introduction de la dimension temporelle dans la modélisation sémantique est loin d'aller de soi et a fait l'objet de plusieurs tentatives concurrentes. Cette dimension est considérée tour à tour comme un système d'indexation produisant une séquence le long du discours, comme une famille de schémas sous-jacents aux faits exprimés, comme une modalité d'évaluation vériconditionnelle, comme un ensemble d'individus logiques sujets à des quantifications et des prédications, comme une projection des paramètres d'un système dynamique sur une structure topologique, ou encore comme un domaine cognitif propre impliqué dans des rapports d'analogie et de métaphore avec d'autres domaines.

D'un point de vue cognitif, le phénomène à expliquer est spectaculaire. Les êtres humains parviennent à communiquer et à argumenter en tenant compte des relations temporelles entre les situations, en atteignant parfois une grande précision. Ils utilisent différents moyens que leur offrent les langues naturelles pour exprimer des nuances temporelles, par exemple la répétition d'une situation. Doit-on considérer qu'ils disposent pour cela de compétences cognitives propres ? Ces compétences sont-elles en rapport avec les représentations postulées dans les différents modèles de la temporalité ?

L'une des difficultés principales de nombreux modèles est qu'ils utilisent des ensembles infinis d'instantanés ou d'intervalles, ce qui pose un problème de plausibilité cognitive. Est-il correct de réifier notre capacité de conceptualisation concernant la temporalité sous la forme d'un ensemble particulier de représentations cognitives ? Cet ensemble constitue-t-il un domaine propre, ou n'est-ce qu'un système de dates et de durées attachées, comme d'autres traits, aux représentations sémantiques ? Dans les deux cas, les mécanismes qui nous permettent de nous repérer dans le temps doivent être spécifiés. L'enjeu, rappelons-le, est de comprendre la facilité, la précision et la diversité avec lesquelles les êtres humains parviennent à exprimer leur perception du temps.

Chapitre 1 :
**Expression du temps dans le
langage**

Introduction

L'expression langagière du temps est un phénomène universel. Dans toutes les sociétés humaines, les individus expriment spontanément les relations temporelles entre les situations qui font l'objet de leur discours. Cette universalité a parfois été occultée par des différences anecdotiques, par exemple des différences linguistiques (au contraire des langues européennes, certaines langues, comme le chinois, n'introduisent pas de morphologie verbale associée au temps) ou culturelles (dans au moins une culture, celle des indiens Aymaras, le passé est gestuellement projeté en avant du corps et le futur en arrière, à l'inverse des cultures européennes¹).

À côté de cette expression spontanée du temps qui s'observe dans le langage et le raisonnement, différentes cultures ont développé des systèmes temporels explicites pour répondre à des motivations religieuses, philosophiques ou scientifiques. On observe, là encore, des différences significatives (certaines cultures, par exemple, ont une conception cosmologique cyclique du temps). Cependant on retrouve dans ces différents systèmes l'idée constante selon laquelle les épisodes du temps sont, au moins localement, strictement ordonnés². L'un des exemples les plus aboutis de ce genre de construction est la dimension du temps utilisée en physique classique, qui repose sur la structure mathématique de l'ensemble des nombres réels. L'une des raisons principales qui pousse les humains à ordonner les épisodes du temps est liée à l'intuition de la causalité. Tout être humain recherche la cause d'un phénomène dans un épisode du temps situé avant le phénomène lui-même³. Tout être humain est surpris si un phénomène se produit avant sa cause. Dans une telle situation, il recherchera d'autres causes, correctement situées avant l'occurrence du phénomène.

Pour expliquer cette capacité humaine de se représenter le temps, différents modèles ont été proposés. Les uns, par exemple, partent d'une représentation temporelle constituée d'instantanés, alors que d'autres considèrent que toute représentation temporelle prend la forme d'un intervalle. D'autres modèles, encore, ne retiennent des relations temporelles que les rapports topologiques que les situations exprimées entretiennent dans le temps. Certains modèles, enfin, observent la ressemblance frappante entre le repérage temporel et le repérage spatial, de telle manière que le premier pourrait n'être qu'une métaphore du second.

Ce chapitre présente, tour à tour, différentes analyses de l'expression du temps dans le langage. Nous nous intéresserons aux approches qui proposent des propriétés générales que doit raisonnablement posséder un système de représentation pour permettre l'expression du temps.

1.1. La temporalité dans les langues humaines

La compétence humaine de communiquer et de raisonner à propos du temps est une capacité cognitive fondamentale. On ne connaît pas de langue qui n'offrirait pas de moyen

¹ Cette particularité nous a été signalée par Rafael Núñez, qui a étudié la gestuelle spontanée des Aymaras.

² Un intéressant contre-exemple, fictif bien entendu, est le cas du peuple uqbar du monde tlön décrit par Jorge Luis Borges (*Fictions* 1944).

³ L'exception des causes dites finales n'est qu'apparente. Ainsi, la "cause" du fait que la girafe naît avec un long cou est qu'un jour, elle pourra attraper les feuilles des grands arbres. Un mécanisme quelconque, que ce soit une intervention divine ou la sélection naturelle, est supposé avoir créé ce phénomène en ayant connaissance du but. La cause reste donc antérieure au phénomène, même si le but lui est ultérieur.

d'introduire la temporalité dans les énoncés⁴. Dans toute langue, un énoncé exprime une situation qui fait généralement l'objet d'une localisation temporelle. Il faut cependant observer que les formes par lesquelles la localisation temporelle des situations est exprimée varient sensiblement d'une langue à l'autre.

The idea of locating situations in time is a purely conceptual notion, and is as such potentially independent of the range of distinctions made in any particular language. It does, however, seem to be the case that all human languages have ways of locating in time. They differ from one another, however, on two parameters. The first, and overall less interesting for our present purpose, is the degree of accuracy of temporal location that is achievable in different languages. The second, and more important, is the way in which situations are located in time, in particular the relative weight assigned to the lexicon and to the grammar in establishing location in time. (COMRIE 1985 [22] p. 7)

La précision de la localisation dépend de la présence, dans une langue donnée, de certaines entités lexicales qui répondent à des besoins culturels particuliers. Les conventions de datation comme les calendriers et les systèmes horaires, les théories scientifiques et les innovations technologiques font émerger des entités lexicales qui sont propre à la culture qui produit ces inventions. Par exemple un mot comme *picoseconde*, la paraphrase de l'expression *10⁻¹² secondes*, est une invention culturelle liée aux besoins de la science et de la technologie. Les différences de précision observées dans les différentes langues ne résident toutefois pas seulement dans les détails des systèmes de datation. Certaines langues n'offrent pas de distinction entre les mots *maintenant* et *aujourd'hui*, alors qu'il s'agit d'une précision courante en français. D'autres langues possèdent un mot pour exprimer l'équivalent de l'expression *l'année dernière*, là où le français utilise un syntagme.

Des différences de précision s'observent aussi dans les mécanismes grammaticaux qu'une langue utilise pour exprimer la localisation temporelle. Par exemple, en français, les formes verbales permettent de faire la distinction entre passé et présent. Pour préciser les distances respectives de deux situations passées par rapport au présent, le français utilise la forme verbale *plus-que-parfait*, qui est censée exprimer une localisation passée antérieure à celle qu'expriment les autres formes passées. D'autres langues offrent le même type de distinction, non sous forme relative, mais sous forme explicite, par exemple en opposant, dans leur grammaire, le passé récent au passé non récent. Des distinctions plus fines peuvent séparer un passé proche, situé dans la même journée, d'un passé moins proche, situé la veille. Certaines langues peuvent ainsi offrir jusqu'à cinq distinctions grammaticales pour l'expression des différentes situations passées.

Il ne s'agit, bien entendu, que de différences dans la précision que peuvent exprimer les mots ou les morphèmes dédiés à l'expression de la temporalité. Le fait remarquable est que, malgré cette diversité apparente, les individus peuvent toujours localiser une situation de manière précise en construisant des expressions composées. Il est même impossible de poser une limite *a priori* sur la précision qui peut être ainsi atteinte dans la localisation temporelle des situations, non par les seules entités lexicales ou mécanismes grammaticaux présents dans une seule clause, mais par l'emploi itéré des moyens, lexicaux ou grammaticaux, offerts par la langue. Dans le chapitre suivant, nous reviendrons sur ce point qui nous semble crucial.

Parmi les formes permettant d'exprimer la temporalité dans une langue donnée, on distingue généralement les moyens lexicaux des moyens grammaticaux. On parle de mécanisme grammatical dès que l'indication temporelle est portée par une construction morphologique, dérivationnelle ou flexionnelle. Les formes verbales, notamment, constituent

⁴ Pour cette affirmation et tous les faits concernant les langues humaines mentionnés dans cette section, nous nous appuyons sur des travaux en linguistique comparative (COMRIE 1976 [21], COMRIE 1985 [22]).

un mécanisme puissant et potentiellement varié d'expression de la temporalité. Il ne s'agit toutefois que d'un moyen parmi d'autres dont l'usage varie. Dans certaines langues, la localisation temporelle est assurée par l'emploi d'auxiliaires ou par la présence de mots réservés appartenant à une classe fermée. D'autres langues, peu nombreuses, présentent même la particularité de n'avoir aucun mécanisme grammatical proprement dit pour la localisation temporelle, si bien que seule la présence de certaines entités lexicales indique la localisation temporelle. Cette opposition classique entre les mécanismes grammaticaux et l'emploi d'entités lexicales spécialisées n'est pas aussi tranchée qu'il y paraît. La distinction, selon nous, doit dépendre du caractère systématique de la localisation temporelle présente dans l'énoncé. Quand la localisation temporelle d'une situation est systématiquement explicitée dans l'énoncé qui l'exprime, il est possible de dire que c'est la grammaire qui est responsable de la localisation temporelle. Sinon, tout élément concernant la temporalité qui apparaît dans un énoncé de manière facultative peut être considéré comme lexical. Selon cette distinction, toutes les langues humaines possèdent des moyens lexicaux servant à la localisation temporelle, et la quasi-totalité des langues utilisent des moyens grammaticaux dans le même but.

La localisation temporelle, telle qu'elle se décline dans les différentes langues, présente deux facettes différentes, que l'on peut représenter à l'aide de deux paramètres, le temps et l'aspect.

Tense relates the time of the situation referred to to some other time, usually to the moment of speaking. [...] Aspects are different ways of viewing the internal temporal constituency of a situation.
(COMRIE 1976 [21] p. 1)

Le premier paramètre est le temps d'énoncé⁵. Il sert à établir une relation d'ordonnement temporel entre différentes situations. Dans le cas le plus simple, la relation concerne la situation exprimée par un énoncé et la situation dans laquelle l'énonciation a lieu (elle a déjeuné). Dans ce cas on parle parfois du temps absolu de l'énoncé, par opposition du temps relatif qui met en relation deux situations exprimées. Le long d'un discours, le temps d'un énoncé doit souvent se comprendre de manière relative, par rapport au temps d'un autre énoncé du même discours (elle a déjeuné ; elle a fait une sieste). Un seul énoncé peut aussi mettre en relation, de manière explicite, deux situations (après avoir déjeuné, elle a fait une sieste). En français ce paramètre de temps d'énoncé est systématiquement déterminé par la forme verbale, bien qu'une variété d'entités lexicales contribuent à préciser la relation d'ordonnement temporel des situations exprimées entre elles ou avec la situation d'énonciation (hier, elle a fait une sieste).

Le paramètre d'aspect établit la manière dont un énoncé exprime le déroulement d'une situation. Il peut se déterminer par d'autres moyens linguistiques que ceux qui déterminent le temps d'énoncé. En français, l'aspect, comme le temps d'énoncé, est présent dans la forme verbale. Une même situation peut être exprimée de manière globale (ce midi, elle a déjeuné à la cantine) ou comme une durée (ce midi, elle déjeunait à la cantine). La manière dont la situation est exprimée peut résulter d'un jeu entre la forme verbale et d'autres entités lexicales de l'énoncé. Ainsi, la forme verbale associée à l'expression des durées peut aussi exprimer une répétition (cette année, elle déjeunait à la cantine). Le paramètre d'aspect peut être aussi important pour préciser la relation entre deux situations. Deux situations peuvent être présentées comme simultanées (quand il est sorti, il a plu) ou incluses l'une dans l'autre (quand il est sorti, il pleuvait).

⁵ Le mot anglais *tense* n'ayant pas d'équivalent en français, nous précisons chaque fois "le temps d'énoncé" pour distinguer le paramètre linguistique associé de la notion générale du temps.

Le temps d'énoncé et l'aspect résument l'essentiel des relations temporelles qui peuvent s'exprimer de manière simple par le langage. Les dispositifs particuliers mis en œuvre, qu'ils soient grammaticaux ou lexicaux, varient d'une langue à l'autre, mais ils sont révélateurs des moyens dont dispose notre système cognitif pour élaborer des relations temporelles. Dans le reste de ce chapitre, nous nous intéressons à différents modèles visant à caractériser les relations temporelles que nous exprimons par le langage.

1.2. La séquence de temps

La relation temporelle la plus simple consiste à localiser une situation par rapport à une autre, en la déclarant comme antérieure, simultanée ou ultérieure. Cette relation produit une séquence où les situations sont ordonnées. Dans plusieurs langues, c'est la forme verbale qui exprime cette relation. La forme verbale nous dit, entre autres, si la situation exprimée par le verbe est antérieure, simultanée ou ultérieure à un repère temporel fourni par le contexte. Le repère le plus naturel est celui du moment de l'énonciation. Ainsi, la situation exprimée par un énoncé est localisée par rapport à la situation dans laquelle il est prononcé. Il suffit cependant d'étudier n'importe quelle langue pour constater que les relations temporelles exprimées sont bien plus riches qu'un simple positionnement par rapport au moment d'énonciation. Comment expliquer, par exemple, la différence entre le passé composé (ils ont vécu heureux) et le passé simple (il vécurent heureux) en français ? L'une des premières solutions proposées fut celle de l'introduction d'un moment de référence comme deuxième repère pour localiser la situation exprimée par un énoncé.

We see that we need three time points even for the distinction of tenses which, in a superficial consideration, seem to concern only two time points. The difficulties which grammar books have in explaining the meanings of the different tenses originate from the fact that they do not recognize the three-place structure of the time determination given in the tenses. (REICHENBACH 1947 [91] p. 289)

La structure proposée pour représenter la séquence de temps d'un énoncé postule trois composants : le point d'énonciation, le point de référence et le point d'événement (REICHENBACH 1947 [91]). Une telle structure peut par exemple expliquer la forme perfective. En anglais, les formes verbales simple et perfective du présent, par exemple les formes *goes* et *has gone* pour le verbe *to go*, ont en commun le fait que, dans les deux cas, le point de référence est localisé comme confondu avec le point d'énonciation. Ces deux formes se distinguent par le fait que, dans la forme perfective, le point d'événement est antérieur au point de référence, alors que, dans la forme simple, le point d'événement se confond avec les deux autres.

Ces trois composants introduisent un système productif qui génère plusieurs configurations, parmi lesquelles on peut retrouver les formes verbales de l'anglais. Les formes de localisation temporelle exprimées dans les autres langues humaines se retrouveraient de manière systématique parmi ces configurations (REICHENBACH 1947 [91]).

E_R_S	E_S,R	E_S_R S,E_R S_E_R
R,E_S	S,R,E	S_R,E
R_E_S R_S,E R_S_E	S,R_E	S_R_E

Dans le tableau ci-dessus les lettres S, R et E représentent respectivement les points d'énonciation, de référence et d'événement. Le signe “_” désigne l'antériorité entre deux points, et le signe “,” désigne le fait que deux points correspondants sont simultanés. La

génération se produit de manière suivante : d'abord, localiser le point de référence par rapport au point d'énonciation, ce qui donne les trois possibilités de passé (R_S), de présent (S, R) et de futur (S_R) ; puis, localiser le point d'événement par rapport au point de référence. Par exemple la configuration R, E_S signifie que le point de référence est antérieur au point d'énonciation, et que le point d'événement et le point de référence sont simultanés. Dans deux cas, la localisation du point d'événement par rapport au point de référence engendre des possibilités supplémentaires. Par exemple, dans le cas où le point de référence est antérieur au point d'énonciation (R_S) la localisation du point d'événement comme ultérieur au point de référence engendre trois configurations (R_E_S ou R_S, E ou R_S_E). Six formes parmi les formes verbales de l'anglais peuvent être identifiées par ce type de configurations (REICHENBACH 1947 [91]).

E_R_S (<i>past perfect</i>)	E_S, R (<i>present perfect</i>)	S_E_R (<i>future perfect</i>)
R, E_S (<i>past</i>)	S, R, E (<i>present</i>)	S_R, E (<i>future</i>)

Ce système constitue un début de modélisation. Il postule l'existence d'une entité abstraite, le point de référence, qui ne possède pas de contrepartie linguistique. Le point d'événement correspond à la situation exprimée par l'énoncé. On en a une trace linguistique dans le verbe. Le point d'énonciation représente le moment de l'acte d'énonciation. Il correspond, par ailleurs, au repère temporel le plus immédiat donné par le contexte. La pertinence de l'hypothèse du point de référence est plus indirecte. Son utilité se justifie dans l'explication de divers phénomènes linguistiques, comme la concordance des temps entre les formes verbales utilisées dans les énoncés qui expriment plusieurs situations.

I had mailed the letter when John came.

La forme verbale du deuxième verbe de cet énoncé est contrainte par le premier. Il est par exemple incorrect d'utiliser une forme présente perfective pour le deuxième verbe. Cette contrainte, qui impose une séquence entre les situations, peut s'expliquer par l'hypothèse de la stabilité du point de référence au sein d'un énoncé (REICHENBACH 1947 [91]). Selon le modèle ci-dessus, les deux formes verbales utilisées dans cet énoncé correspondent aux configurations suivantes.

$E1_R1_S$
 $E2, R2_S$

La contrainte vient du fait que les deux points de références $R1$ et $R2$ sont superposés, ce qui exclut, entre autres, la configuration $R2, S$. On est donc conduit à utiliser la forme de passé prétérit pour le deuxième verbe. Ainsi, la stabilité du point de référence suffit à expliquer la séquence de temps qui se traduit dans les contraintes sur l'usage des formes verbales.

La stabilité du point de référence peut toutefois être contournée si une entité lexicale impose un nouveau point de référence (REICHENBACH 1947 [91]).

He was healthier when I saw him than he is now.

La séquence de temps de cet énoncé peut être représentée par le schéma suivant, où les deux points de références $R1$ et $R2$ sont superposés, tandis que le troisième point de référence $R3$, introduit par le mot *now*, doit être considéré comme différent des deux autres.

$R1, E1_S$
 $R2, E2_S$
 $S, R3, E3$

La structure présentée ci-dessus ne peut, telle quelle, rendre compte des énoncés qui comportent une forme verbale progressive. Une première approximation consiste à remplacer certains points de la structure de départ par des périodes (REICHENBACH 1947 [91]). La notion de période correspond à une sorte d'étirement de temps. Elle exprime tantôt une durée dans laquelle une situation dure, tantôt la répétition d'une situation. Dans le premier cas, la période est un point étendu ; dans le deuxième cas, la période est une suite de points. Par exemple, on aura les configurations suivantes pour les formes progressives en anglais (REICHENBACH 1947 [91]).

EE_R_S	EE_S,R	S_EE_R
<i>(past perfect progressive)</i>	<i>(present perfect progressive)</i>	<i>(future perfect progressive)</i>
RR, EE_S	S, R \in EE	S_RR, EE
<i>(past progressive)</i>	<i>(present progressive)</i>	<i>(future progressive)</i>

Dans ce tableau, les étirements des points, les périodes, sont représentées par des doubles lettres. On voit qu'une dissymétrie apparaît dans le tableau. Là où la période d'événement est superposé au point de référence, le point de référence devient lui aussi une période. De plus, dans le cas du présent, une nouvelle relation entre les situations apparaît, celle de l'inclusion, représentée par le signe "∈". Le statut de ces étirements et des relations qui opèrent dessus n'est pas totalement clair à ce stade. Nous reviendrons plus tard sur ces relations qui viennent s'ajouter à celles de l'ordonnancement temporel.

L'intérêt premier de ce modèle est de permettre de caractériser, en n'utilisant qu'un nombre minimal d'éléments, les relations d'ordonnancement temporel qui peuvent s'exprimer à l'aide des mécanismes grammaticaux présents dans les différentes langues. Son prolongement naturel s'est réalisé dans un vaste programme de recherche visant à caractériser les contraintes temporelles imposées par les entités lexicales.

1.3. Les schémas temporels

Le modèle présenté dans la section précédente présente la forme verbale comme un moyen de localiser les situations exprimées par les verbes les uns par rapport aux autres. Il ignore un certain nombre de contraintes liées aux verbes par lesquelles les situations sont exprimées. Par exemple, toute situation n'est pas libre d'avoir un point ou une période d'événement propre.

The fact that verbs have tenses indicates that considerations involving the concept of time are relevant to their use. These considerations are not limited merely to the obvious discrimination between past, present, and future; there is another, a more subtle dependence on that concept: the use of a verb may also suggest the particular way in which that verb presupposes and involves the notion of time. (VENDLER 1967 [107] p. 97)

L'idée est que les verbes, ou les groupes verbaux contenant certains compléments, introduisent des schémas temporels qui déterminent de manière systématique l'application de telles ou telles formes verbales (VENDLER 1967 [107]). Les schémas temporels déterminent en outre l'emploi des prépositions ou des adverbes susceptibles d'accompagner le verbe. Certains projets de recherche visent donc à classifier les différents schémas temporels contenus dans le lexique.

La première dichotomie qui vient à l'esprit est celle entre les verbes qui permettent un emploi continu ou qui ne le permettent pas (VENDLER 1967 [107]). Prenons pour exemple les

deux verbes *push* et *believe* en anglais. L'acceptabilité des interactions suivantes révèle leur différence dans leur rapport avec la continuité temporelle.

What are you doing? - I am pushing it.
* What are you doing? - I am believing it.
* Do you push it? - Yes I do.
Do you believe it? - Yes I do.

La deuxième et la troisième interaction semblent mal formées. Ceci suggère que ces deux verbes n'expriment pas la continuité temporelle d'une situation de la même manière. L'idée est que le verbe *push*, de même que les verbes *run*, *draw*, *et cætera*, expriment des processus qui se déroulent dans le temps et qui consistent en phases successives qui se suivent les unes les autres. Ce n'est pas le cas du verbe *believe*, ni celui des verbes *know*, *reach*, *et cætera*. Ces verbes peuvent se rapporter à une situation sur un certain point ou pendant une certaine période, mais n'expriment en rien un déroulement quelconque. À cette première distinction liée à la présence ou l'absence d'un processus, nous pouvons ajouter d'autres discriminations.

Considérons d'abord les verbes comme *push* ou *draw* qui permettent un emploi continu. Lorsqu'ils sont attribués à un processus, certains de ces verbes seulement peuvent également s'appliquer à chaque partie propre de ce processus. Ce test permet par exemple de séparer les deux situations exprimées par les syntagmes *pushing the cart* et *drawing the circle*. La première peut être découpée en tronçons temporels qui, tous, continuent à être exprimé par l'expression *pushing the cart*. Il n'en est pas de même de la seconde. Les interactions suivantes illustrent cette idée.

For how long did you push the cart? - I was pushing it for half an hour.
* For how long did you draw the circle? - I was drawing it for twenty seconds.
* How long did it take to push the cart? - It took me half an hour to push it.
How long did it take to draw the circle? - It took me twenty seconds to draw it.

De nouveau, la deuxième et la troisième interaction semblent bizarres. La raison en est que le processus qui consiste à pousser un chariot peut être interrompu à n'importe quel moment sans changer de nature, alors que le processus qui consiste à tracer un cercle doit s'arrêter à un moment précis pour que l'on puisse dire que le processus a eu lieu. Cette distinction peut être résumée en disant que les verbes qui expriment des processus sont de deux types, les verbes d'activité et les verbes d'accomplissement (VENDLER 1967 [107]). Ainsi, le syntagme *pushing a cart* exprime une activité, tandis que le syntagme *drawing a circle* exprime un accomplissement. Il est important de noter que ce genre de trait n'est pas porté seulement par le verbe, dans la mesure où les compléments peuvent avoir un effet sur les schémas temporels. Par exemple, la forme verbale *running* peut exprimer une activité tandis que le syntagme *running a mile* exprime un accomplissement.

La deuxième famille de verbes, ceux qui, comme le verbe *believe*, n'expriment pas de processus se déroulant dans le temps, donne également lieu à une dichotomie. Parmi ces verbes, certains présentent un aspect ponctuel, tandis que les autres s'étalent sur des périodes de temps plus ou moins grandes. Prenons le cas des situations exprimées par les syntagmes *reach the top* et *believe in the stork*. La première semble concerner un moment précis du temps, tandis que la deuxième peut durer. Les interactions suivantes illustrent cette idée.

For how long did you believe in the stork? - Till I was seven.
* For how long did you reach the top? - Till noon sharp.
* At what time did you believe in the stork? - At I was seven.
At what time did you reach the top? - At noon sharp.

Cette distinction peut se formuler en distinguant les verbes d'état des verbes d'achèvement (VENDLER 1967 [107]). Ainsi, le syntagme *reach the top* exprime un achèvement, alors que le syntagme *believe in the stork* exprime un état. Certains emplois des verbes d'état ressemblent à des achèvements sans en être. Par exemple, dans l'énoncé *and then suddenly I knew the solution*, la situation exprimée semble être localisée comme un point précis. Cependant, ce n'est pas exact, car dix minutes plus tard l'individu peut toujours formuler l'énoncé *I know the solution* en exprimant ainsi le même état. L'énoncé initial exprimait plutôt une situation comme celle exprimée par l'énoncé *I did not know the solution before*.

Cette classification en quatre schémas temporels apparaît rapidement comme imparfaite. Certains verbes montrent des propriétés relatives à plusieurs schémas. Prenons les énoncés suivants.

He thinks that John is arrogant.
He is thinking about the arrogance of John.

L'usage des prépositions joue ici un rôle essentiel. Ainsi, là où le groupe verbal *think that* exprime un état, le groupe verbal *think about* présente un aspect délibéré qui en fait une activité (VENDLER 1967 [107]). Un deuxième exemple montre comment, à l'inverse, un verbe d'activité peut devenir un verbe d'état.

He is smoking a cigar.
He smokes cigars.

Ici la différence réside dans la détermination du complément. Dans le premier énoncé il s'agit d'un processus particulier, tandis que dans le cas du deuxième énoncé, il est question d'une habitude, c'est-à-dire un état de faits (VENDLER 1967 [107]). C'est l'usage du complément indéterminé qui fait que la situation exprimée par un verbe d'activité est perçue comme un état. Nous observons donc une sorte de mixité entre verbes d'état et verbes d'activité, quoique les échanges ne sont pas complètement symétriques. Dans les paires d'exemples précédentes, la manière dont le verbe change de fonction n'est pas identique. Cette différence peut être expliquée en utilisant les notions d'état générique et état spécifique (VENDLER 1967 [107]). Le verbe *smoke*, dans son emploi premier, reste une activité. Dans son emploi en tant qu'état, cette activité est distribuée sur une durée de temps. Pendant cette durée, de manière régulière, le processus exprimé par le verbe se produit. Un tel état, dit spécifique, correspond à seulement certaines périodes de cette durée. Dans le cas du groupe verbal *think that*, l'état exprimé n'est en aucune manière le résultat de la distribution sur une durée de l'activité exprimée par le groupe verbal *think about*. C'est un état générique, qui s'applique de manière globale à cette durée. Ce genre d'observation permet de restaurer la validité de la classification. Les verbes ne changent pas de catégorie par hasard, il existe une logique permettant d'expliquer le fait que dans des contextes différents, ils puissent correspondre à des schémas différents.

Cette approche lexicale de la temporalité a donné lieu à de nombreux développements en sémantique linguistique. L'objectif est de pouvoir déduire les relations temporelles entre les situations exprimées dans un énoncé non seulement à partir des marqueurs temporels morphologiques ou lexicaux, mais également à partir de la reconnaissance des schémas portés par les verbes et certains mots susceptibles d'affecter leur sens. Nous aurons toutefois l'occasion, dans la suite de notre travail, de critiquer ce type de taxonomie rigide attachée au lexique.

Dans ce qui précède, nous avons considéré les relations temporelles pour elles-mêmes, sans chercher à les coupler aux autres relations exprimées par les éléments de l'énoncé pris indépendamment de leur temporalité. Nous envisageons cette synthèse dans le cadre des structures logiques d'interprétation.

1.4. L'interprétation modale du temps

L'énoncé type, objet de l'analyse logique, est généralement atemporel. L'indication du temps qui est portée, selon la langue concernée, par des mots particuliers, des marques morphologiques ou des mécanismes syntaxiques spécialisés, doit être incorporée dans le langage atemporel de base. La question est de savoir comment modifier le système logique pour ce faire. Si l'on part de situations que l'on cherche à localiser dans le temps les unes par rapport aux autres, les situations deviennent l'objet d'une prédication temporelle. De même que pour une entité X il est possible d'avoir une prédication du genre X est rouge, on peut dire X est présent dès que X désigne une situation. De même que l'on peut exprimer des relations explicites entre deux entités comme Y contient Z , on aura des relations portant sur des situations comme Y précède Z . Les situations sont donc réifiées de manière à pouvoir entrer comme argument dans les prédications temporelles. Une telle solution, bien que naturelle au premier abord, ne va pas sans poser de problèmes.

'Is present', 'is past', etc., are only quasi-predicates, and events only quasi-subjects. ' X 's starting to be Y is past' just means 'It has been that X is starting to be Y ', and the subject here is not ' X 's starting to be Y ' but X . [...] It is X which comes to have started to be Y , [...] the other entities are superfluous, and we see how to do without them, how to stop treating them as subjects, when we see how to stop treating their temporal qualifications ('past', etc.) as predicates, by rephrasings which replace them with propositional prefixes ('It has been that', etc.) analogous to negation. (PRIOR 1967 [85] p. 18)

Cet argument commence par refuser de réifier les situations : il n'y a pas de situations, il n'y a que des entités ; il n'y a pas de relation qui concerne le temps, il n'y a que les relations atemporelles logiques. La différence d'approche ne se limite pas à ce refus. Le fait de considérer l'indication temporelle comme un préfixe propositionnel, au même titre que la négation, est porteur de nombreuses conséquences. La temporalité apparaît dans la transcription logique comme un opérateur, au même titre que les opérateurs logiques. En d'autres termes, le traitement de la dimension temporelle est renvoyé dans la procédure d'interprétation des énoncés. Les énoncés continuent à représenter des relations atemporelles entre des entités, et c'est leur évaluation sémantique qui est assujettie au temps. Une autre façon de présenter cette idée est de dire que l'on effectue un traitement intensionnel de la temporalité : les extensions des symboles logiques se déterminent en fonction du temps. Cette façon de penser la temporalité en fait une modalité. La logique temporelle apparaît ainsi comme une logique modale particulière (PRIOR 1967 [85]). Dans un modèle modal temporel, la notion de monde possible correspond à celui de l'état du monde à un moment donné et la relation d'accessibilité s'identifie à celle de l'évolution temporelle d'un état vers les états qui le suivent.

Si l'on applique cette idée à la logique d'ordre zéro, alors la valeur de vérité d'une proposition est déterminée à chaque moment de manière différente. Pour ce faire, on définit une structure $\mathbf{M} = (M, <)$, où M est l'ensemble des moments du temps, et $<$ est une relation binaire qui représente l'évolution du temps⁶. On définit ensuite une famille de valuations v_m qui déterminent la valeur de vérité des propositions à chaque moment m appartenant à l'ensemble M . La valuation v_m associe à toute proposition atomique ϕ une valeur binaire : dire que $v_m(\phi) = 1$ signifie que la proposition ϕ est vraie au moment m . L'interprétation des

⁶ Pour le moment, rien n'est dit sur la nature de ces moments et leurs relations. La caractérisation mathématique de la structure temporelle sera abordée dans la section suivante.

propositions composées se fait selon les définitions habituelles des connecteurs logiques (les connecteurs écrits en gras sont les connecteurs sémantiques définis, par exemple, à l'aide des tables de vérité).

$$\begin{aligned} v_m(\neg\varphi) &= \mathbf{non} \ v_m(\varphi) \\ v_m(\varphi \wedge \psi) &= v_m(\varphi) \ \mathbf{et} \ v_m(\psi) \\ v_m(\varphi \vee \psi) &= v_m(\varphi) \ \mathbf{ou} \ v_m(\psi) \\ v_m(\varphi \supset \psi) &= \mathbf{si} \ v_m(\varphi) \ \mathbf{alors} \ v_m(\psi) \end{aligned}$$

Ce système d'indexation de la structure d'interprétation fournit un système d'une richesse sans commune mesure avec les nuances qui peuvent être exprimées dans une langue donnée. Une manière naturelle pour se limiter à des relations plus sommaires consiste à ne retenir que les relations de type passé, présent et futur. On introduit pour cela deux nouveaux opérateurs temporels P et F. Si la proposition φ correspond à un énoncé exprimé au présent, le résultat de l'application de ces opérateurs à cette proposition, les propositions $P\varphi$ et $F\varphi$, sont censées exprimer le même énoncé, respectivement, au passé simple et au futur simple. La vérité de ces propositions à un moment m est déterminée de manière suivante.

$$\begin{aligned} v_m(P\varphi) &= 1 \text{ si et seulement s'il existe un moment } m', \text{ tel que } m' < m \text{ et } v_{m'}(\varphi) = 1 \\ v_m(F\varphi) &= 1 \text{ si et seulement s'il existe un moment } m', \text{ tel que } m < m' \text{ et } v_{m'}(\varphi) = 1 \end{aligned}$$

Les opérateurs P et F correspondent respectivement, en termes linguistiques, aux expressions il fut vrai que φ , et il sera vrai que φ . Ces opérateurs sont qualifiés d'opérateurs faibles. Il est possible de définir des opérateurs duaux H et G, dits les opérateurs forts, qui expriment le fait qu'une proposition devient définitivement vraie après le moment courant ou qu'elle a toujours été vraie avant ce moment. La vérité de ces propositions à un moment m est déterminée de manière suivante.

$$\begin{aligned} v_m(H\varphi) &= 1 \text{ si et seulement si pour tout moment } m', \text{ si } m' < m \text{ alors } v_{m'}(\varphi) = 1 \\ v_m(G\varphi) &= 1 \text{ si et seulement si pour tout moment } m', \text{ si } m < m' \text{ alors } v_{m'}(\varphi) = 1 \end{aligned}$$

Les opérateurs H et G correspondent, respectivement, aux expressions il fut toujours vrai que φ , et il sera toujours vrai que φ . On peut constater que ces nouveaux opérateurs se déduisent des précédents.

$$\begin{aligned} H\varphi &\equiv \neg P\neg\varphi \\ G\varphi &\equiv \neg F\neg\varphi \end{aligned}$$

Parfois un énoncé au présent est prononcé pour signifier que l'énoncé est vrai de manière atemporelle, qu'il s'agit d'une vérité logique. Dans le cadre de logique temporelle, ce type d'énoncé est interprété par une proposition valide, c'est-à-dire une proposition qui est vraie de toute éternité.

$$\models_{\mathcal{M}} \varphi \text{ si et seulement si pour tout moment } m, v_m(\varphi) = 1$$

Cette éternité peut s'exprimer à l'aide des opérateurs définis ci-dessus.

$$\models_{\mathcal{M}} \varphi \text{ si et seulement s'il existe un moment } m, \text{ tel que } v_m(\varphi) = 1, v_m(G\varphi) = 1 \text{ et } v_m(H\varphi) = 1$$

De même, on peut parler d'une conséquence logique si la vérité d'un énoncé résulte de la vérité d'un ensemble d'hypothèses et ceci indépendamment du moment où les hypothèses sont affirmées.

$$\Gamma \models_{\mathcal{M}} \varphi \text{ si et seulement si pour tout moment } m, \text{ si } v_m(\psi) = 1 \text{ pour toute proposition } \psi \text{ appartenant à } \Gamma \text{ alors } v_m(\varphi) = 1$$

Une fois ces définitions précisées, il faut encore ajouter aux axiomes de la logique dans laquelle on se situe de nouveaux axiomes propres aux opérateurs temporels, de manière à permettre le type d'inférences que les humains réalisent sur le temps. Ces axiomes vont par paires, l'un pour le passé, l'autre pour le futur. Les axiomes les plus naturels sont les suivants.

$$\begin{aligned} \varphi &\supset \text{HF}\varphi \\ \varphi &\supset \text{GP}\varphi \\ \text{H}(\varphi \supset \psi) &\supset (\text{H}\varphi \supset \text{H}\psi) \\ \text{G}(\varphi \supset \psi) &\supset (\text{G}\varphi \supset \text{G}\psi) \end{aligned}$$

Le deuxième axiome, par exemple, précise que si une proposition est vraie maintenant, alors elle aura toujours été vraie. Le troisième nous dit que ce qui a toujours résulté d'un fait qui a toujours été vrai, a toujours été vrai. À ces axiomes s'ajoutent deux schémas d'inférence permettant de faire le lien entre la validité (atemporelle) d'une proposition et son éternelle validité.

$$\begin{aligned} \varphi &\vdash \text{H}\varphi \\ \varphi &\vdash \text{G}\varphi \end{aligned}$$

Ces axiomes précisent que s'il existe une preuve pour la validité d'un énoncé, alors on dispose d'une preuve pour sa validité passée et future. L'ensemble de ces définitions permet de construire une logique temporelle minimale (PRIOR 1967 [85]). Dans ce cadre minimal des logiques modales, aucune supposition n'est faite sur la nature de la relation d'accessibilité. Si l'on veut préciser la nature de l'évolution temporelle que l'on imagine, il faut ajouter des axiomes portant sur les propriétés de la relation d'accessibilité. Nous en parlerons dans la section suivante à propos des instants et des intervalles.

Le système modal qui vient d'être décrit permet d'engendrer des propositions complexes censées représenter la signification temporelle des énoncés du langage naturel. Or, il se révèle souvent insuffisant. En particulier, on souhaite souvent exprimer la vérité d'une proposition entre le moment d'énonciation et le moment où une autre proposition était vraie. Il est démontré que dans certaines conceptions de l'évolution temporelle, il n'est pas possible de représenter une telle relation temporelle à l'aide des opérateurs classiques de la logique temporelle (KAMP & REYLE 1993 [56]). Pour cette raison deux nouveaux opérateurs temporels binaires S et U sont définis, que l'on peut paraphraser respectivement par les expressions ψ est vrai depuis que φ fut vraie et ψ sera vrai jusqu'à ce que φ sera vrai. Ces opérateurs, au contraire des quatre opérateurs de base, s'appliquent à un couple de propositions.

$$v_m(S\varphi\psi) = 1 \text{ si et seulement s'il existe un moment } m', \text{ tel que } m' < m \text{ et } v_{m'}(\varphi) = 1, \text{ et pour tout moment } m'', \text{ si } m' < m'' \text{ et } m'' < m \text{ alors } v_{m''}(\psi) = 1$$

$$v_m(U\varphi\psi) = 1 \text{ si et seulement s'il existe un moment } m', \text{ tel que } m < m' \text{ et } v_{m'}(\varphi) = 1, \text{ et pour tout moment } m'', \text{ si } m < m'' \text{ et } m'' < m' \text{ alors } v_{m''}(\psi) = 1$$

Ces deux opérateurs S et U suffisent pour définir une logique temporelle, car les opérateurs classiques peuvent être définis à partir d'eux.

$$\begin{aligned} \text{P}\varphi &\equiv \text{S}\varphi(\varphi \vee \neg\varphi) \\ \text{F}\varphi &\equiv \text{U}\varphi(\varphi \vee \neg\varphi) \end{aligned}$$

L'introduction de ces opérateurs augmente le pouvoir d'expressivité de la logique temporelle. Cependant, il reste des problèmes inhérents à l'interprétation du temps comme une modalité.

Le système des opérateurs tire sa richesse du fait que l'imbrication des opérateurs sur des propositions permet de produire des propositions de plus en plus complexes de manière récursive. Or, cette richesse n'est pas toujours compatible avec ce que les humains expriment avec le langage naturel. Un exemple nous est fourni par l'itération des opérateurs. L'interprétation modale du temps autorise à faire agir un nombre arbitraire d'opérateurs sur une proposition. Or, dans la plupart des cas, le résultat d'une telle itération n'a pas d'équivalent linguistique exact (KAMP & REYLE 1993 [56]). Il est clair que la proposition $PHFGPHF\phi$ ne peut pas être raisonnablement exprimée sous forme langagière. Toutefois, l'absence de correspondance entre les opérateurs itérés et les énoncés linguistiques n'est pas due qu'à des problèmes de performance humaine limitée. Certaines des itérations les plus simples apparaissent déjà comme peu pertinentes.

Une traduction plausible des propositions $FP\alpha$ ou de $FH\alpha$ est le futur antérieur. Cette forme verbale représente un fait futur, évalué comme antérieur à un référentiel, futur lui aussi (il aura terminé avant 5h / avant que tu arrives). Les deux formes logiques mentionnées couvrent, outre la période future considérée, tout le passé. Or, les instants qui précèdent le référentiel d'énonciation ne sont pas pertinents dans la situation suggérée par la forme linguistique. Ces deux formes, bien que parmi les plus simples qui peuvent être exprimées à l'aide des opérateurs de base, n'ont pas d'équivalent linguistique simple. Pour exprimer le futur antérieur sous forme logique, il faut ajouter une conjonction pour indiquer le moment où la proposition commence à être vraie, ou plutôt utiliser des combinaisons avec les opérateurs binaires qui, par nature, font le lien entre les deux propositions. De même, des propositions comme $PF\phi$ et $PG\phi$ semblent pouvoir exprimer des relations conditionnelles. Une formule comme PF (beau-père) semble fournir une traduction logique acceptable d'une expression linguistique statique comme ex-futur-beau-père. Cependant, dans ce cas comme lorsqu'il s'agit de rendre ce que nous exprimons par le mode conditionnel en français, il faut supposer une bifurcation dans l'évolution temporelle, ce qui nécessite d'ajouter des axiomes à la logique temporelle pour préciser la structure ramifiée du temps. Mentionnons encore le fait que des propositions du genre $PP\phi$, ou $PH\phi$ apportent peu par rapport à celles de $P\phi$ et $H\phi$ respectivement, car elles introduisent un décalage par rapport au moment d'énonciation sans pouvoir préciser ce décalage. Les expressions linguistiques qui utilisent par exemple la forme plus-que-parfait réalisent bien davantage : elles précisent un nouveau référentiel, ou établissent un lien temporel entre deux faits différents (il était parti avant dix heures / avant qu'elle soit arrivée). Les différentes marques linguistiques de temps, présentes dans ce type d'énoncé, expriment des relations temporelles entre des situations, ce que la superposition itérée d'opérateurs a du mal à rendre.

Ces exemples suggèrent que la possibilité d'itérer l'application des opérateurs engendre un ensemble trop riche de possibilités qui n'ont pas leur équivalent exact dans les langues naturelles, sans pour autant parvenir à rendre l'éventail des relations temporelles exprimées par le langage⁷.

La richesse d'expression temporelle du langage naturel est liée, entre autres, à l'utilisation des référentiels temporels et aux liens que ces référentiels établissent entre les propositions. Par exemple, le moment d'énonciation n'a pas pour seul rôle la détermination de la valuation des énoncés. Il est accompagné d'un contexte susceptible d'affecter l'interprétation temporelle de la forme verbale, notamment lorsque cette interprétation fait intervenir le phénomène d'indexicalité (KAMP & REYLE 1993 [56]). Considérons les interactions suivantes.

⁷ Ce problème reste entier pour la logique du premier ordre. Même si l'interprétation y est plus riche que dans la logique propositionnelle, le temps continue d'être exprimé au niveau modal.

- (1) Finalement il est parti en vacances l'été dernier ?
Non. Il a dit qu'il était malade.
- (2) Pourquoi n'est-il pas venu à la réunion d'hier ?
Il a dit qu'il était malade.
- (3) Je ne le vois pas dans la pièce. Pourquoi est-il parti ?
Il a dit qu'il était malade.

Si la proposition φ dénote l'énoncé il est malade, à partir de l'énoncé il a dit qu'il était malade, on peut, selon le contexte, inférer dans (1) que $\text{PP}\varphi$, dans (2) que $\text{P}\varphi$, et dans (3) que φ . La proposition logique que l'on peut attacher à un énoncé langagier dépend donc d'un décodage préalable des référentiels temporels. Or ceux-ci dépendent du contexte, comme le montrent les exemples précédents. De manière équivalente, l'interprétation des entités lexicales associées au temps peut dépendre de la structure temporelle de l'énoncé. Dans l'exemple qui suit, l'interprétation du mot minuit doit être déduite du temps verbal.

Il rentrera à minuit.
Il est rentré à minuit.

La définition du moment d'énonciation, essentielle à l'application des opérateurs temporels modaux, ne s'arrête pas aux limites du syntagme verbal ou de l'énoncé. Dans l'exemple suivant, le référentiel temporel est repris de manière anaphorique d'une proposition à l'autre.

Il est rentré à minuit. Il m'a appelé tout de suite.

Ainsi, le moment passé, auquel l'énoncé il est rentré fait référence, doit être utilisé pour calculer celui qui est introduit dans l'énoncé il m'a appelé. Or, ce phénomène de reprise anaphorique des référentiels temporels pose des difficultés aux logiques temporelles (KAMP & REYLE 1993 [56]). Le problème est augmenté par le fait que les relations temporelles entre propositions sont variables. Dans les exemples suivants, la permutation dans l'ordre de prononciation des énoncés change le type de relation qu'ils entretiennent entre eux. Dans le premier cas il s'agit d'une simultanéité, tandis que dans le deuxième cas une inclusion est suggérée.

Elle est sortie voir un ami. Il pleuvait.
Il pleuvait. Elle est sortie voir un ami.

Les référentiels temporels introduits par des mots contribuent pour une part essentielle à la richesse de la temporalité qui s'exprime dans le langage naturel. Or, les opérateurs de la logique temporelle sont, en quelque sorte, amnésiques. Chaque application d'un opérateur masque le référentiel temporel précédent. Il est donc impossible de le reprendre de manière anaphorique.

Dans la logique temporelle, tout est organisé pour que le temps n'intervienne que dans les opérateurs qui préfixent une proposition atemporelle. Dès que l'on veut faire intervenir le temps de manière explicite, il faut au préalable préciser la structure de la dimension temporelle utilisée. Le premier choix qu'il faut opérer est celui de l'ontologie temporelle de base, qui est généralement de type intervalle ou de type instant.

1.5. Les instants et les intervalles

Le repérage temporel prend typiquement deux formes. On peut repérer un instant par rapport à un autre instant, en précisant qu'il est situé en même temps, avant ou après. On peut aussi repérer un intervalle par rapport à un autre intervalle, en précisant qu'il est inclus dans l'autre, qu'il l'inclut, ou qu'ils se chevauchent. Ces deux formes de repérage peuvent être modélisées à l'aide des structures mathématiques, c'est-à-dire des ensembles munis de relations qui lient leurs éléments les uns aux autres.

A logical study of Time presupposes a formal pattern for its subject. [...] Time [can be] considered as a totality of temporal individuals connected by certain relations – as a structure in the model–theoretic sense, that is. [...] Both as regards temporal individuals and relations, various interesting conceptual possibilities exist.

(VAN BENTHEM 1983 [104] p. 1)

Ainsi, il est possible de caractériser les structures mathématiques adéquates aux deux formes de repérage temporel, que nous allons appeler, respectivement, des structures de points et des structures de périodes⁸.

Points

Considérons la structure $\mathbf{P}(\mathcal{P}, <)$, où \mathcal{P} est un ensemble non vide d'individus, appelés points, et $<$ est une relation binaire, appelée relation de précédence. Les intuitions que l'on peut avoir sur la nature et l'organisation des instants du temps peuvent être décrites par les propriétés de cette relation.

La relation de précédence est conçue intuitivement comme un ordre. En mathématique un ordre est avant tout caractérisé par la propriété de transitivité. Si un point précède un autre, il précède aussi tous les points que ce deuxième précède.

$$\forall p \forall q \forall r (p < q \wedge q < r) \supset p < r$$

Notre intuition sur la relation de précédence suggère une deuxième propriété, celle d'irréflexivité, selon laquelle aucun point ne se précède.

$$\forall p \neg p < p$$

La propriété d'asymétrie, selon laquelle deux points ne peuvent pas se précéder mutuellement, peut être déduite des deux propriétés précédentes.

$$\forall p \forall q p < q \supset \neg q < p$$

Ces trois propriétés caractérisent la relation de précédence comme un ordre strict partiel. Cela suffit pour représenter bon nombre d'inférences intuitives concernant l'ordonnement des instants du temps. Cependant l'intuition suggère également plusieurs autres propriétés pour la relation de précédence entre les points.

La première intuition est qu'il est possible de se déplacer au sein d'une structure de points, selon un principe appelé connexion.

⁸ Pour les définitions et les théorèmes de cette section, nous reprenons les travaux sur la variété des ontologies temporelles en théorie des modèles (VAN BENTHEM 1983 [104]).

Soit $\mathcal{P}(\mathcal{P}, <)$ une structure de points. Pour deux points quelconques p et q appartenant à l'ensemble \mathcal{P} , il existe une séquence finie de points r_1, \dots, r_k appartenant au même ensemble, telle que $r_1 = p$ et $r_k = q$, et que pour chaque i entre 1 et k , soit $r_i < r_{i+1}$, soit $r_{i+1} < r_i$.

Ainsi, l'ensemble \mathcal{P} ne peut pas être partitionné en sous-ensembles sans liens de précédence entre eux. Ce principe de connexion peut être précisé en dotant la relation de précédence de la propriété de direction.

$$\begin{aligned} \forall p \forall q \exists r \ p < r \wedge q < r \\ \forall p \forall q \exists r \ r < p \wedge r < q \end{aligned}$$

Cette propriété autorise par exemple des structures de points sous forme de graphe dans lesquelles deux points ne sont comparables que par rapport à un troisième. Une version encore plus contrainte du principe de connexion est celle où, simplement, tous les points se succèdent. Cela correspond à la chronologie de base où l'on peut toujours décider, entre deux points distincts, lequel est antérieur à l'autre. Il s'agit ici d'attribuer à la relation de précédence la propriété de linéarité.

$$\forall p \forall q \ p < q \vee p = q \vee q < p$$

Avec cette propriété, la relation de précédence devient un ordre strict total. Cette contrainte est parfois trop forte, notamment pour exprimer les conditionnels ou le futur vu comme une modalité. Le temps, dans ce cas, doit être doté d'une structure d'arbre, permettant des branchements à droite. Pour ce faire, on doit remplacer la propriété de linéarité par celle de linéarité à gauche.

$$\forall p \forall q \forall r \ (p < r \wedge q < r) \supset (p < q \vee p = q \vee q < p)$$

À l'inverse, une structure d'arbre avec des branchements à gauche, s'obtient à l'aide de la propriété de linéarité à droite.

$$\forall p \forall q \forall r \ (r < p \wedge r < q) \supset (p < q \vee p = q \vee q < p)$$

Une autre réflexion qui peut être faite à propos des structures de points concerne leur début et leur fin. Hors de toute considération cosmologique, on peut toujours concevoir, intuitivement, un point qui précède un point donné dans une structure et un autre qui le suit. Cette absence de borne peut être formulée à l'aide de la propriété de succession.

$$\begin{aligned} \forall p \exists q \ q < p \\ \forall p \exists q \ p < q \end{aligned}$$

Cette propriété confère un caractère infini aux structures de points. Une autre manière d'introduire l'infini dans une structure de points est d'affirmer qu'il existe une infinité de points entre deux points quelconques. Cela revient à doter la relation de précédence de la propriété de densité.

$$\forall p \forall q \ p < q \supset \exists r \ p < r \wedge r < q$$

Si l'on refuse une telle infinité à une structure de points, on peut au contraire considérer qu'elle n'offre qu'une précision finie. La relation de précédence peut être ainsi dotée de la propriété de discrétion.

$$\begin{aligned} \forall p \forall q \ p < q \supset \exists r \ p < r \wedge \neg \exists s \ p < s \wedge s < r \\ \forall p \forall q \ p < q \supset \exists r \ r < q \wedge \neg \exists s \ r < s \wedge s < q \end{aligned}$$

Une autre caractéristique que l'on peut souhaiter attribuer à une structure de points est que toute division exhaustive de cette structure à une partie antérieure et une partie ultérieure, suppose un point de séparation : il n'y en a pas de lacune. Il s'agit de doter la relation de précédence de la propriété de continuité.

$$\forall A (\forall p \forall q ((Ap \wedge \neg Aq) \supset p < q) \wedge \exists r Ar \wedge \exists r \neg Ar) \supset \\ (\exists s (As \vee \neg As) \wedge (\forall t s < t \supset \neg At) \wedge (\forall t t < s \supset At))$$

Dans une logique d'ordre deux, cette propriété exprime le fait que pour tout prédicat unaire A , si l'expression Ap est vraie avant d'être fausse, il existe soit un point où elle est encore vraie, mais devient fausse pour tout point ultérieur, soit un point où elle est déjà fausse, mais restait vraie pour tout point antérieur.

Outre ces propriétés qui viennent d'être mentionnées et qui concernent la relation de précédence entre les points, d'autres principes peuvent être formulés concernant les structures de points. Ces principes introduisent des morphismes sur ces structures⁹. Ils sont, au départ, indépendants de l'axiomatisation qui détermine la nature de la relation de précédence, mais il est possible de les retrouver dans certains modèles d'une telle axiomatique, notamment les modèles mathématiques dits standard.

Un premier principe exprime l'intuition selon laquelle on peut renverser la direction du temps sans en modifier la structure. Il s'agit du principe de symétrie.

Soit $\mathcal{P}(\mathbb{P}, <)$ une structure de points. Pour tout point p appartenant à l'ensemble \mathbb{P} , les deux structures $(\{q:q < p\}, <)$ et $(\{q:p > q\}, >)$, où la relation $>$ est exactement l'opposée de la relation de précédence $<$, sont isomorphes.

En d'autres termes, la structure des points ultérieurs à un point donné est identique à la structure des points antérieurs à ce point. Le principe de symétrie impose évidemment des conditions sur le nombre de points dans les deux directions. Par exemple, la propriété de succession bilatérale de la relation de précédence rend la symétrie possible. On peut formuler un deuxième principe plus fort, celui d'isotropie.

Soit $\mathcal{P}(\mathbb{P}, <)$ une structure de points. Pour tout point p appartenant à l'ensemble \mathbb{P} , les deux structures $(\{q:q < p\}, <)$ et $(\{q:p < q\}, <)$ sont isomorphes.

Ce principe stipule que, vu d'un point, les structures du passé et du futur sont superposables par une opération de projection, sans renversement du temps. Ce principe est rendu possible par les propriétés de succession et de densité de la relation de précédence.

Une autre intuition concerne les projections d'une structure de points sur elle-même. Cette idée est formulée par le principe d'homogénéité.

Soit $\mathcal{P}(\mathbb{P}, <)$ une structure de points. Pour deux points quelconques p et q appartenant à l'ensemble \mathbb{P} , il existe un automorphisme de \mathcal{P} qui projette p sur q .

Autrement dit, on peut projeter n'importe quel point sur n'importe quel autre en conservant la structure globale intacte. Un principe encore plus fort, concernant la projection d'une structure de points sur ses parties propres, est celui de réflexion.

⁹ Les morphismes sont des applications établissant des liens entre les éléments d'ensembles munis de relations. Ils appliquent les individus d'un ensemble sur ceux d'un autre, en conservant les relations analogues établies entre ces individus dans les deux structures. Deux structures analogues sont considérées comme isomorphes s'il existe entre elles un morphisme bijectif, dont l'application réciproque est aussi un morphisme. On appelle automorphisme les isomorphismes d'un ensemble structuré sur lui-même.

Soit $\mathcal{P}(\mathcal{P}, <)$ une structure de points. Pour deux points quelconques p et q appartenant à l'ensemble \mathcal{P} , tels que $p < q$, il existe un isomorphisme entre \mathcal{P} et sa sous-structure $(\{r : p < r \wedge r < q\}, <)$.

Les structures de points fournissent un système de dates qui permet de représenter certains aspects temporels des énoncés du langage. Cependant, de nombreux énoncés portent, directement ou implicitement, sur des durées. Il est donc naturel de s'intéresser à des structures dans lesquelles les intervalles constituent les entités de base.

Périodes

Considérons la structure $\mathcal{P}(\mathcal{P}, \subseteq, <)$, où \mathcal{P} est un ensemble non vide d'individus, appelés périodes, \subseteq est une relation binaire, appelée relation d'inclusion, et $<$ est une relation binaire, appelée relation de précédence. Les intuitions que l'on peut avoir sur la nature et l'organisation des intervalles du temps peuvent être décrites par les propriétés de ces deux relations.

Considérons tout d'abord les propriétés qui doivent être attribuées à la relation d'inclusion. Cette relation représente l'intuition selon laquelle les périodes, au contraire des points, peuvent s'inclure les unes les autres. Ceci crée une hiérarchie d'inclusion qui correspond à un ordre. Ainsi, la relation d'inclusion est, à son tour, caractérisée par la propriété de transitivité.

$$\forall p \forall q \forall r (p \subseteq q \wedge q \subseteq r) \supset p \subseteq r$$

À la différence de la relation de précédence, la relation d'inclusion n'est pas un ordre strict. On peut sans problème concevoir qu'une période s'inclut elle-même. La relation d'inclusion doit donc posséder la propriété de réflexivité.

$$\forall p p \subseteq p$$

En acceptant la propriété de réflexivité, on perd celle d'asymétrie. Pour conserver l'idée de l'absence de symétrie dans l'ordre, il faut attribuer une troisième propriété à la relation d'inclusion, celle d'anti-symétrie.

$$\forall p \forall q (p \subseteq q \wedge q \subseteq p) \supset p = q$$

Ainsi, seules les périodes identiques peuvent s'inclure mutuellement. Les trois propriétés de transitivité, de réflexivité, et d'anti-symétrie font de la relation d'inclusion un ordre large partiel.

À partir de cet ordre, on peut définir deux autres relations entre les périodes. La première est la relation de chevauchement. Deux périodes se chevauchent lorsqu'il existe une troisième période qui est incluse dans les deux.

$$\forall p \forall q p \circ q \stackrel{\text{def}}{=} \exists r r \subseteq p \wedge r \subseteq q$$

Intuitivement, la relation de chevauchement entre deux périodes définit une troisième période. Cette opération devient possible si l'on attribue à la relation d'inclusion la propriété de conjonction.

$$\forall p \forall q p \circ q \supset \exists r r \subseteq p \wedge r \subseteq q \wedge \forall s (s \subseteq p \wedge s \subseteq q) \supset s \subseteq r$$

Autrement dit, le chevauchement de deux périodes définit une plus grande période qui est incluse dans ces deux périodes. Avec cette propriété, la relation d'inclusion peut donner lieu à une opération d'intersection semblable à l'intersection ensembliste. Cette opération aura les propriétés d'absorption, de commutativité et d'associativité. Une différence notable avec

l'intersection ensembliste est que l'intersection entre deux périodes qui ne se chevauchent pas n'est pas définie, car il n'y a pas de période nulle semblable à l'ensemble vide. Il s'agit donc d'une opération d'intersection partielle.

$$\forall p \forall q \forall r \ r = p \cap q \equiv_{\text{def}} r \subseteq p \wedge r \subseteq q \wedge \forall s \ (s \subseteq p \wedge s \subseteq q) \supset s \subseteq r$$

Une autre propriété importante à formuler sur la relation d'inclusion prescrit que si toutes les parties d'une période chevauchent une autre période, alors la première est incluse dans la deuxième. Il s'agit de la propriété de liberté.

$$\forall p \forall q \ (\forall r \ r \subseteq p \supset r \subseteq q) \supset p \subseteq q$$

Cette propriété confère une constituance par rapport aux périodes incluses : si une période p n'est pas incluse dans une deuxième période q , alors il existe une période r , incluse dans p , responsable de ce fait, c'est-à-dire que r ne chevauche pas q .

À partir de l'ordre large partiel défini par la relation d'inclusion, on peut définir une deuxième relation entre les périodes, la relation de recouvrement. Cette relation impose qu'il existe une troisième période qui recouvre les deux périodes en relation.

$$p \cup q \equiv_{\text{def}} \exists r \ p \subseteq r \wedge q \subseteq r$$

De manière parallèle à ce qui a été dit pour la relation de chevauchement, la relation de recouvrement entre deux périodes peut définir une troisième période. Pour rendre possible une telle opération, la relation d'inclusion doit être dotée de la propriété de disjonction.

$$\forall p \forall q \ p \cup q \supset \exists r \ p \subseteq r \wedge q \subseteq r \wedge \forall s \ (p \subseteq s \wedge q \subseteq s) \supset r \subseteq s$$

Autrement dit, le recouvrement de deux périodes définit une plus petite période qui inclut les deux premières. Cette propriété de la relation d'inclusion peut donner lieu, à son tour, à une opération d'union.

$$\forall p \forall q \forall r \ r = p \cup q \equiv_{\text{def}} p \subseteq r \wedge q \subseteq r \wedge \forall s \ (p \subseteq s \wedge q \subseteq s) \supset r \subseteq s$$

Cette opération ressemble à l'union ensembliste, avec la différence que, dans le cas des périodes, le résultat peut inclure des périodes qui ne sont pas incluses dans les deux périodes de départ. Il est possible de définir une nouvelle opération, celle de somme, qui exclut une telle possibilité.

$$\forall p \forall q \forall r \ r = p + q \equiv_{\text{def}} r = p \cup q \wedge \forall s \ s \subseteq r \supset (s \subseteq p \vee s \subseteq q)$$

Dans cette nouvelle construction, la période r se découpe exhaustivement en deux parties correspondant aux périodes p et q . Les deux opérations d'union et de somme ont, à leur tour, les propriétés d'absorption, de commutativité et d'associativité.

Contrairement à la relation de chevauchement, la relation de recouvrement peut être conçue entre n'importe quel couple de périodes. Pour le permettre, il suffit de doter la relation d'inclusion de la propriété de direction.

$$\forall p \forall q \ p \cup q$$

Cette propriété peut donner un caractère arbitrairement grand à la structure des périodes, c'est-à-dire qu'il est possible de construire des périodes de plus en plus grandes. De manière inverse, on peut dire qu'il y a toujours des périodes de plus en plus petites. Cette idée suggère la propriété de descente infinie.

$$\forall p \forall q \ q \subseteq p \supset \exists r \ r \subseteq q \wedge \neg r = q$$

Il est parfaitement possible de refuser la propriété précédente. On peut très bien imaginer que les périodes ne peuvent pas être plus petites qu'un grain défini. Dans ce cas la relation d'inclusion sera dotée de la propriété d'atomicité.

$$\forall p \exists q q \subseteq p \wedge \forall r r \subseteq q \supset r = q$$

Considérons maintenant la relation de précédence. Intuitivement, il s'agit d'un ordre strict partiel, c'est-à-dire une relation possédant les mêmes propriétés de transitivité, d'irréflexivité, et d'asymétrie que la précédence entre les points. L'absence de borne dans une structure de périodes peut être caractérisée, comme dans le cas des points, par les propriétés de succession de la relation de précédence. De même, la propriété de direction de la relation de précédence, comme dans le cas des points, peut rendre possible le principe de connexion pour une structure de périodes. La propriété de linéarité, quant à elle, se formule par une propriété mixte qui lie les deux relations d'inclusion (sous la forme de la relation de chevauchement) et de précédence.

$$\forall p \forall q p < q \vee q < p \vee p \circ q$$

On peut voir ainsi dans la relation de précédence l'équivalent d'un ordre strict total, puisque toutes les périodes sont, sauf quand elles se chevauchent, nécessairement comparables. Une exigence plus forte peut être formulée pour une structure de périodes, qui précise que les relations de précédence et de chevauchement, et *a fortiori* d'inclusion, sont incompatibles. Les relations de précédence et d'inclusion peuvent ainsi être liées par la propriété de séparation.

$$\forall p \forall q p < q \supset \neg p \circ q$$

L'alternative entre un ordre dense ou discret pour les points trouve un parallèle dans les propriétés de descente infinie ou d'atomicité de la relation d'inclusion, qui déterminent l'existence ou non de périodes de plus en plus petites. En ce qui concerne la relation de précédence, son caractère discret peut s'exprimer par la propriété de voisinage : toute période possède deux voisins immédiats.

$$\forall p \forall q p < q \supset \exists r p < r \wedge \neg \exists s p < s \wedge s < r$$

$$\forall p \forall q p < q \supset \exists r r < q \wedge \neg \exists s r < s \wedge s < q$$

Cette propriété correspond assez bien à l'intuition que l'on peut avoir de l'agencement des intervalles du temps. Le choix inverse, qui confère un caractère dense à la relation de précédence, semble contre-intuitif. Il suppose qu'il existe une troisième période entre deux périodes quelconques. Un tel choix demande une réflexion supplémentaire sur la notion même de voisinage¹⁰. Il existe une manière plus naturelle d'exprimer une propriété de densité en rapport avec la relation de précédence, en la liant à la relation d'inclusion.

$$\forall p \exists q \exists r p = q \cup r \wedge q < r$$

Cette propriété stipule que toute période p possède une aile gauche et une aile droite, représentées respectivement par les périodes q et r . Il s'agit d'une propriété mixte, où la relation d'inclusion intervient par le biais de l'opération d'union. Noter que cette propriété de densité autorise la présence d'un fossé entre les deux périodes q et r . Pour exclure une telle possibilité, on peut formuler la propriété de densité renforcée, où la relation d'inclusion intervient par le biais de l'opération de somme.

$$\forall p \exists q \exists r p = q + r \wedge q < r$$

¹⁰ Nous reviendrons plus loin sur la notion topologique de voisinage et sa pertinence au sujet de la modélisation des relations temporelles exprimées par les énoncés langagiers.

Cette nouvelle propriété suppose que toute période p peut être découpée de manière exhaustive en deux périodes q et r qui sont immédiatement voisines l'une de l'autre.

Les relations de précédence et d'inclusion peuvent aussi être liées pour renforcer l'idée exprimée par la propriété de liberté. Il s'agit d'affirmer que si une période p ne précède pas une deuxième période q , alors il existe deux périodes r et s , respectivement incluses dans p et dans q , responsables de ce fait, c'est-à-dire qu'il n'existe aucune période incluse dans r précédant une période incluse dans s . En ajoutant cette propriété à la propriété de liberté définie ci-dessus, nous obtenons la propriété de liberté renforcée.

$$\forall p \forall q \neg p < q \supset \exists r \exists s r \subseteq p \wedge s \subseteq q \wedge \forall t \forall u (t \subseteq r \wedge u \subseteq s) \supset \neg t < u$$

Une autre propriété mixte est celle qui impose à la relation de précédence de se propager aux périodes incluses dans une période donnée. Il s'agit de la propriété de monotonie.

$$\begin{aligned} \forall p \forall q p < q \supset \forall r r \subseteq p \supset r < q \\ \forall p \forall q p < q \supset \forall r r \subseteq q \supset p < r \end{aligned}$$

La dernière propriété mixte que l'on peut formuler est celle de convexité, qui stipule que si une période en recouvre deux autres, elle recouvre également toutes les périodes intermédiaires.

$$\forall p \forall q \forall r (p < q \wedge q < r) \supset \forall s (p \subseteq s \wedge r \subseteq s) \supset q \subseteq s$$

Enfin, de même que nous avons des principes concernant les morphismes dans les structures de points, nous pouvons formuler ici le principe de réflexion pour une structure de périodes.

Soit $\mathcal{P}(\mathcal{P}, \subseteq, <)$ une structure de périodes. Pour toute période p appartenant à l'ensemble \mathcal{P} , la restriction de \mathcal{P} à l'ensemble des sous-périodes de p est isomorphe à \mathcal{P} .

Il résulte de ce qui précède que les points et les périodes sont deux moyens cohérents de représenter le temps. On pourrait s'en étonner, tant l'intuition qui sous-tend ces deux approches est différente : les points sont censés représenter des dates, alors que les périodes représentent des durées. Il peut être réconfortant pour l'intuition de savoir qu'il existe plusieurs moyens d'établir une équivalence entre les deux approches.

Points versus Périodes

Même si instants et intervalles apparaissent comme de nature radicalement différente, il existe un lien entre les deux : l'intuition mathématique suggère qu'un intervalle peut apparaître comme un ensemble constitué d'instants et, à l'inverse, qu'un instant peut être vu comme ce qui est commun à un ensemble d'intervalles emboîtés. Cette intuition est confirmée par le fait qu'il est possible de définir les structures de points et les structures de périodes les unes à partir des autres.

Pour exposer cette correspondance, nous nous basons sur les définitions minimales suivantes pour les structures de points et les structures de périodes.

Une structure de points $\mathcal{P}(\mathcal{P}, <)$ est définie par la donnée d'un ensemble non vide \mathcal{P} muni d'une relation d'ordre strict partiel $<$ appelée la relation de précédence.

Une structure de périodes $\mathcal{P}(\mathcal{P}, \subseteq, <)$ est définie par la donnée d'un ensemble non vide \mathcal{P} muni d'une relation d'ordre strict partiel $<$, appelée la relation de précédence, et d'une relation d'ordre large partiel, appelée la relation d'inclusion, définissant une opération d'intersection partielle, ces deux relations étant liées par la propriété mixte de monotonie.

La manière la plus simple de définir des périodes à partir des points consiste à considérer les premiers comme des ensembles des seconds. Cependant, tout ensemble de points ne possède pas les propriétés suffisantes pour servir de base à la définition d'une période. Une de ces propriétés est celle de la convexité.

Soit $\mathcal{P}(\mathcal{P}, <)$ une structure de points. Un sous-ensemble c de l'ensemble \mathcal{P} est convexe, si et seulement si pour deux points quelconques p et q appartenant à c , et pour tout point r appartenant à \mathcal{P} , si $p < r$ et $r < q$, alors r appartient à c .

Pour tout sous-ensemble d'une structure d'ordre strict partiel, il existe une clôture convexe. Cette clôture s'obtient par une procédure qui consiste à ajouter à ce sous-ensemble tous les éléments intermédiaires entre chaque paire de ces éléments.

Ainsi, à partir d'une structure de points, il est possible de définir une structure de périodes dont les éléments ne sont que des sous-ensembles convexes non vides de cette structure de points.

Soit $\mathcal{P}(\mathcal{P}, <)$ une structure de points. La structure de périodes convexes $\mathcal{CP}(\mathcal{C}, \subseteq_c, <_c)$ introduite par \mathcal{P} est définie par la donnée de \mathcal{C} , l'ensemble des sous-ensembles convexes non vides de l'ensemble \mathcal{P} , de la relation d'inclusion ensembliste \subseteq_c , et d'une relation de précédence $<_c$. Cette relation est telle que $c <_c d$ pour deux sous-ensembles convexes quelconques c et d appartenant à \mathcal{C} , si et seulement si pour deux points quelconques p et q appartenant respectivement à c et à d , $p < q$.

On peut vérifier que \mathcal{CP} satisfait les conditions d'une structure de périodes définies ci-dessus. La relation d'ordre large est donnée par l'inclusion ensembliste entre sous-ensembles convexes. De même, l'intersection partielle est issue de l'intersection ensembliste restreinte aux résultats non vides, par le fait que l'intersection de deux ensembles convexes est convexe. La transitivité et l'irréflexivité de la relation $<_c$ s'héritent de celles de la relation $<$. La propriété de monotonie découle naturellement du fait que la relation \subseteq_c correspond à l'inclusion ensembliste. De plus, on remarque que \mathcal{CP} possède la propriété de convexité.

Construire une structure de périodes à partir des sous-ensembles convexes de points est donc possible, mais l'exigence de convexité n'est pas nécessaire. Il suffit que soit garantie l'existence d'une opération partielle d'intersection. Ceci peut être obtenu directement par la définition d'une structure d'intervalles de points.

Soit $\mathcal{P}(\mathcal{P}, <)$ une structure de points. Une structure d'intervalles de points $\mathcal{IP}(\mathcal{I}, \subseteq_i, <_i)$ introduite par \mathcal{P} est définie par la donnée de \mathcal{I} , un ensemble de sous-ensembles non vides de l'ensemble \mathcal{P} , clos par la formation d'intersections ensemblistes non vides, de la relation d'inclusion ensembliste \subseteq_i , et d'une relation de précédence $<_i$. Cette relation est telle que $i <_i j$ pour deux intervalles quelconques i et j appartenant à l'ensemble \mathcal{I} , si et seulement si pour deux points quelconques p et q , appartenant respectivement à i et à j , $p < q$.

On vérifie facilement que \mathcal{IP} satisfait les conditions d'une structure de périodes définies ci-dessus. L'existence d'une opération partielle d'intersection vient de la clôture de

l'ensemble \mathbb{I} qui fait que l'intersection ensembliste non vide de deux éléments de cet ensemble lui appartient aussi.

L'opération inverse, qui consiste à définir une structure de points à partir d'une structure de périodes, n'est pas aussi facile. Une première idée consiste à définir les points comme les plus petites périodes. C'est évidemment possible quand la structure de périodes possède la propriété d'atomicité. Ainsi, à partir d'une structure de périodes, il est possible de définir une structure de points dont les éléments ne sont que des atomes de cette structure de périodes.

Soit $\mathbf{P}(\mathbb{P}, \subseteq, <)$ une structure de périodes dotée de la propriété d'atomicité. La structure d'atomes de périodes $\mathbf{AP}(\mathbb{A}, <_a)$ introduite par \mathbf{P} est définie par la donnée de \mathbb{A} , l'ensemble des atomes de l'ensemble \mathbb{P} , et d'une relation de précédence $<_a$ qui est la restriction de la relation $<$ de \mathbb{P} à \mathbb{A} .

On vérifie facilement que \mathbf{AP} est une structure de points. Il suffit d'observer que la relation $<_a$ est une relation d'ordre strict partiel, ce qui provient du fait qu'il s'agit de la restriction d'une relation d'ordre strict partiel.

De plus, on peut vérifier qu'une structure d'atomes de périodes offre une représentation du temps qui, sous certaines conditions, est la même que celle de la structure de périodes qui l'introduit. Pour ce faire, il suffit de l'utiliser pour reconstituer une nouvelle structure de périodes, et montrer que cette dernière structure est isomorphe à la structure de périodes initiale. Il faut toutefois que celle-ci possède la propriété de liberté renforcée.

Soit $\mathbf{AP}(\mathbb{A}, <_a)$ une structure d'atomes de périodes introduite par une structure de périodes $\mathbf{P}(\mathbb{P}, \subseteq, <)$ dotée des propriétés d'atomicité et de liberté renforcée. La structure d'intervalles d'atomes $\mathbf{IAP}(\mathbb{I}, \subseteq_i, <_i)$ introduite par \mathbf{AP} est définie par la donnée de \mathbb{I} , l'ensemble de toutes les projections $i = \text{atomes}(p)$, où $\text{atomes}(p)$ est l'ensemble de tous les atomes inclus dans une période quelconque p appartenant à l'ensemble \mathbb{P} , de la relation d'inclusion ensembliste \subseteq_i , et d'une relation de précédence $<_i$. Cette relation est telle que $i <_i j$ pour deux intervalles quelconques i et j appartenant à l'ensemble \mathbb{I} , si et seulement si pour deux atomes quelconques a et b , appartenant respectivement à i et à j , $a <_a b$.

Cette définition produit bien une structure d'intervalles de points, car l'ensemble \mathbb{I} , ainsi défini, est clos pour la formation d'intersections ensemblistes. Pour deux périodes quelconques p et q appartenant à l'ensemble \mathbb{P} , si l'intersection ensembliste des deux intervalles $i = \text{atomes}(p)$ et $j = \text{atomes}(q)$ existe, alors il existe au moins un atome qui est inclus à la fois dans p et dans q , si bien que l'intersection entre p et q existe. Il est facile de vérifier que l'intervalle correspondant aux atomes inclus dans la période $p \cap q$ est équivalent à l'intersection ensembliste entre i et j . Donc cette dernière appartient à \mathbb{I} .

Pour que la structure de périodes ainsi obtenue soit isomorphe à la structure de départ, il faut premièrement que la projection atomes soit injective, c'est-à-dire que pour deux périodes quelconques différentes p et q appartenant à l'ensemble \mathbb{P} , les projections $\text{atomes}(p)$ et $\text{atomes}(q)$ soient deux intervalles différents appartenant à l'ensemble \mathbb{I} . Ceci nécessite que la relation \subseteq possède au moins la propriété de liberté. La propriété d'anti-symétrie implique que $\neg p \subseteq q$ ou $\neg q \subseteq p$. La propriété de liberté, elle, implique que si $\neg p \subseteq q$, alors il existe des périodes incluses dans p qui ne chevauchent pas q , et donc qu'il existe des atomes inclus dans p qui ne sont pas inclus dans q , si bien que l'on a $\neg \text{atomes}(p) = \text{atomes}(q)$. Un argument équivalent peut être formulé pour le cas où $\neg q \subseteq p$. Dans ce raisonnement, la

propriété de liberté permet de considérer les atomes comme les constituants des périodes de \mathcal{P} , ce qui permet de les faire correspondre aux éléments appartenant aux intervalles de \mathbf{IAP} .

Dans un deuxième temps, il faut que la projection atomes engendre une relation \subseteq_i isomorphe à la relation \subseteq . La transitivité de la relation d'inclusion \subseteq suffit pour démontrer que, pour deux périodes quelconques p et q , si $p \subseteq q$, alors $\text{atomes}(p) \subseteq_i \text{atomes}(q)$. Pour démontrer la clause réciproque, c'est-à-dire, pour deux périodes quelconques p et q , si $\text{atomes}(p) \subseteq_i \text{atomes}(q)$ alors $p \subseteq q$, il faut faire intervenir la propriété de liberté comme précédemment.

Dans un troisième temps, il faut que la projection atomes engendre une relation $<_i$ isomorphe à la relation $<$. La propriété de monotonie suffit pour démontrer que, pour deux périodes quelconques p et q , si $p < q$ alors $\text{atomes}(p) <_i \text{atomes}(q)$. La démonstration de la clause réciproque, à savoir, le fait que pour deux périodes quelconques p et q , si $\text{atomes}(p) <_i \text{atomes}(q)$, alors $p < q$, nécessite l'usage de la deuxième partie de la propriété de liberté renforcée. Supposons que p ne précède pas q . Alors il existe deux périodes r et s , telles que $r \subseteq p$ et $s \subseteq q$, et il n'existe aucune période incluse dans r précédant une période incluse dans s . Or deux telles périodes ne peuvent pas exister car tous les atomes de la période p précèdent les atomes de la période q .

La construction qui précède permet de montrer qu'il existe une structure de points équivalente à toute structure de périodes sous certaines conditions, notamment que la structure de périodes possède la propriété d'atomicité. Or, cette dernière exigence peut être évitée. Pour définir des points à partir de périodes dépourvues d'atomes, on utilise la notion de filtre.

Soit $\mathcal{P}(\mathcal{P}, \subseteq, <)$ une structure de périodes. Un sous-ensemble f de l'ensemble \mathcal{P} est un filtre, si et seulement si pour toute période p appartenant à l'ensemble f , toutes les périodes q , telles que $p \subseteq q$, appartiennent à f , et pour deux périodes quelconques p et q appartenant à f , la période $p \cap q$ existe et appartient à f .

Notamment, il est possible de définir un filtre f_p pour toute période p comme l'ensemble de toutes les périodes q telles que $p \subseteq q$. Il est intéressant de constater que les filtres présentent la propriété de décomposition, c'est-à-dire que pour tout filtre f , deux périodes quelconques p et q appartiennent à ce filtre si et seulement si la période $p \cap q$ lui appartient.

Ainsi, à partir d'une structure de périodes, il est possible de définir une structure de points dont les éléments ne sont que des filtres de cette structure de périodes, sans qu'aucune propriété supplémentaire ne soit nécessaire pour celle-ci.

Soit $\mathcal{P}(\mathcal{P}, \subseteq, <)$ une structure de périodes. La structure de filtres de périodes $\mathbf{FP}(\mathcal{F}, <_f)$ introduite par \mathcal{P} est définie par la donnée de \mathcal{F} , l'ensemble des filtres non vides de l'ensemble \mathcal{P} , et d'une relation de précédence $<_f$ telle que $f <_f g$ pour deux filtres quelconques f et g appartenant à \mathcal{F} , si et seulement s'il existe deux périodes p et q appartenant respectivement à f et à g , telles que $p < q$.

Pour vérifier que \mathbf{FP} est une structure de points, il suffit d'observer que la relation $<_f$ est une relation d'ordre strict partiel. Supposons, que $f <_f g$ et $g <_f h$, pour trois filtres quelconques f et g et h . On aura $p < q$ et $r < s$, où les périodes q et r appartiennent à g , et les périodes p et s appartiennent respectivement à f et à h . Or, par la propriété de monotonie, p précède la période $q \cap r$, qui, à son tour, précède s . Par transitivité de la relation $<$, p précède donc s , si bien que $f <_f h$. En outre, si $f <_f f$, alors deux périodes p et q

appartenant à f sont telles que $p < q$. Par la propriété de monotonie, la période $p \cap q$ se précède elle-même, ce qui n'est pas possible en raison d'irréflexivité de la relation $<$. Donc la relation $<_f$ est bien un ordre strict.

Pour montrer que la structure de filtres de périodes représente bien la même structure temporelle que la structure de périodes de départ, le mieux est de montrer que l'on peut reconstituer, à partir de la première, une structure isomorphe à la deuxième.

Soit $\mathbf{FP}(F, <_f)$ une structure de filtres de périodes introduite par une structure de périodes $\mathbf{P}(P, \subseteq, <)$. La structure d'intervalles de filtres $\mathbf{IFP}(I, \subseteq_i, <_i)$ introduite par \mathbf{FP} est définie par la donnée de I , l'ensemble de toutes les projections $i = \text{filtres}(p)$, où $\text{filtres}(p)$ est l'ensemble de tous les filtres auxquels appartient une période quelconque p appartenant à l'ensemble P , de la relation d'inclusion ensembliste \subseteq_i , et d'une relation de précédence $<_i$. Cette relation est telle que $i <_i j$ pour deux intervalles quelconques i et j appartenant à l'ensemble I , si et seulement si pour deux filtres quelconques f et g appartenant respectivement à i et à j , $f <_f g$.

Pour vérifier que \mathbf{IFP} est bien une structure d'intervalles de points, il faut montrer que l'ensemble I contient toutes les intersections ensemblistes non vides de ses éléments. Pour deux périodes quelconques p et q appartenant à l'ensemble P , si l'intersection des intervalles deux $i = \text{filtres}(p)$ et $j = \text{filtres}(q)$ existe, alors il existe un filtre qui contient p et q . Ce filtre contient donc aussi la période $p \cap q$. Par la propriété de décomposition, on vérifie facilement que l'intervalle $\text{filtres}(p \cap q)$ est équivalent à l'intersection ensembliste entre i et j . Donc cette dernière appartient à I .

Pour qu'il y ait isomorphisme, il faut que premièrement que la projection filtres soit injective. Pour deux périodes quelconques différentes p et q appartenant à l'ensemble P , les intervalles associés, $i = \text{filtres}(p)$ et $j = \text{filtres}(q)$, doivent être différents. La propriété d'anti-symétrie implique que $\neg p \subseteq q$ ou $\neg q \subseteq p$. Or si $\neg p \subseteq q$, alors on connaît au moins un filtre auquel appartient p , à savoir le filtre f_p , et auquel q n'appartient pas. Un argument équivalent peut être formulé pour le cas où $\neg q \subseteq p$. Donc les intervalles i et j sont différents.

Il faut ensuite que la projection filtres engendre une relation \subseteq_i isomorphe à la relation \subseteq . Par définition des filtres, pour deux périodes quelconques p et q , si $p \subseteq q$ alors $\text{filtres}(p) \subseteq_i \text{filtres}(q)$. L'existence des filtres f_p et f_q suffit pour montrer que, pour deux périodes quelconques p et q , si $\text{filtres}(p) \subseteq_i \text{filtres}(q)$, alors $p \subseteq q$.

Enfin, il faut que la projection filtres engendre une relation $<_i$ isomorphe à la relation $<$. Pour deux périodes quelconques p et q , si $p < q$, alors par la définition des relations $<_f$ et $<_i$, $\text{filtres}(p) <_i \text{filtres}(q)$. Inversement, pour deux périodes quelconques p et q , si $\text{filtres}(p) <_i \text{filtres}(q)$, alors en particulier $f_p <_f f_q$. Il existe donc deux périodes r et s , appartenant respectivement à f_p et à f_q telles que $r < s$. Comme $p \subseteq r$ et $q \subseteq s$, on obtient par la propriété de monotonie que $p < q$.

La construction qui précède permet de montrer qu'il existe une structure de points équivalente à toute structure de périodes, sans imposer aucune exigence supplémentaire sur cette dernière. Or la contrepartie de cette flexibilité est que les structures de points ainsi construites présentent un défaut intrinsèque : elles contiennent trop d'éléments.

Pour exposer ce problème, prenons l'exemple d'une structure de périodes finie $\mathbf{P}(P, \subseteq, <)$, telle que $P = \{p_1, p_2, p_3, p_4\}$, et que $p_4 \subseteq p_2$, $p_2 \subseteq p_1$, et $p_3 \subseteq p_1$. Les ensembles $f_1 = \{p_1\}$, $f_2 = \{p_1, p_2\}$, $f_3 = \{p_1, p_3\}$, et $f_4 = \{p_1, p_2, p_4\}$ constituent les

filtres définis à partir de \mathcal{P} . La structure d'intervalles de filtres qui en résulte contient les ensembles $i_1 = \{f_1, f_2, f_3, f_4\}$, $i_2 = \{f_2, f_4\}$, $i_3 = \{f_3\}$, et $i_4 = \{f_4\}$. Il est facile de vérifier que la même structure d'intervalles de filtres pourrait résulter des seuls trois filtres f_2 , f_3 , et f_4 . Autrement dit, le filtre f_1 est redondant par rapport aux trois autres filtres.

Pour remédier à ce défaut, on peut modifier la construction précédente en introduisant la notion d'ultrafiltre.

Soit $\mathcal{P}(\mathcal{P}, \subseteq, <)$ une structure de périodes. Un sous-ensemble u de l'ensemble \mathcal{P} est un ultrafiltre, si et seulement si u est un filtre, et pour toute période p n'appartenant pas à u , il existe une période q appartenant à u , telle que $\neg p \circ q$.

L'intérêt principal de l'emploi des ultrafiltres réside dans le fait que chaque fois qu'ils contiennent une période, ils contiennent au moins l'une des périodes qui la constituent. Cette idée de constituance peut être formulée à l'aide de l'opération de somme. Supposons que pour une période quelconque p appartenant à un ultrafiltre u , il existe deux périodes q et r , telles que $p = q + r$. On peut vérifier que soit q appartient à u , soit r appartient à u , soit les deux lui appartiennent. Car, supposons qu'aucune des deux périodes q et r n'appartienne au filtre u . Selon la définition ci-dessus, il existe deux autres périodes s et t appartenant à u , telles que $\neg s \circ q$ et $\neg t \circ r$. Mais, u étant un filtre, la période $s \cap t \cap p$ existe et appartient à u . Or, par la définition de l'opération de somme, cette dernière période doit avoir une intersection avec au moins l'une des deux périodes q et r , ce qui est contradictoire avec notre dernière supposition.

Dans l'exemple ci-dessus, le filtre $f_1 = \{p_1\}$ n'est pas un ultrafiltre : il ne contient aucune des deux branches constituantes de la période p_1 . C'est pour cette raison qu'il est redondant par rapport aux deux ultrafiltres f_2 et f_3 qui, eux, contiennent chacun une branche possible de cette période, à savoir, respectivement, p_2 et p_3 . Ainsi le filtre f_1 représente un point surnuméraire par rapport à la structure de points optimale associée à la structure de périodes de départ.

Ainsi, à partir d'une structure de périodes, il est possible de définir une structure de points dont les éléments ne sont que des ultrafiltres de cette structure de périodes.

Soit $\mathcal{P}(\mathcal{P}, \subseteq, <)$ une structure de périodes. La structure d'ultrafiltres de périodes $\mathcal{UP}(\mathcal{U}, <_u)$ introduite par \mathcal{P} est définie par la donnée de \mathcal{U} , l'ensemble des ultrafiltres de l'ensemble \mathcal{P} , et d'une relation de précédence $<_u$ telle que $u <_u v$ pour deux ultrafiltres quelconques u et v , si et seulement s'il existe deux périodes p et q appartenant respectivement à u et à v , telles que $p < q$.

Avant de vérifier que \mathcal{UP} conduit à une structure de points, encore faut-il s'assurer que cette définition a un contenu, autrement dit que ces ultrafiltres de périodes existent. C'est bien le cas, car il est possible de démontrer que tout filtre est inclus dans un ultrafiltre. Ce théorème, qui est un équivalent de l'axiome du choix, se démontre en constatant qu'à partir d'un filtre, on peut construire une chaîne de filtres qui s'ordonnent de manière croissante pour l'inclusion. Cette chaîne admet un élément maximal qui est un ultrafiltre. À partir de là, on peut définir une structure d'ultrafiltres pour n'importe quelle structure de périodes.

La vérification du fait que \mathcal{UP} est une structure de points s'effectue exactement de la même manière que cette même vérification pour une structure de filtres de périodes. Pour montrer que la structure d'ultrafiltres de périodes représente bien la même structure temporelle que la structure de périodes de départ, il suffit de l'utiliser pour reconstituer une nouvelle structure de périodes, et montrer que cette dernière structure est isomorphe à la

structure de périodes initiale. Il faut toutefois que celle-ci possède la propriété de liberté renforcée.

Soit $\mathcal{UP}(U, <_u)$ une structure d'ultrafiltres de périodes introduite par une structure de périodes $\mathcal{P}(P, \subseteq, <)$ dotée de la propriété de liberté renforcée. La structure d'intervalles d'ultrafiltres $\mathcal{IUP}(I, \subseteq_i, <_i)$ introduite par \mathcal{UP} est définie par la donnée de I , l'ensemble de toutes les projections $i = \text{ultrafiltres}(p)$, où $\text{ultrafiltres}(p)$ est l'ensemble de tous les ultrafiltres auxquels appartient une période quelconque p appartenant à l'ensemble P , de la relation d'inclusion ensembliste \subseteq_i , est d'une relation de précédence $<_i$. Cette relation est telle que $i <_i j$ pour deux intervalles quelconques i et j appartenant à l'ensemble I , si et seulement si pour deux ultrafiltres quelconques u et v appartenant respectivement à i et à j , $u <_u v$.

La vérification du fait que \mathcal{IUP} est une structure d'intervalles de points isomorphe à la structure de périodes initiale, s'effectue par les mêmes étapes que cette même vérification pour une structure de filtres de périodes. Parmi ces étapes, seules celles pour lesquelles l'hypothèse de l'existence des filtres du type f_p dans toute structure de filtres de périodes a été utilisée, nécessite une démonstration différente. Car ces filtres ne sont pas obligatoirement des ultrafiltres : dans l'exemple ci-dessus le filtre f_1 est un filtre de ce type, et ce n'est pas un ultrafiltre.

Il faut donc démontrer, dans un premier temps, que, pour deux périodes quelconques p et q appartenant à P , si $\text{ultrafiltres}(p) \subseteq_i \text{ultrafiltres}(q)$, alors $p \subseteq q$.

Pour ce faire, supposons qu'il existe deux périodes p et q , telles que $\neg p \subseteq q$ et $\text{ultrafiltres}(p) \subseteq_i \text{ultrafiltres}(q)$. Selon la propriété de liberté de la relation \subseteq il existe une troisième période r , telle que $r \subseteq p$ et $\neg r \cap q$. Les périodes r et p appartiennent au filtre f_r qui est inclus dans un ultrafiltre u . Or, la période q est exclue de cet ultrafiltre. Donc, il existe au moins un ultrafiltre u qui appartient à l'intervalle $\text{ultrafiltres}(p)$, et qui n'appartient pas à l'intervalle $\text{ultrafiltres}(q)$, ce qui est contradictoire avec notre hypothèse initiale.

Dans un deuxième temps, il faut démontrer que, pour deux périodes quelconques p et q appartenant à P , si $\text{ultrafiltres}(p) <_i \text{ultrafiltres}(q)$, alors $p < q$.

Pour ce faire, supposons qu'il existe deux périodes p et q , telles que $\neg p < q$ et $\text{ultrafiltres}(p) <_i \text{ultrafiltres}(q)$. Selon la propriété de liberté renforcée de la relation \subseteq il existe deux périodes r et s , telle que $r \subseteq p$ et $s \subseteq q$, et aucune sous-période de la période r ne précède une sous-période de la période s . De même, il existe au moins deux ultrafiltres u et v , auxquels appartiennent respectivement r et s : ce sont les ultrafiltres qui incluent les filtres f_r et f_s . Or, si $u <_u v$, alors u et v contiennent respectivement deux périodes m et n , telles que $m < n$. Ainsi, les périodes $m \cap r$ et $n \cap s$ existent, et selon la propriété de monotonie $m \cap r < n \cap s$. Mais ceci est contradictoire avec la contrainte prononcée à propos des périodes r et s . Ainsi, il existe deux ultrafiltres, u et v , appartenant respectivement aux intervalles $\text{ultrafiltres}(p)$ et $\text{ultrafiltres}(q)$, tels que $\neg u <_u v$, ce qui est contradictoire avec notre hypothèse initiale.

Le développement qui précède montre que, sous certaines conditions minimales, les structures de points et les structures de périodes sont deux manières équivalentes de représenter le temps. Ceci suggère que la distinction intuitive entre les instants et les intervalles peut disparaître lorsqu'il s'agit d'une caractérisation mathématique de la structure du temps en termes de points ou de périodes. Cependant, cette distinction, loin d'être inhérente à l'organisation d'une dimension temporelle réifiée, est inspirée par le fait que les

énoncés du langage peuvent exprimer les situations, tantôt comme des changements ponctuels, tantôt comme des périodes de stabilité. Pour une modélisation mathématique plausible de l'expression langagière du temps, la tâche essentielle est donc de rendre compte de cette dualité.

La structure temporelle et les situations

L'adoption d'une structure temporelle, qu'il s'agisse d'une structure de points ou d'une structure de périodes, n'est qu'une première étape dans la représentation des situations exprimées par le langage. Il faut, de plus, intégrer cette structure dans le système logique qui modélise les énoncés langagiers.

Dans le cadre d'une logique d'ordre un, pour représenter la stabilité d'une situation, il semble plausible d'établir un lien entre la valeur de vérité d'un prédicat et un ensemble d'instantanés du temps. De même, pour représenter un changement de situation, on peut relier un instant du temps à la modification de valeur de vérité d'un prédicat. Or, ces deux opérations se révèlent inadéquates lorsqu'il s'agit d'exprimer la stabilité et le changement couramment véhiculés par le langage.

Considérons les énoncés suivants.

- (1) Ce matin, il n'était pas à l'université.
- (2) Cet après-midi, il était à l'université.
- (3) À midi, il est arrivé à l'université.

On peut représenter ces trois énoncés à l'aide d'un prédicat α qui représente la situation d'être à l'université¹¹. La constante t_0 représente l'instant présent.

$$\exists t_1 \exists t_2 \exists t_3 \ t_3 < t_2 < t_1 < t_0 \wedge \\ \forall t \ (t_3 < t < t_2 \supset \neg \alpha(t)) \wedge (t_2 < t < t_1 \supset \alpha(t))$$

Ce type de représentation oblige à traiter différemment le groupe verbal être à l'université, qui est traduit par le prédicat α , et le groupe verbal arriver à l'université, qui est rendu par une expression logique composée. De plus, il faut obligatoirement que l'arrivée corresponde à un instant du temps. Or, cela rend impossible la représentation d'un énoncé comme le téléphone a sonné pendant qu'il arrivait.

Le problème vient du fait que, dans cette démarche, la distinction intuitive entre la stabilité d'une situation et son changement, est représentée par l'emploi respectif d'un ensemble d'instantanés et d'un instant isolé. Si l'on renonce à ce choix, il faut supposer l'existence de deux instantanés passés distincts t_4 et t_5 , tels que $t_5 < t_4$, et qui prendront la place de t_2 , respectivement, dans les clauses associées aux expressions $\neg \alpha(t)$ et $\alpha(t)$. Ainsi, l'arrivée peut être représentée par une expression $\beta(t)$, qui aura la valeur vraie pour l'ensemble des instantanés t qui se trouvent entre t_5 et t_4 . Or, dans une telle représentation, il est impossible d'attribuer une valeur de vérité à l'expression $\alpha(t)$ pour le même ensemble d'instantanés.

Pour remédier à ce défaut, il est légitime de se tourner vers une représentation à base d'intervalles. Nous obtenons alors le type de représentation suivante pour les énoncés ci-dessus, où la constante t_0 désigne l'intervalle présent.

$$\exists t_2 \exists t_1 \ (t_2 < t_1 < t_0 \wedge \neg \alpha(t_2) \wedge \alpha(t_1))$$

¹¹ Nous considérons le cas où les éléments de la structure temporelle peuvent être des arguments des prédicats ; le cas où cette structure n'intervient que dans la procédure d'évaluation conduit à des conclusions similaires.

Une représentation correcte nécessite que la relation entre les intervalles t_2 et t_1 soit précisée. Trois cas peuvent être considérés. Dans le premier cas, on utilise la propriété de voisinage pour dire que l'intervalle t_1 est un voisin immédiat de l'intervalle t_2 . Dans ce cas, l'absence d'intervalle nul fait qu'un prédicat β censé représenter l'arrivée du personnage ne peut pas être situé temporellement. Dans le deuxième cas, il existe un nouvel intervalle t_3 telle que $t_2 < t_3 < t_1$, et l'expression $\beta(t_3)$ est vraie. Le problème est que l'on ne sait pas quelle valeur de vérité attribuer à l'expression $\alpha(t_3)$. Le troisième cas correspond au chevauchement des intervalles t_2 et t_1 . Appelons l'intervalle t_4 l'intersection des intervalles t_2 et t_1 . Ce cas est tout aussi problématique que le précédent, car l'expression $\alpha(t_4)$ est supposée recevoir deux valeurs de vérité contradictoires, dans la mesure où l'intervalle t_4 est inclus à la fois dans l'intervalle t_2 et dans l'intervalle t_1 .

On constate que l'expression des situations à l'aide de prédicats et d'une structure temporelle est loin d'aller de soi. On en vient donc à considérer l'idée que les situations elles-mêmes puissent être des objets logiques pouvant se prêter à des quantifications. De plus, il semble nécessaire de distinguer deux types différents de situations, les états et les événements, afin de rendre compte des intuitions associées aux notions de stabilité et de changement.

1.6. Les états et les événements

L'idée selon laquelle les situations doivent apparaître en tant qu'entités à part entière dans l'interprétation logique des énoncés du langage naturel a été proposée pour une raison indépendante des considérations qui précèdent, à savoir la représentation logique des énoncés portant sur des actions. Pour les verbes dits d'action, la représentation logique classique qui consiste à traduire un verbe sous la forme d'un prédicat ayant une arité égale au nombre d'arguments du verbe n'est pas satisfaisante (DAVIDSON 1980 [24]). Prenons les énoncés suivants.

- (1) Jean-Louis apprécie un Bordeaux dont j'oublie l'appellation exacte.
- (2) Jean-Louis boit un Bordeaux dont j'oublie l'appellation exacte.

Les formes grammaticales de ces deux énoncés sont identiques. Pourtant, parmi les deux énoncés suivants, seul le deuxième peut être logiquement déduit¹².

- (3) Jean-Louis apprécie.
- (4) Jean-Louis boit.

Certes, le verbe apprécier est généralement utilisé avec un complément. Il faut cependant expliquer pourquoi l'énoncé (4) constitue une conclusion correcte. Représentons les énoncés (1) à (4) en logique du premier ordre.

- (1') $\exists x \exists y \text{ bordeaux}(x) \wedge \text{apprécier}(\text{Jean-Louis}, x) \wedge \text{appelation_exacte}(y, x) \wedge \text{oublier}(\text{Laleh}, y)$
- (2') $\exists x \exists y \text{ bordeaux}(x) \wedge \text{boire}(\text{Jean-Louis}, x) \wedge \text{appelation_exacte}(y, x) \wedge \text{oublier}(\text{Laleh}, y)$
- (3') $\text{apprécier}(\text{Jean-Louis})$
- (4') $\text{boire}(\text{Jean-Louis})$

¹² Le présent du verbe boire est utilisé ici dans le sens du syntagme est en train de boire.

Le problème que nous avons posé plus haut revient à se demander comment l'expression (4') peut être dérivée de l'expression (2'). Ce problème est lié à un autre problème, celui du traitement des compléments circonstanciels et des adverbes. Dans les exemples suivants, l'énoncé (8) peut être déduit de chacun des énoncés (5), (6) et (7).

- (5) Jean-Louis boit un Bordeaux sur une terrasse du café.
- (6) Jean-Louis boit un Bordeaux dans une coupe.
- (7) Jean-Louis boit tranquillement un Bordeaux.
- (8) Jean-Louis boit un Bordeaux.

Dans ces exemples, le nombre de compléments du verbe boire ne change pas. La question est de savoir comment représenter les prédications correspondantes et le fait que l'énoncé (8) puisse en être conclu. On peut imaginer de traduire les énoncés (5) et (6) à l'aide d'un prédicat boire à trois places dont la troisième correspond, respectivement, à l'endroit indiqué ou au récipient utilisé. Cependant, rien n'indique la manière dont un énoncé qui contient le prédicat boire binaire est dérivé d'un énoncé comportant le prédicat boire ternaire. Dans le cas de l'énoncé (7), on peut imaginer utiliser un prédicat de deuxième ordre pour qualifier le prédicat boire. Encore une fois, on ne sait pas comment en conclure (8) qui correspond à l'action effectuée tranquillement.

La clé du problème réside justement dans le fait de savoir sur quoi porte l'adverbe. Une réponse plausible est de dire que l'adverbe porte sur un événement (DAVIDSON 1980 [24]). L'idée est qu'un verbe d'action comme boire introduit un événement qui devient argument du prédicat associé. L'adverbe introduit une prédication sur cet événement. La représentation des énoncés (5) et (6) et (7) et (8) peut alors prendre la forme suivante.

- (5') $\exists e \exists x \exists u \text{ boire}(\text{Jean-Louis}, x, e) \wedge \text{bordeaux}(x) \wedge \text{terrasse_du_cafe}(u) \wedge \text{endroit}(e, u)$
- (6') $\exists e \exists x \exists v \text{ boire}(\text{Jean-Louis}, x, e) \wedge \text{bordeaux}(x) \wedge \text{coupe}(v) \wedge \text{moyen}(e, v)$
- (7') $\exists e \exists x \exists w \text{ boire}(\text{Jean-Louis}, x, e) \wedge \text{bordeaux}(x) \wedge \text{tranquille}(w) \wedge \text{manière}(e, w)$
- (8') $\exists e \exists x \text{ boire}(\text{Jean-Louis}, x, e)$

Il est facile de voir que l'expression (8') peut être dérivée de chacune des trois expressions précédentes. Autrement dit, les énoncés (5) (6) (7) peuvent être paraphrasés comme suit.

Il existe un événement qui consiste dans le fait que Jean-Louis boit un bordeaux et que ...

Ainsi, de ces trois énoncés, on peut toujours conclure la paraphrase de l'énoncé (8).

Il existe un événement qui consiste dans le fait que Jean-Louis boit.

L'introduction des événements comme individus logiques à part entière simplifie donc significativement les choses. Cette méthode, que nous venons d'utiliser pour les adverbes et les compléments circonstanciels, peut être étendue au traitement des compléments obligatoires, ce qui donne les représentations logiques suivantes pour les énoncés (2) et (4). Ces traductions permettent de produire la dérivation de l'expression (4'') à partir de l'expression (2'').

- (2'') $\exists e \exists x \exists y \text{ bordeaux}(x) \wedge \text{boire}(\text{Jean-Louis}, e) \wedge \text{objet}(e, x) \wedge \text{appellation_exacte}(y, x) \wedge \text{oublier}(\text{Laleh}, y)$
- (4'') $\exists e \text{ boire}(\text{Jean-Louis}, e)$

Ce genre de démarche repose sur une caractérisation logique adéquate des mots qui peuvent introduire un événement. Notamment, la caractérisation des verbes dits d'action n'est pas sans poser de problème, y compris de nature philosophique. En particulier, les mêmes verbes d'action n'introduisent pas toujours les mêmes événements. Prenons l'énoncé Jean-Louis boit souvent un bordeaux dont j'ai oublié le nom. Peut-on dire que l'événement sur lequel porte l'adverbe souvent est l'acte de boire de Jean-Louis ? Il porte plutôt sur l'acte de boire un bordeaux particulier. Il est donc plus plausible de dire que l'événement introduit par cet énoncé peut être exprimé par l'énoncé Jean-Louis boit un bordeaux et non pas par l'énoncé Jean-Louis boit. Pourtant, dans les énoncés de type (7) comportant un adverbe de manière, l'événement qualifié est plutôt exprimé par l'énoncé Jean-Louis boit, car la conclusion Jean-Louis boit tranquillement semble correcte. Les indices syntaxiques sont donc insuffisants pour décider du type d'événement à introduire dans la représentation logique. Prenons un autre exemple. Il semble que n'importe quel verbe qui, normalement, ne désigne pas une action puisse introduire un événement. Dans l'énoncé Jean-Louis apprécie vraiment un bordeaux dont j'ai oublié le nom, l'adverbe vraiment ne porte pas sur un événement. Il qualifie plutôt l'intensité du prédicat binaire $\text{apprécier}(x, y)$. Dans l'énoncé Jean-Louis apprécie souvent un bordeaux dont j'ai oublié le nom, il existe une interprétation où l'adverbe souvent porte sur l'événement de Jean-Louis appréciant le bordeaux. Ce fait induit à penser qu'il existe un prédicat à trois places $\text{apprécier}(x, y, e)$. Le verbe apprécier n'est donc plus seulement une relation entre deux individus, mais désigne un événement possible dans laquelle le prédicat binaire $\text{apprécier}(x, y)$ peut avoir lieu. Les énoncés qui introduisent un événement ne sont donc pas limités à ceux qui comportent des verbes dits d'action.

Part of what we must learn when we learn the meaning of any predicate is how many places it has, and what sorts of entities the variables that hold these places range over. Some predicates have an event place, some do not. (DAVIDSON 1980 [24] p. 119)

Dans cette perspective, la caractérisation des énoncés qui introduisent des événements impose donc une réflexion philosophique approfondie sur les notions d'action et d'intention (DAVIDSON 1980 [24]).

Nous ne rentrons pas dans ce débat. Pour notre propos, il faut se demander en quoi les énoncés qui introduisent des événements expriment la temporalité différemment des autres. En particulier, nous devons les opposer à des énoncés qui introduisent des états, par exemple les énoncés construits autour d'une copule. Il est facile de représenter l'énoncé Jean-Louis est amateur de vin à l'aide d'une relation d'appartenance de l'individu Jean-Louis à la classe d'individus qui tombent sous la propriété d'être amateur de vin. Cette appartenance apparaît comme un état de fait. Cet état de fait peut changer quand l'individu cessera d'apprécier le vin. Jusqu'à ce jour, il existe une continuité qui se résume dans la validité de cette relation d'appartenance.

La même idée s'applique à bon nombre de verbes. Quand on prononce l'énoncé Jean-Louis préfère le bordeaux au Bourgogne, on énonce un état de fait. Le penchant de Jean-Louis peut être localisé dans le temps, comme dans l'énoncé depuis peu Jean-Louis préfère le Bordeaux au Bourgogne. Par contre, dans l'énoncé récemment Jean-Louis s'est mis à apprécier le Bordeaux, on marque un changement dans le comportement de l'individu. Il est toujours question d'un fait, mais on se focalise sur la survenue de ce fait et sur la différence entre cet état de fait et celui qui le précédait. Un tel changement d'un état à un autre apparaît comme un événement.

Events involve some kind of change, whereas states do not: that a state obtains over some interval *i* means that some condition remains in force for the duration of *i*. The occurrence of an event, in contrast, seems to imply that some condition, which obtains when the event begins, is terminated by the event and gets replaced by another, “opposite” condition.

(KAMP & REYLE 1993 [56] p. 507)

Nous obtenons ainsi une caractérisation de l’opposition état/événement qui s’exprime en termes de continuité/discontinuité. Un état correspond à une situation continue, dans laquelle une certaine relation est maintenue, alors qu’un événement est la marque d’une discontinuité dans cette relation. Cette distinction ressemble à l’opposition entre les points et les périodes, si bien que l’on pourrait être tenté de faire correspondre les états aux périodes et les événements aux points. Ce n’est malheureusement pas si simple. Les notions de points et de périodes peuvent être définies l’une à partir de l’autre. En revanche, nous avons deux intuitions différentes concernant la continuité et la discontinuité. La discontinuité n’est pas l’élément constitutif de la continuité, et la continuité seule ne peut engendrer la discontinuité. La différence de nature entre états et événements semble incontournable. Pour autant, les notions temporelles d’état et d’événement ne sont pas sans entretenir un rapport étroit, au point d’apparaître comme deux notions duales. En particulier, un état peut être borné par deux événements, et un événement peut être encadré par deux états.

Elle a habité dans cette ville depuis le décès de son grand-père jusqu’à la fin de ses études.
Ayant passé les concours avec succès, elle est partie de chez ses parents pour s’installer.

Dans cet exemple, le séjour dans la ville est un état borné par deux événements, sans doute deux déménagements. L’événement du décès sert de repère pour le premier, la fin des études permet de localiser le second. L’événement de la réussite dans les concours apparaît comme une frontière, terminant un état et commençant un autre, tous deux des domiciles. Chaque événement repéré joue le rôle d’une bifurcation. Sans cette bifurcation, l’état précédant l’événement se serait poursuivi. La réalité cognitive de la bifurcation apparaît encore plus clairement lorsque l’on fait des inférences sur la branche irréaliste.

Si elle n’avait pas réussi, elle ne serait pas partie.

Cette notion de bifurcation éclaire la dualité état/événement. L’état présente une stabilité intrinsèque. Seul un événement peut y mettre fin. Mais en tant que bifurcation, l’événement ne dure pas, il conduit à un changement. Le caractère insécable d’un événement le rend fondamentalement différent d’un intervalle constitué d’un ensemble d’instant. Un événement ne peut pas non plus être considéré comme un instant tandis que les états seraient des intervalles, car si un intervalle est constitué d’instant, un état n’est pas un ensemble de bifurcations possibles.

Sur quels indices faire reposer la distinction états/événements ? D’un point de vue linguistique, il semble qu’elle soit portée essentiellement par les verbes.

Cet après-midi il restera à la maison.
Ce soir il sortira de la maison.

La nature du verbe est donc importante, mais sa forme l’est tout autant, au point de pouvoir inverser la catégorie que l’on assigne normalement au verbe.

Il sortait de la maison quand l’un des invités l’a appelé.
Pour une fois, il est resté à la maison l’après-midi !

Cette possibilité de changer la nature événementielle des verbes par l’emploi d’une forme où d’une conjonction appropriée induit à penser que la distinction entre les états et les

événements n'est pas attachée aux mots, ni même à la situation désignée, mais au "regard" que l'on porte sur cette situation et à la manière dont ce regard est rendu par les mots.

States differ from events, we said, in that states involve the continuation of some condition whereas events involve its abrogation. But a condition is something conceptual, something that has a much to do with the way in which we choose to see reality as with reality itself. So we should not be surprised to find that the same bit of reality can be conceptualized either as event- or as state-like, depending on how we look at it. This conceptual dimension to the distinction between events and states is reflected by the way we speak.

(KAMP & REYLE 1993 [56] p. 507)

Une même scène peut être décrite par deux énoncés il lui tendit la main et il lui tendait la main. Ce qui distingue ces deux manières d'exprimer la situation n'est pas tant que la première induit une idée de ponctualité tandis que la seconde induit une idée de durée. Dans le premier cas, la scène est considérée de l'extérieur : on peut parler d'un avant et d'un après, si bien que la scène marque une discontinuité, un point de changement. Elle apparaît comme un événement. En revanche, dans le deuxième cas, on se place mentalement à l'intérieur de la scène. Cela induit un état, une continuité, car le point de vue intérieur laisse apparaître une période de stabilité dans la situation.

La distinction événement/état repose donc sur la notion de point de vue. Dans le cas d'un état, le point de vue est intérieur, ce qui permet ensuite d'explorer d'éventuelles bifurcations, sachant que les propriétés de l'état perdurent. Dans le cas d'un événement, en revanche, le point de vue est extérieur, si bien que les éventuelles bifurcations internes sont, à ce stade, inaccessibles.

Nous aurons l'occasion de revenir sur cette notion fondamentale de point de vue dans le chapitre suivant. Si la distinction état/événement ne correspond pas à la distinction point/période, il est légitime de se demander comment elle se traduit dans une structure temporelle. La section suivante suggère qu'elle peut être vue comme une opposition de nature topologique.

1.7. La topologie des situations

La représentation des relations logico-temporelles entre états et événements peut être obtenue en partant d'une représentation dynamique des situations en jeu, puis en effectuant une projection de ces aspects dynamiques sur une structure temporelle. La stabilité des valeurs de certains paramètres dans le domaine dynamique se projette, sur la structure temporelle, par un état, alors que la projection d'un changement des mêmes valeurs produit un événement sur la structure temporelle. On peut s'intéresser à une troisième catégorie de situation : un processus est la projection, sur la structure temporelle, du passage des valeurs de certains paramètres dans le domaine dynamique d'une stabilité initiale vers une stabilité finale (DESCLÈS 1989 [27]). Alors qu'un événement représente l'occurrence d'un changement, un processus représente sa trajectoire. Le contraste est illustré par l'emploi des formes progressives (he was swimming) et simples (he swam) en anglais. Le caractère duratif du processus peut être rapproché de celui de l'état en observant que le processus exprime, dans le domaine dynamique, la stabilité de la dérivée de certains paramètres. Il s'agit de se demander, à partir des définitions dynamiques qui correspondent aux états, événements, et processus, comment ces situations se différencient lorsqu'elles sont projetées sur une structure temporelle.

The situations are considered in linguistics as the denotations of predicative relations and are expressed by means of sentences and texts. The description of these complex denotations is represented by the *phase-portrait* of a dynamic system. This phase-portrait is an abstract space filled with the trajectories of the processes, occurrences of events and constant trajectories of states. To study the grammatical meanings, we project these trajectories and occurrences on a temporal referential. In this temporal referential, each process, event and state are represented by the intervals with different topological properties. (DESCLÈS 1989 [27] p. 169)

La projection de la représentation dynamique sur la structure temporelle permet de produire des relations logiques liées à la temporalité. Ainsi, une situation, caractérisée par un ensemble de valeurs de paramètres du domaine dynamique, peut se caractériser, après projection, par l'application d'un prédicat à un ensemble d'éléments de la structure temporelle. Si l'on fonde la classification état/événement/processus sur l'existence de propriétés dynamiques différentes, il faut montrer comment ces différences se traduisent sur la structure temporelle. Il semble plausible de considérer que ces différences sont de nature topologique. Pour cela, on considère une structure dans laquelle les situations sont représentées par des ensembles bornés d'instantants contigus (DESCLÈS 1989 [27]).

Soit T un ensemble non vide d'instantants, muni d'une relation d'ordre strict total $<$.

Un sous-ensemble I de T est convexe, si et seulement si pour deux instantants quelconques i et j appartenant à I , pour tout instant t appartenant à T , si $i < t$ et $t < j$, alors t appartient à I .

Soit I un sous-ensemble de T . Pour tout instant m appartenant à T , m est un minorant de I (respectivement un majorant de I), si et seulement si $m < i$ (respectivement $i < m$) pour tout instant i appartenant à I .

Soit I un sous-ensemble de T . Pour tout instant b appartenant à T , b est une borne inférieure de I (respectivement une borne supérieure de I), si et seulement si pour tout m minorant de I (respectivement majorant de I), $m < b$ (respectivement $b < m$).

Un sous-ensemble convexe I de T est, par définition, un intervalle ouvert, si et seulement s'il existe une borne inférieure b_1 et une borne supérieure b_2 pour I , tels que b_1 et b_2 n'appartiennent pas à I ¹³.

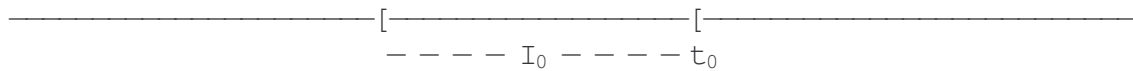
Un sous-ensemble convexe I de T est, par définition, un intervalle fermé à gauche (respectivement intervalle fermé à droite), si et seulement s'il existe un intervalle ouvert J , tel que I est égal à J augmenté de sa borne inférieure b_1 (respectivement de sa borne supérieure b_2).

Un sous-ensemble convexe I de T est, par définition, un intervalle fermé, si et seulement s'il est fermé à la fois à gauche et à droite.

Sur la base de cette structure, il est possible de distinguer les différents types de situations. Un état est représenté par un intervalle ouvert, un processus est représenté par un intervalle fermé à gauche, et un événement est représenté par intervalle fermé (DESCLÈS 1989 [27]). Le cas particulier d'un intervalle fermé réduit à un seul instant représente un événement ponctuel. Nous pouvons schématiser ces distinctions sur une ligne

¹³ Il est facile d'observer que les bornes d'un intervalle, si elles existent, sont uniques.

droite, censée représenter l'ensemble des instants ordonnés, sur laquelle nous superposons la notation classique par crochets des segments ouverts $]b_1, b_2[$ et segments fermés $[b_1, b_2]$. La première situation à représenter est le processus d'énonciation (DESCLÈS 1989 [27]).

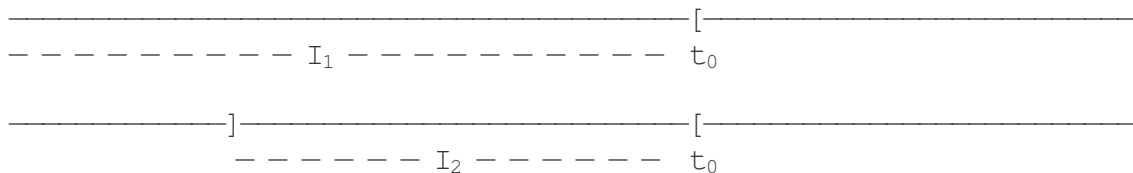


Le segment I_0 représente le processus d'énonciation, et le point t_0 représente l'instant de ce processus par rapport auquel les situations exprimées vont être représentées. Cet instant apparaît toujours dans les schémas et peut être appelé l'instant d'énonciation (DESCLÈS 1989 [27]). L'instant correspondant au début du processus I_0 n'apparaît dans les schémas que quand il est pertinent.

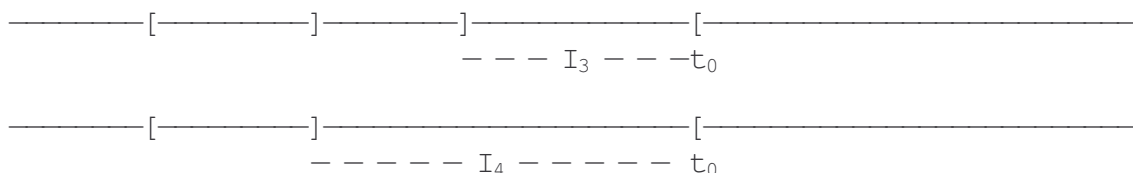
Prenons les exemples suivants.

- (1) Whales are mammals.
- (2) The door is open.
- (3) The door is opened.
- (4) He has opened the door.
- (5) He is opening the door.
- (6) He was opening the door when she arrived.
- (7) He opened the door.

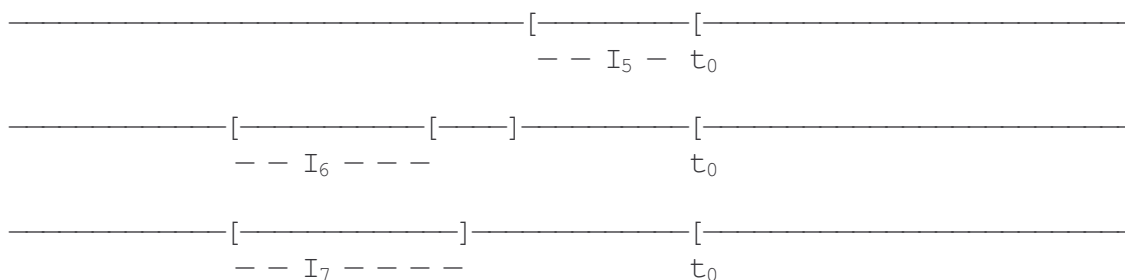
Les situations exprimées dans les énoncés (1) et (2) sont des états, représentés donc par des intervalles ouverts. Leur borne supérieure correspond toujours à l'instant d'énonciation. La différence est que le premier peut être vu comme un état permanent, dont la borne inférieure n'est pas précisée, tandis que le deuxième est un état contingent (DESCLÈS 1989 [27]).



Les énoncés (3) et (4) expriment aussi des états. Dans les deux cas, et contrairement aux énoncés (1) et (2), il s'agit d'un état causé par un événement, l'ouverture de la porte. Les représentations des énoncés (3) et (4) comportent donc un événement, plus un état résultant. L'énoncé (4) diffère de l'énoncé (3) par le fait que l'événement comporte un agent explicite. On peut qualifier les états exprimés dans les énoncés (3) et (4), respectivement, comme état passif et état résultant (DESCLÈS 1989 [27]). Dans le premier cas, on sait implicitement que l'événement existe et qu'il précède l'état, sans pouvoir dire qu'ils sont contigus. Dans le cas de l'état résultant, c'est l'état qui est implicite et sa borne inférieure est indiscernable de la borne supérieure de l'événement qui le produit.



L'énoncé (5) peut être représenté par un processus progressif (DESCLÈS 1989 [27]). Sa borne supérieure coïncide avec l'instant d'énonciation. L'énoncé (6) représente un processus du même type dont la borne supérieure est repérée par un événement antérieur à l'instant d'énonciation. Enfin, le processus exprimé en (5), lorsqu'il est terminé, donne un événement qui est exprimé par l'énoncé (7).



Le choix de ce système de représentation repose sur le fait que les intervalles ouverts peuvent donner lieu à la définition d'un espace topologique¹⁴. Ceci est dû au fait que, dans un ensemble totalement ordonné, la famille des sous-ensembles convexes qui ne contiennent pas leurs bornes constitue la base d'une topologie¹⁵. Ainsi, tout ouvert non vide de l'espace topologique résultant peut être écrit comme une union de tels intervalles. Notamment les intervalles ouverts eux-mêmes font partie des ouverts de cet espace topologique. Nous pouvons résumer cette construction comme suit.

Soit T un ensemble non vide d'instants, muni d'une relation d'ordre strict total $<$. Il existe un espace topologique (T, \mathcal{O}) , tel que tout ouvert de la topologie \mathcal{O} est l'union des sous-ensembles convexes de T qui ne contiennent ni leur borne inférieure, ni leur borne supérieure.

Dans l'espace topologique (T, \mathcal{O}) il est possible d'appliquer les opérateurs d'intérieur et de fermeture à tout sous-ensemble de T ¹⁶. Notamment, ces opérateurs peuvent être appliqués aux intervalles ouverts et fermés définis ci-dessus (DESCLÈS 1989 [27]).

Soit I un intervalle de T .

L'intérieur I° de I est un intervalle ouvert que I inclut, et qui inclut tout autre intervalle ouvert que I inclut.

La fermeture \bar{I} de I est un intervalle fermé qui inclut I , et qu'inclut tout autre intervalle fermé qui inclut I .

Le choix de représenter les états et les événements, respectivement, par des intervalles ouverts et fermés, entraîne que l'intérieur d'un état est identique à ce même état, et la fermeture d'un événement est identique à ce même événement. L'utilité de ces notions apparaît avec la possibilité d'avoir une double lecture de la même situation, que l'on décrit tantôt comme un état, tantôt comme un événement. L'événement correspond à la fermeture de l'état. En incluant l'état, son commencement ainsi que son achèvement, il offre une

¹⁴ Un espace topologique peut être défini par un ensemble d'ouverts. Étant donné un ensemble T , la structure (T, \mathcal{O}) est un espace topologique si \mathcal{O} est un sous-ensemble de l'ensemble $P(T)$, des parties de T , possédant les propriétés suivantes : les ensembles T et \emptyset appartiennent à \mathcal{O} ; toute intersection finie d'éléments de \mathcal{O} appartient à \mathcal{O} ; toute union d'éléments de \mathcal{O} appartient à \mathcal{O} . L'ensemble \mathcal{O} est appelé une topologie de l'ensemble sous-jacent T . Les éléments de \mathcal{O} sont appelés les ouverts de cette topologie.

¹⁵ Soit (T, \mathcal{O}) un espace topologique. Un sous-ensemble B de l'ensemble $P(T)$, des parties de T est une base pour la topologie \mathcal{O} , si et seulement si tout ouvert non vide appartenant à \mathcal{O} peut être construit par une union des éléments de B .

¹⁶ La définition des opérateurs d'intérieur et de fermeture se fait à partir de la notion de voisinage. Pour un élément quelconque t de T , un sous-ensemble V de T est appelé voisinage de t , s'il existe un ouvert \circ appartenant à \mathcal{O} tel que t appartient à \circ , et que \circ soit inclus dans V . L'intérieur I° d'un sous-ensemble I de T est l'ensemble de tous les éléments de T dont au moins un voisinage est inclus dans I . La fermeture \bar{I} d'un sous-ensemble I de T est l'ensemble de tous les éléments de T dont aucun voisinage n'a d'intersection vide avec I . Il est facile de voir que l'intérieur d'un ensemble est nécessairement inclus dans cet ensemble, tandis que sa fermeture peut contenir des éléments qui ne lui appartiennent pas.

description globale du changement. Inversement, comme l'intérieur d'un événement est ce même événement privé de son commencement et de son achèvement, il constitue un état décrivant le déroulement de cet événement.

Une telle description rencontre cependant quelques difficultés, en particulier, en ce qui concerne les événements ponctuels. Un intervalle fermé qui se résume à un seul instant ne peut pas être transformé en un intervalle ouvert. Or, le langage permet de transformer n'importe quel événement, aussi court qu'il soit, en l'exprimant sous la forme d'une durée, c'est-à-dire une succession d'instant. Un problème similaire apparaît quand on s'intéresse au début ou la fin d'un état. Intuitivement le début et la fin d'un état correspondent chacun à un événement. Or le seul intervalle qui peut être engendré, grâce aux opérations topologiques, à partir de l'instant qui correspond à la borne, est ce même instant.

Ces difficultés nous ramènent au problème, soulevé plus haut, concernant la représentation des situations par des instants isolés. Peut-on parler de l'instant précis où un état se termine pour laisser la place à sa négation ? Peut-on affirmer que l'événement qui est sous-jacent à ce changement peut se réduire à un seul instant ?

Une solution élégante consiste à remplacer les opérateurs habituels d'intérieur et de fermeture par d'autres opérateurs. De nouveaux opérateurs sont proposés dans le cadre de la locologie (DE GLAS 1992 [26]). Le but est de définir une forme de relation de proximité entre les instants qui représente la notion de grain. Cette relation remplacera la notion de voisinage qui est à la base des opérateurs topologiques habituels.

L'idée consiste à considérer les bornes d'un intervalle comme un nouvel intervalle. Ceci revient à modifier le grain d'observation : ce qui pouvait apparaître comme ponctuel à l'échelle de l'intervalle entier apparaît maintenant comme ayant un intérieur, ou plutôt un "cœur". À chaque grain d'observation, un halo est associé à chaque instant. De même que le grain est plus ou moins fin, ce halo peut être plus ou moins large. Un tel halo peut être calculé à l'aide d'une relation qui devra avoir les propriétés suivantes (DE GLAS 1992 [26]).

Soit λ une relation définie sur un ensemble T . On note $\lambda(t)$ l'ensemble de tous les éléments x appartenant à T tels que $t\lambda x$. La relation λ est, par définition, une locologie sur T , si et seulement si :

λ est réflexive ;

λ n'est pas maigre, *id est* il n'existe pas de t appartenant à T tel que $\lambda(t) = \{t\}$;

λ inclut une relation symétrique qui satisfait les deux conditions précédentes.

Sur la base de cette relation, deux opérateurs, appelés cœur et ombre, peuvent être définis (DE GLAS 1992 [26]).

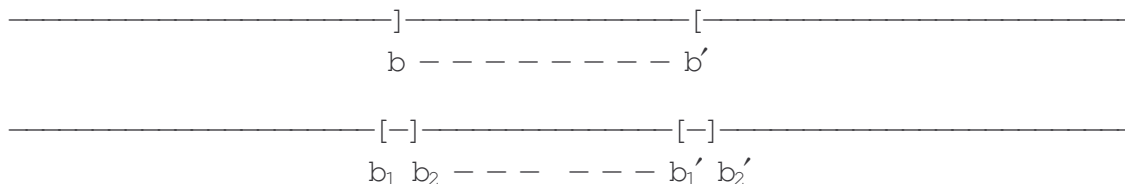
Soit T un ensemble totalement ordonné d'instant. Soit I un sous-ensemble de T .

Pour tout instant t appartenant à T , t appartient à l'ensemble $h(I)$, le cœur de I , si et seulement si $\lambda(t)$ est inclus dans I .

Pour tout instant t appartenant à T , t appartient à l'ensemble $s(I)$, l'ombre de I , si et seulement si l'intersection de $\lambda(t)$ et de I n'est pas vide.

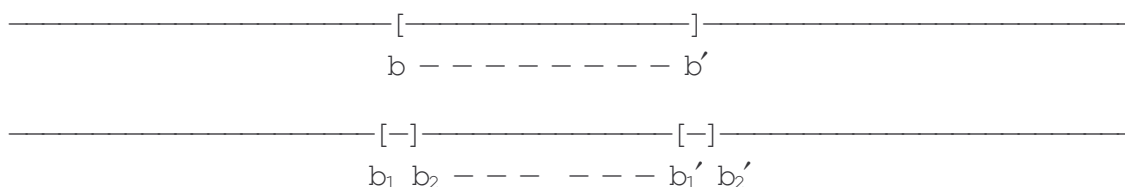
Ces deux opérateurs sont similaires aux opérateurs topologiques, avec la différence notable que les opérateurs topologiques sont idempotents : l'intérieur d'un ensemble est son propre intérieur, alors que le cœur d'un ensemble inclut son propre cœur. De même, l'ombre de l'ombre inclut l'ombre. Ceci est dû au rôle de la relation non maigre λ qui représente l'idée intuitive de changement de grain. Ainsi, le calcul du cœur ou de l'ombre d'un intervalle introduit implicitement un nouveau grain d'observation. Ce calcul peut être itéré pour obtenir des grains d'observation de plus en plus fin.

Si l'on applique les opérateurs locologiques à la structure des intervalles définie ci-dessus, le début et la fin d'un état peuvent être vus comme des événements.



L'intervalle $[b_1, b_2]$ est défini comme l'ombre de $\{b\}$. Il constitue un événement qui correspond au commencement de l'état de départ. De même, l'intervalle $[b_1', b_2']$, défini comme l'ombre de $\{b'\}$ constitue un événement qui correspond à l'achèvement de l'état de départ. L'intervalle $[b_1, b_2']$ est l'événement qui englobe ce même état. Noter qu'en raison du changement de grain, cet intervalle est différent de celui engendré par la fermeture de l'intervalle $]b, b' [$.

De la même manière, le début et la fin d'un événement peuvent être vus comme des événements.



Cette fois, $\{b\}$ correspond au cœur de l'intervalle $[b_1, b_2]$. Cet intervalle constitue un événement qui correspond au commencement de l'événement de départ. De même, $\{b'\}$ correspond au cœur de l'intervalle $[b_1', b_2']$ qui constitue un événement correspondant à l'achèvement de l'événement de départ. L'intervalle $[b_2, b_1']$ correspond à l'état qui décrit le déroulement de ce même événement. En raison du changement de grain, cet intervalle diffère de celui engendré par l'intérieur de l'intervalle $[b, b']$.

Cette modélisation présente l'avantage de proposer une description plausible, de nature topologique, des notions d'état et d'événement en évitant, apparemment, les difficultés inhérentes à une ontologie d'instant. Nous verrons cependant, dans le prochain chapitre, que les choses ne sont pas si simples, en raison du fait que les opérations de changement de grain ont lieu sur la même structure et peuvent être itérées à l'infini.

Outre son intérêt technique, l'usage de notions topologiques comme la proximité, l'intérieur, ou la frontière a l'avantage de mettre au premier plan la similitude intuitive entre les traitements cognitifs du temps et de l'espace.

1.8. Le temps, l'analogie et la métaphore

L'analogie entre l'expression du temps et celle de l'espace est manifeste. Cette analogie est fortement marquée dans le langage, par des interprétations parallèles des notions temporelle et spatiale d'étendue et de localisé. Les exemples suivants illustrent ce propos.

- Le positionnement d'un localisé dans une étendue.
- Cette grotte se situe dans la région de l'Ardèche.
- Le mariage a eu lieu pendant le mois de janvier.

La répétition d'un localisé dans une étendue.

Il y a des villages tout au long de la route.

Il y a eu des accidents au cours de son voyage.

La délimitation d'une étendue par deux localisés.

La forêt commence au pied des montagnes, et s'étend jusqu'au bord de la mer.

La fête a commencé au coucher du soleil, et a duré jusqu'à l'aube.

La séparation de deux étendues par un localisé.

La ville se situe entre la forêt et le désert.

Le cambriolage a eu lieu entre le matin et l'après-midi.

Le chevauchement de deux étendues.

L'herbe avance jusqu'au milieu de la cour.

Les vacances se sont poursuivies jusqu'à la moitié de l'automne.

L'ordonnancement de deux localisés.

Son bureau est à droite de la fenêtre.

Son départ a eu lieu après la chute du gouvernement.

Cette analogie étroite entre la représentation spatiale et la représentation temporelle est précieuse pour notre compréhension de cette dernière. Prenons l'exemple de l'opposition entre le localisé et l'étendu. La localisation spatiale d'un objet demande un regard extérieur. L'objet est considéré dans sa globalité, en tant que discontinuité dans le paysage qui l'entoure. En revanche, la perception d'une étendue suppose que le point de vue change et se situe à l'intérieur pour la parcourir. Ainsi, dans l'énoncé cette grotte se situe dans la région de l'Ardèche, le point de vue est externe à la grotte et interne à la région. Cette différence entre les points de vue extérieur et intérieur est exactement la même que celle qui s'applique à la distinction fondamentale entre les catégories temporelles d'événement et d'état.

Les exemples ci-dessus révèlent une autre parenté entre les représentations spatiale et temporelle. Des verbes comme commencer, s'étendre, avancer sont des expressions dynamiques, qui supposent un mouvement. Le mouvement est ainsi systématiquement employé, tant dans le domaine spatial que temporel, pour exprimer des étendues et des durées. L'omniprésence de la notion de mouvement s'observe également dans les emplois métaphoriques qui font intervenir des mouvements dits fictifs, comme dans les énoncés la moquette avance jusqu'au milieu de la pièce et la crise se rapproche à grands pas.

Les phénomènes liés à l'analogie et à la métaphore sont au centre des préoccupations d'un ensemble d'auteurs, regroupés dans le courant dit grammaires cognitives. Parmi les outils principaux pour rendre compte de ces phénomènes, il y a les deux notions de projection (*mapping*) et fusion (*blending*) (FAUCONNIER 1997 [33]).

L'idée est que les analogies et les métaphores mêlent des expressions langagières qui trouvent leurs significations dans des domaines cognitifs distincts. Les domaines de la temporalité et de la spatialité en sont des exemples. Or, les locuteurs utilisent parfois le vocabulaire associé à un domaine pour parler d'un autre domaine. Cette capacité cognitive est décrite comme un mécanisme cognitif de projection des structures du premier domaine sur le deuxième (FAUCONNIER 1997 [33]). Le premier domaine est appelé domaine source et le deuxième domaine cible. Prenons l'énoncé le calme revient. Habituellement, ce sont des personnes ou des objets mobiles qui sont susceptibles de revenir. Or, ici, le mot revenir est appliqué dans le domaine des propriétés.

De même, les locuteurs peuvent utiliser des expressions associées à deux domaines différents pour exprimer une idée indépendante de ces deux domaines. Ce genre d'expression

langagière s'explique par un deuxième mécanisme cognitif, le mécanisme de fusion, qui consiste à projeter les deux domaines sur un troisième, le domaine mixte (*blend*). Celui-ci hérite partiellement des structures des deux domaines et obtient ses propres structures émergentes (FAUCONNIER 1997 [33]). Considérons l'énoncé *ce geste lui coûtera cher*. Dans un premier domaine, les objets sont comparés entre eux et évalués par un prix. Un deuxième domaine concerne les actions et leurs conséquences. Dans le troisième domaine, qui résulte de la fusion, la conséquence du geste se trouve évaluée, comme le serait un objet, mais en tant qu'action, comme chère. Ces mécanismes de projection et de fusion peuvent être utiles pour analyser certaines constructions où le temps est utilisé métaphoriquement comme l'espace.

Prenons l'exemple suivant.

I can't catch up with myself.

Le locuteur exprime qu'il se trouve en retard par rapport à ce que la situation exige de lui. Par exemple, il devrait être déjà en train de rédiger une conclusion, alors qu'il écrit encore le corps du texte. Essayons d'expliquer, étape par étape, les opérations mentales qui nous permettent de comprendre cet énoncé.

Considérons trois domaines : le domaine des individus i ; le domaine des temps t ; et le domaine des situations s . On peut projeter les individus sur les temps : $F(i) = t$. Cela peut correspondre au temps des activités quotidiennes, comme dans l'énoncé *il faut que j'organise mon temps*. Le résultat de la projection peut aussi être l'époque dans laquelle l'individu vit, comme dans l'énoncé *pour son temps c'était un progressiste*. On peut aussi projeter les individus sur les situations : $G(i) = s$. Par exemple, la situation peut être celle dans laquelle on est impliqué quotidiennement, comme dans l'énoncé *il est impossible de travailler dans sa situation*, ou bien la condition générale dans laquelle l'individu se trouve, comme dans l'énoncé *dans leur situation ces jeunes ne peuvent que se révolter*. On peut réaliser une troisième projection en associant les temps aux situations $H(t) = s$. Les années soixante sont associées à l'action du mouvement pacifiste, le milieu de l'après-midi est associé au goûter.

L'exemple du retard sur l'emploi du temps peut être décrit en utilisant ces trois projections. L'individu i qui parle se voit associer un temps $F(i)$ ainsi qu'une situation $G(i)$. Or, son emploi du temps associe à son temps une situation $H(F(i))$ qui est différente de sa situation effective $G(i)$. Par projection directe, l'individu est associé à une situation qui diffère de celle que lui alloue une projection composée qui passe par son temps puis par la situation programmée pour ce temps. Ainsi, les deux pronoms *I* et *myself* évoquent, respectivement, les situations $G(i)$ et $H(F(i))$. L'emploi du temps de l'individu impose un ordre entre les différentes situations dans lesquelles il peut se trouver, ce qui rend possible leur projection sur des positions dans le domaine de l'espace. Ceci crée la possibilité d'employer des relations spatiales à propos des deux pronoms.

L'interprétation pourrait en rester là, comme celle qui concerne l'énoncé *she is behind of herself*. Pour expliquer l'emploi métaphorique du verbe *catch*, on peut considérer qu'il y a fusion des deux domaines d'individus et de l'espace. La structure qui résulte de cette fusion permet le mouvement d'une position vers une autre. Les deux positions associées aux deux situations, effective et programmée, de l'individu sont donc fusionnées avec deux individus distincts donnant lieu à un nouveau domaine, où peut alors se jouer une capture fictive entre les pronoms *I* et *myself*.

Conclusion

Dans la dernière section de ce chapitre, nous avons souligné l'aspect métaphorique lié à la notion même de localisation temporelle. Cette localisation est souvent comprise comme une activité de placement. Notre capacité à repérer les situations dans le temps est donc souvent représentée, dans les différents modèles, par le placement des situations sur une dimension temporelle.

La première question à résoudre, dans un tel schéma, est celle de la structure de cette dimension temporelle. Est-elle constituée d'instants qui peuvent se précéder ou coïncider ? Ou est-elle constituée d'intervalles qui peuvent s'inclure, se chevaucher ou se précéder en étant disjoints ? Nous avons pu vérifier que le choix présente peu d'enjeu, dans la mesure où chaque type de structure peut se définir à partir de l'autre.

Une deuxième question que nous avons abordée dans ce chapitre concerne l'opposition état/événement à propos de la nature des situations localisées elles-mêmes : Les états représentent une continuité qui résulte d'un point de vue intérieur ; les événements une discontinuité qui résulte d'un point de vue extérieur. Ce fait, qu'une même situation perçue puisse donner lieu à deux lectures radicalement différentes, nous paraît être à la base de la conceptualisation du temps. La modélisation de ces différentes "lectures" et des relations qu'elles engendrent fait l'objet du chapitre qui suit.

Chapitre 2 :
La leçon de Zénon

Introduction

Tout modèle visant à rendre compte de l'aptitude humaine à comprendre les relations temporelles exprimées par le langage doit expliquer comment nous parvenons, à partir de notre maîtrise d'une langue d'une part, et de nos connaissances encyclopédiques et épisodiques d'autre part, à localiser correctement les situations, dans le temps, les unes par rapport aux autres. Un tel modèle doit également prédire notre capacité d'effectuer des inférences à partir de cette localisation. L'intérêt de ce problème de modélisation est qu'il porte sur des phénomènes peu ambigus. L'ordonnement relatif de deux situations est, dans la plupart des contextes, une information binaire : soit A est avant B , soit c'est l'inverse. Toute erreur aura des conséquences probables sur les inférences que l'individu effectuera, notamment en bloquant toute causalité de A vers B , si A est compris, à tort, comme ultérieur à B . L'efficacité des narrations démontre que les individus sont experts dans la compréhension des composantes temporelles des énoncés. La tâche de modélisation est, avant tout, d'expliquer cette expertise.

Nous nous intéressons ici à une modélisation cognitive. L'objectif n'est pas d'imaginer un système théorique permettant de fonder des raisonnements scientifiques sur le temps. Il s'agit de cerner les moyens par lesquels les êtres humains parviennent à raisonner et à communiquer à propos de situations situées dans le temps. Notamment, il faut déterminer la nature de la structure de mémoire où sont stockées les dates et les durées des situations mémorisées ou imaginées. Ensuite, il faut comprendre comment cette structure de mémoire est lue et utilisée, autrement dit comment elle s'interface avec le langage et le raisonnement. Le problème est délicat, car la modélisation cognitive s'interdit de postuler des structures de mémoire infiniment grandes ou des mécanismes non constructifs, structures et mécanismes qui ne pourraient être hébergés par un cerveau humain¹. Cette restriction disqualifie, en tant que modèles cognitifs, la plupart des formalisations qui ont été proposées pour cerner la notion du temps sur le plan technique ou théorique. La réflexion développée dans ce chapitre va nous conduire à remettre en question certains présupposés classiques concernant la localisation temporelle des situations.

2.1. Le dilemme de la granularité

La difficulté première d'un modèle de la compétence humaine de communiquer et de raisonner à propos du temps réside dans le choix d'une structure de représentation permettant la localisation des situations dans le temps. En l'absence d'une telle structure, on voit mal comment les jugements sur l'ordonnement des situations pourraient être effectués. Nous savons que l'indépendance des États-Unis est antérieure à la révolution française parce que nous l'avons appris ; nous savons que la guerre du Golfe est postérieure à la révolution iranienne parce que nous nous en souvenons ; nous savons que la mort d'Aristote est postérieure à la naissance d'Alexandre parce que nous jugeons que le contraire serait contradictoire avec le fait que le premier fut le précepteur du second. Quelle est le mécanisme qui nous permet de comparer ces faits ? Quelle est la nature des représentations à l'œuvre lorsque nous effectuons la déduction concernant la mort d'Aristote ? Font-elles partie de la même structure que celles qui sont évoquées lorsque nous évaluons le temps qu'il nous faut

¹ Un exemple de ce type de postulat est remarquablement illustré par le comportement de Funes, le personnage fictif décrit par Jorge Luis Borges (*Artifices* 1944).

pour passer à la librairie avant d'aller prendre le train de 17h12 ? Un dispositif mécanique ou cognitif capable de faire ces jugements doit disposer d'une structure de mémoire comportant au minimum des indications d'ordonnement. La question à laquelle nous sommes confrontés est de savoir quelles sont les caractéristiques minimales dont une telle structure doit être dotée.

Il ne s'agit pas ici de seulement définir une structure à partir de ses propriétés. En mathématique, il est courant de définir des structures par un jeu fini d'axiomes. Ainsi, la structure des nombres réels peut être définie par une liste limitée de propriétés d'une relation d'ordre, comme celles de la transitivité et de la continuité. Cette capacité de notre esprit à définir des structures ne signifie pas que ces structures sont matériellement représentées dans le cerveau qui les a conçues. Ainsi, notre cerveau, quoique limité par le nombre fini de ses synapses ou de ses atomes, peut néanmoins concevoir l'infini continu de l'ensemble des nombres réels. Or, le problème qui nous occupe n'est pas un problème de conceptualisation, mais un problème de modélisation cognitive. La question n'est pas de savoir ce que notre cerveau parvient à conceptualiser, mais de comprendre les moyens qu'il met en œuvre pour le faire. Notre cerveau parvient à raisonner et à communiquer à propos du temps. Pour ce faire, il doit utiliser une mémoire temporelle, c'est-à-dire une structure cognitive capable de stocker des relations temporelles entre les situations mémorisées, perçues ou imaginées. En tant que telle, cette structure est supposée "matériellement" représentée. En d'autres termes, un certain nombre de neurones, de synapses et d'atomes sont requis pour chacun des éléments de cette mémoire. La question qui se pose est de déterminer la forme et la dimension de cette mémoire temporelle.

Un premier constat semble être que la mémoire temporelle doit être dotée d'un ordre total². Tout individu accepte que de deux situations quelconques A et B, l'une a dû précéder l'autre. Dans le cas où A et B sont susceptibles de se chevaucher, soit leurs débuts D_A et D_B , soit leurs fins F_A et F_B , sont dans un rapport d'antériorité, sauf à accepter une simultanéité parfaite. La situation exclue est celle d'une incomparabilité de principe. En d'autres termes, la structure de la mémoire temporelle est contrainte par notre capacité de comparaison. Si les situations sont, par principe, toutes comparables, donc si la mémoire temporelle est totalement ordonnée, ressemble-t-elle à une ligne, et si c'est le cas, combien cette ligne comporte-t-elle d'éléments ?

Une réponse immédiate à cette question est qu'il s'agit d'une structure finie. Comme la mémoire temporelle doit avoir une représentation matérielle dans le cerveau, il est exclu qu'elle comporte un ensemble infini d'éléments. La conséquence est que la précision avec laquelle nous pouvons conceptualiser le temps est elle-même finie. Il doit exister un grain élémentaire, un atome de durée en deçà duquel nous ne pouvons pas concevoir l'ordonnement des situations. Les situations qui ne diffèrent que d'une durée inférieure à ce grain doivent être conçues comme simultanés. Le dilemme de la granularité vient de ce que l'observation de notre pouvoir de localisation temporelle contredit l'existence d'un tel grain.

L'existence d'un grain temporel va de pair avec celle d'un horizon de précision. Or, il semble que notre capacité de comparaison ne soit pas astreinte à une telle limitation. Considérons la phase suivante.

Il y a quinze milliards d'années, trois secondes après le début de l'univers, la symétrie entre la matière et l'antimatière s'est brisée.

Une telle phrase n'est pas particulièrement délicate à comprendre. En particulier, l'ordonnement relatif du moment d'énonciation et des deux situations mentionnées, le

² Même si l'on prend en compte l'expression des situations conditionnelles, chaque branche de la mémoire temporelle doit être totalement ordonnée.

début de l'univers et la brisure de symétrie, ne fait pas mystère³. Le fait que nous comprenions cette phrase sans effort particulier est riche d'enseignements. Pour donner un sens à cette phrase, nous sommes capables de discerner des situations séparées de quelques secondes, chose que nous réalisons couramment dans la vie quotidienne. Si l'on imagine que les différentes situations doivent être positionnées sur une mémoire temporelle censée héberger, avec une certaine échelle et une certaine précision, l'ordonnement de tous les faits qui ont pu se dérouler depuis le début de l'univers, il faut que cette mémoire temporelle contienne au moins un demi-milliard de milliards d'éléments représentant chacun une seconde. Or, ce nombre est déjà trois ordres de grandeur au-dessus du nombre de synapses d'un cerveau ! L'idée d'une structure de mémoire à l'image du déroulement temporel est donc absurde, non seulement en raison de ses dimensions prohibitives, mais également pour la rigidité qu'elle impose : pour comprendre la phrase de notre exemple, il faudrait positionner le début de l'univers au moins à une seconde près. Or, si une précision d'une seconde est requise à la fin de la phrase, pour distinguer le début de l'univers et la brisure de symétrie, elle est inacceptable au début. L'auditeur non spécialiste de cette phrase n'a que faire qu'il s'agisse de quinze milliards d'années ou de seize, ce qui veut dire qu'il est prêt à tolérer une erreur de $3 \cdot 10^{16}$ secondes ! De plus, le positionnement ultra-précis du début de l'univers sur une échelle absolue est une tâche non seulement hors de portée de la science actuelle, mais de plus sans intérêt pour la compréhension de la phrase de notre exemple.

Une solution permettant d'éviter l'absurdité de l'hypothèse d'une mémoire temporelle de dimension gigantesque représentant l'entièreté du déroulement des situations, mémorisées ou imaginées, consiste à la considérer comme une mémoire logarithmique. La plupart de nos perceptions possèdent la propriété d'être sensibles aux variations relatives et non aux variations absolues. Ainsi, le doublement de la puissance d'un son est perçu comme un ajout de trois décibels, son décuplement par l'ajout de dix décibels. Si l'on applique ce principe logarithmique à la perception du temps, il semble que l'on évite le problème d'une structure de mémoire de taille exorbitante. Le rapport entre les deux durées de notre exemple passe de 10^{17} à une valeur raisonnable, par exemple 170. Malheureusement, comme nous allons le voir, une telle solution perd d'un côté ce qu'elle gagne de l'autre.

Considérons les deux situations de notre exemple, le commencement C de l'univers et la brisure B de symétrie. La phrase énonce que C et B sont séparés de trois secondes. Nous pouvons imaginer sans peine le discours d'un physicien qui se plairait à décrire une multitude de faits ayant eu lieu entre le commencement de l'univers et la brisure de symétrie. De manière plus systématique, nous pouvons reprendre le raisonnement effectué il y a vingt-quatre siècles par Zénon d'Elée. Nous sommes capables de concevoir un fait A_1 s'étant produit entre C et B et distinct de C et de B. Puisque C et A_1 sont distincts, nous pouvons de même concevoir qu'un fait A_2 se soit produit entre C et A_1 . Si l'on réitère le processus, on en arrive à considérer une suite de faits A_1, A_2, A_3, \dots . La question qui se pose alors est de savoir comment ces faits sont représentés dans la mémoire temporelle. En d'autres termes, il s'agit de savoir comment la structure de la mémoire temporelle est contrainte par notre capacité d'insertion.

Les structures de points et de périodes que nous avons considérées précédemment se classent en deux groupes : les structures denses et les structures discrètes (CF. CHAPITRE 1). Si l'on considère une structure discrète de situations, alors tout élément p , pour qui il existe un autre élément q tel que $p < q$, possède un plus proche voisin à droite.

$$\exists r \ p < r \wedge \neg \exists s \ p < s \wedge s < r$$

³ L'énoncé ne prétend pas à l'exactitude par rapport aux sciences physiques. Le lecteur spécialiste en cosmologie pourra rectifier les valeurs temporelles indiquées.

Le raisonnement de Zénon rend donc la propriété de discrétion, ou celle de voisinage, inacceptable, puisque celles-ci interdisent qu'une situation A soit insérée entre p et r .

Un problème analogue se pose si l'on s'intéresse aux durées qui séparent les situations considérées par Zénon. Considérons pour cela une structure de périodes constituée par les durées bornées par deux situations. Le raisonnement de Zénon empêche de conférer la propriété d'atomicité à une telle structure (CF. CHAPITRE 1). Si l'on considère une structure atomique, alors toute période p inclut un atome q qui n'inclut aucune autre période.

$$\forall r \quad r \subseteq q \supset r = q$$

Or, si q appartient à notre structure de périodes, c'est qu'elle est bornée par deux situations A et B . Le raisonnement de Zénon nous oblige à accepter l'existence d'une situation A_1 ultérieure à A et antérieure à B . Or, la période correspondant à la durée entre A et A_1 est strictement incluse dans q , ce qui est contradictoire avec l'hypothèse d'atomicité.

Le raisonnement de Zénon requiert donc la propriété de densité. Le fait que tout être humain soit capable de suivre le raisonnement de Zénon démontre que les modèles discrets, quelle que soit leur utilité d'un point de vue technique, sont dépourvus de plausibilité cognitive. Nous pouvons toujours, par la pensée, imaginer et verbaliser une situation située entre deux situations qui nous sont données.

Si nous considérons, de plus, que la mémoire temporelle peut être "plongée" dans la structure des nombres réels, qui représente par exemple le temps physique adopté par l'observateur scientifique, alors l'argument de Zénon oblige à considérer des durées de taille zéro. L'hypothèse, pour un i quelconque, de l'existence d'un A_{i+1} entre C et un A_i fait que la suite $(A_i - A_{i+1})$ tend vers zéro. Ceci est dû à la compacité du segment $[C, B]$ du temps physique : la suite A_i admet un point d'accumulation, et comme elle est décroissante, elle converge. Cet argument permet de conclure que les durées que le sujet doit se représenter sont, pour un observateur scientifique, arbitrairement petites en taille. Autrement dit, cet observateur doit non seulement considérer que la mémoire temporelle est infinie, mais qu'en outre elle ne comporte pas de grain. En d'autres termes, il n'existe pas d'horizon de précision dans les durées que nous pouvons concevoir.

Compte tenu du fait que tout être humain peut comprendre le raisonnement de Zénon, la question devient la suivante : peut-on représenter une structure qui respecte le raisonnement de Zénon sur une mémoire physique ? La réponse est "non !". À partir du moment où la relation de précédence est physiquement représentée, autrement dit si l'affirmation "A précède B" suppose que les situations A et B correspondent à des représentations mises en relation en mémoire, alors le caractère infini de la suite $\{A_i\}$, imposé par le raisonnement de Zénon, demande que la structure de mémoire comporte une infinité d'éléments distincts, ce qui est absurde. L'absurdité commence à partir du moment où l'on accepte que le raisonnement de Zénon engendre une infinité de représentations distinctes. L'exigence de continuité suggérée par Zénon oblige, de surcroît, à considérer que certaines durées matériellement représentées dans le cerveau du sujet sont, pour un observateur externe, de taille arbitrairement petite.

On pourrait imaginer échapper à l'exigence d'une structure de mémoire infinie en supposant que la procédure de localisation temporelle s'effectue sur une structure externe que l'esprit humain se contente de lire sans l'héberger. Le cerveau humain posséderait ainsi un mécanisme de repérage qui s'opère sur la structure du temps tel qu'il existe dans le monde extérieur. Indépendamment des problèmes considérables que poserait une telle solution sur le plan de l'épistémologie, constatons qu'elle ne résout pas le dilemme de la granularité. Pour satisfaire aux exigences du raisonnement de Zénon, la procédure de lecture doit atteindre des précisions arbitrairement grandes. Or, l'exécution d'une telle procédure demanderait un temps

non borné. Les humains peuvent sauter à des échelles arbitrairement petites en une seule étape, comme le suggère notre exemple sur le début de l'univers. Une procédure de repérage sur une structure externe est incapable d'une telle opération, car elle requiert un positionnement parfait à chaque étape, ce qui est irréalisable en un temps borné. Dans notre exemple, si le début de l'univers n'est pas localisé de manière parfaite, il n'existe aucune garantie que son positionnement précédera le moment de la brisure de symétrie.

Il ne s'agit pas ici de mettre en question la capacité humaine de concevoir une structure infinie : au contraire, nous allons tenter de montrer comment cette conceptualisation a lieu. En mathématique, une structure infinie peut être décrite par une axiomatique. Cette axiomatique peut être implémentée dans un dispositif artificiel par un programme. Cependant, le temps que mettra un programme de repérage, par exemple pour décider si un nombre qui lui est donné est plus petit que le nombre π , ne peut être borné, sauf si la précision avec laquelle les nombres sont représentés est également bornée, ce qui revient à opter pour l'atomisme. Ce qui est inaccessible à un dispositif matériel n'est pas d'héberger un programme de repérage dans une structure dense, mais de mettre en œuvre ce programme en un temps borné.

Le dilemme de la granularité nous laisse avec deux options tout aussi absurdes l'une que l'autre : l'existence d'un grain impose une atomicité inacceptable à la mémoire temporelle, tandis que l'absence de tout grain suppose que cette mémoire soit dense et infinie. Le fait d'introduire une imprécision dans la manière dont nous conceptualisons les durées ne change rien. Par exemple, le modèle locologique peut laisser supposer que le raisonnement de Zénon est bloqué, même en l'absence de la propriété de voisinage (CF. CHAPITRE 1). Ainsi, la situation A_{i+1} ne pourrait être conçue car les bornes de tout intervalle étant "épaisses", il semble en résulter une imprécision intrinsèque sur les durées qui empêche de séparer les situations C et A_i dès qu'ils sont suffisamment proches. En réalité, l'existence de bornes épaisses pour les durées ne change pas fondamentalement le raisonnement de Zénon. Dans la mesure où la borne possède un cœur, elle possède elle-même des bornes, qui ont elles-mêmes un cœur, et ainsi de suite. Aucune structure physique ne peut engendrer une telle itération. Une remarque analogue s'applique à des modèles qui représenteraient les durées par des segments flous. L'argument de Zénon réfute toute existence de flou absolu. Il est toujours possible de séparer mentalement deux situations, quelle que soit leur proximité temporelle.

La seule solution que nous pouvons envisager pour échapper au dilemme de la granularité consiste à autoriser une réutilisation des emplacements de mémoire. Ainsi, par exemple, les deux relations $A_2 < A_1$ et $A_{K+1} < A_K$ pourraient utiliser les mêmes emplacements m_1 et m_2 . Les situations que Zénon nous demande de considérer n'auraient donc qu'une existence temporaire. Une telle hypothèse a des conséquences qui vont nous amener à revoir radicalement la manière dont les relations temporelles sont cognitivement représentées. Pour que la confusion ne s'instaure pas entre les situations perçues et les représentations temporaires, il nous faudra imaginer deux mémoires temporelles distinctes. C'est le modèle que nous proposons dans la suite de ce chapitre.

2.2. Procédure de localisation temporelle

Le moyen que nous proposons pour échapper au dilemme de la granularité consiste à considérer deux types de supports de représentation du temps, dont aucun n'a le statut de mémoire globale. La conséquence de ce choix est que l'idée d'un temps linéaire et dense apparaîtra comme une extrapolation produite par notre cognition plutôt qu'une structure physiquement représentée dans notre cerveau. Si nous renonçons à l'idée d'une mémoire temporelle globale, nous devons répondre à deux questions : (1) Comment les situations de la mémoire épisodique et de la mémoire encyclopédique sont-elles datées ? (2) Comment

parvenons-nous à raisonner sur le temps et à juger de l'ordonnement relatif de deux situations données ?

Temps qualitatif

Les êtres humains ont une perception des durées. Nous ne confondons pas la durée d'un voyage de Téhéran à Paris avec le temps que met une pomme pour tomber de l'arbre, ou avec le laps de temps entre l'enfoncement d'une touche sur le clavier et l'apparition du caractère correspondant sur l'écran. La perception des durées peut s'exercer "en temps réel" : c'est cette capacité qui nous permet de rester en vie lorsque nous traversons une rue. La perception des durées peut s'exercer sur des situations mémorisées : nous nous remémorons la durée exagérée du discours de départ d'un professeur. Enfin, de même que les aspects qualitatifs de la perception se retrouvent dans la simulation mentale, la perception des durées intervient dans les scènes que nous imaginons : de même que nous pouvons nous imaginer en train de traverser Paris à pied, nous pouvons imaginer le temps que nous mettrions pour cela. Nous appelons temps qualitatif l'ensemble des représentations et des mécanismes qui nous permettent de percevoir, d'estimer, d'évoquer, et d'anticiper les durées de manière intuitive.

Notre capacité de représentation qualitative du temps ne semble pas limitée à la perception et la simulation des durées. Nous possédons en outre la capacité d'agencer temporellement certains faits de notre mémoire encyclopédique. Nous pouvons nous figurer la suite de certains événements qui ont jalonné l'histoire de notre pays, même si bien entendu nous n'en avons pas été témoins. Cette capacité repose sur le lien associatif que nous avons établi, par apprentissage, entre chaque événement historique et celui qui le suit, dans notre mémoire. Les limites de cette capacité d'organiser nos connaissances et nos souvenirs en chaînes associatives temporelles s'observent dans les erreurs couramment commises. Par exemple, nous pouvons donner une succession d'événements politiques qui se sont produits en Europe aux dix-septième et dix-huitième siècles. Nous pouvons faire de même pour la succession des œuvres philosophiques majeures de cette époque, des œuvres théâtrales, des œuvres musicales. En revanche, il est souvent difficile de synchroniser ces suites entre elles ; si bien que de nombreuses personnes, mélangeant les différents classicismes, situent Bach à l'époque de Molière, alors que le second a disparu soixante-dix-sept ans avant le premier.

Les chaînes associatives temporelles sont analogues à d'autres chaînes associatives que nous apprenons par cœur, comme l'alphabet, la suite des nombres ou la suite des mots bijou, caillou, chou, Le parcours de telles chaînes présente des aspects qualitatifs analogues à la perception du temps : nous savons sans calcul que les lettres d et s sont plus éloignées que les lettres v et y dans l'alphabet. Il en est de même pour les chaînes temporelles que nous avons apprises. Par extension, nous inclurons ces chaînes associatives temporelles dans le temps qualitatif, compris au sens large.

La définition précise des différentes capacités qui sous-tendent le temps qualitatif est du ressort de la psychologie (JOHNSON & FOLEY & SUENGAS & RAYE 1988 [53]). Nous retenons que notre représentation qualitative du temps est approximative et parcellaire. Nous n'avons qu'une idée grossière du temps qui s'est écoulé depuis le moment où nous avons regardé notre montre. Nos associations ne nous donnent qu'une image très lacunaire de la chaîne des situations qui ont jalonné notre dernier voyage. Comment expliquer que nous parvenions à raisonner efficacement sur le temps à partir d'une mémoire temporelle si peu fiable ? Nous pouvons comparer notre représentation qualitative du temps à la situation qui prévaut dans un atelier de montage de cinéma. L'opérateur dispose d'une multitude de bouts de films et de séquences de diapositives montrant le genre de plans que l'on peut voir dans des documentaires ou dans des fictions. L'opérateur est capable d'associer telle image de telle séquence avec une autre séquence. Il peut ainsi monter quelques séquences pour réaliser un

passage cohérent. Si notre perception du temps est fidèle à cette analogie où tout est approximatif, lacunaire et sujet aux erreurs d'association, il faut montrer comment nous parvenons à raisonner de manière précise sur le temps et expliquer d'où nous vient l'idée d'un temps linéaire.

Une autre manière de présenter la différence entre la représentation qualitative du temps et la localisation temporelle exprimée par le langage consiste à voir la première comme intuitive et implicite, alors que la seconde est, par nature, explicite. C'est cette propriété qui rend la seconde propre à être communiquée verbalement, alors que le premier demeure une expérience privée. Alors que la perception des durées est une expérience "à la première personne", nous pouvons exprimer par des mots une relation de précédence comme celle exprimée par la phrase la révolution a précédé la guerre. Dans notre modèle, ce passage à la verbalisation est rendu possible par la mise en œuvre de cartes temporelles.

Cartes temporelles

Notre modèle repose sur l'idée que les représentations qui sous-tendent le temps qualitatif ne sont pas les seules mises en œuvre dans le raisonnement temporel. Nous postulons l'existence d'une structure de représentation, qui peut être qualifiée de mémoire de travail⁴, constituée de quelques registres qui permettent de conserver actives des représentations de type qualitatif. Cette mémoire héberge les grilles temporelles que nous allons considérer maintenant.

Considérons la phrase avant le repas, elle est sortie pour acheter des cigarettes. L'auditeur de cet énoncé comprend qu'il a affaire à deux situations, le repas R et la sortie S en vue de l'achat des cigarettes. Il sait que S précède R d'une durée D de l'ordre de l'heure. Comment sait-il que D n'est pas de l'ordre de mille ans ou d'une microseconde ? Une réponse précise nécessiterait des investigations psychologiques. La question a été abordée dans le cas du raisonnement spatial (TVERSKY 1993 [103]). La distance désignée par un syntagme comme à droite de l'arbre n'est généralement ni un millimètre, ni un kilomètre. Elle dépend de la taille du repère, l'arbre, mais aussi de celle de l'objet localisé : en cherchant à droite de l'arbre en question, on regardera plus loin pour localiser une voiture que s'il s'agit de localiser un stylo égaré. Elle peut aussi dépendre de la distance entre les interlocuteurs et le repère. Des paramètres analogues doivent s'appliquer à la localisation temporelle. Nous reparlerons plus loin de l'influence, sur les durées à considérer, des situations localisées. Nous retenons ici le fait que les sujets sont capables de déterminer une durée D plausible en utilisant leur estimation qualitative des ordres de grandeur en jeu et de leurs connaissances concernant les situations concernées. Dans notre exemple, la durée typique d'un repas et la durée typique d'un achat de cigarettes induisent l'auditeur à penser que D est du même ordre de grandeur que ces deux laps de temps.

Nous modélisons la performance de l'auditeur qui saisit les relations temporelles de notre exemple en supposant qu'il applique une grille temporelle aux représentations qualitatives, évoquées par les mots de l'énoncé. Dans cet exemple, il s'agit d'une grille de séparation qui permet de séparer et d'ordonner les événements S et R dans le temps. Nous symbolisons cette grille comme deux ronds séparés par un segment, ce que nous notons $\circ\text{---}\circ$. La grille devient une carte lorsque ses deux ronds renvoient à des situations du temps qualitatif et sa ligne à une durée. La carte temporelle relative à notre exemple peut donc se dessiner par le schéma suivant.

⁴ Le terme "mémoire de travail", importé de l'informatique, est largement utilisé en neuropsychologie (BADDELEY 1992 [2]). Nous comprenons ce terme comme référant à une base de registres, analogues à des pointeurs informatiques. En aucun cas nous ne supposons la possibilité d'une mémoire temporaire dans laquelle des contenus pourraient être recopiés.



La durée D détermine l'échelle de la carte temporelle, de même que la distance réelle entre deux villes détermine l'échelle d'une carte sur laquelle figurent ces deux villes comme deux ronds reliés par une ligne symbolisant une route. La même grille s'appliquait à deux reprises dans notre exemple sur le début de l'univers, produisant deux cartes. La première carte, séparant le début de l'univers du moment d'énonciation, produisait une échelle de l'ordre de la dizaine de milliards d'années. La deuxième carte, qui permettait de séparer le début de l'univers de la brisure de symétrie, était à l'échelle de la seconde.

Notre modèle prévoit un nombre limité de grilles temporelles. Ainsi, une grille d'inclusion, notée $--o--$, permet de traiter les relations temporelles comme celle exprimée dans la phrase *elle est née pendant l'été 1350*. Dans ce cas, un événement, la naissance, est localisé au sein d'un état caractérisé par une date, l'été 1350. L'échelle est donnée, ici, par la durée considérée par le mot *été*. D'autres grilles sont possibles, comme celles que nous notons $o----$ et $----o$, qui permettent de traiter des phrases comme *elle est née au début / à la fin de l'été 1350*. Nous utiliserons également une grille d'extension, notée $=====$, pour signifier qu'un état s'étend au moins à un autre, comme dans la phrase *elle a été malade pendant deux jours*, où la durée de l'état de maladie évoqué est étendue de manière à englober la durée des deux jours. Enfin, nous utiliserons une grille de superposition, notée $--8--$, pour représenter l'aspect temporel des phrases comme *il est arrivé quand je suis partie*.

La notion de grille remplit la fonction de mémoire de travail. Les éléments d'une grille sont autant de registres qui peuvent "pointer", au sens informatique du terme, vers des éléments du temps qualitatif. Nous appelons moments les registres qui pointent vers des événements et qui sont désignés par des ronds dans notre symbolisation des grilles. Nous appelons époques les registres qui pointent vers des états et qui sont désignés par un segment dans une carte donnée.

Le fait de considérer des cartes temporelles peut laisser supposer que le modèle est sujet au dilemme de la granularité. Comment représenter des durées de temps de plus en plus fines sur une carte temporelle ? L'hypothèse fondamentale du modèle est qu'une carte n'offre que les moments et les époques que comporte sa grille. Tout autre traitement nécessite un changement de carte, comme nous le verrons.

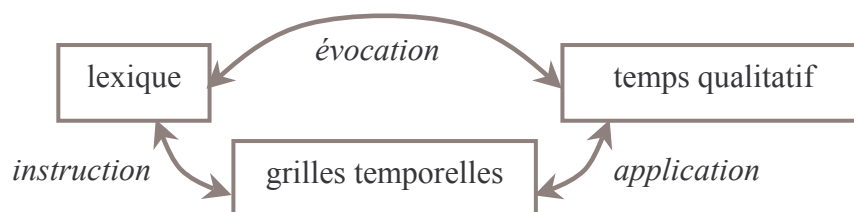
Du langage vers le temps qualitatif

Nous devons maintenant indiquer comment les deux systèmes que nous avons distingués, le système du temps qualitatif et le système des grilles temporelles, fonctionnent conjointement pour permettre la gestion des aspects temporels du langage et du raisonnement. Nous commençons par décrire comment les composants d'un énoncé simple contribuent à la formation d'une carte temporelle.

Certains éléments lexicaux d'un énoncé peuvent évoquer, par association, des épisodes mémorisés ou imaginés, ou des scènes prototypiques. Ainsi, dans la phrase *avant le repas, elle est sortie pour acheter des cigarettes*, le mot *repas* peut évoquer l'épisode du repas qui vient d'avoir lieu. D'autres éléments lexicaux peuvent infléchir ces associations. Les mots *hier*, *avant*, *ensuite*, *déjà*, *ancien*, *quand*, *pendant* *et cætera*, ainsi que certaines marques morphologiques comme le temps verbal, donnent des indications précises sur la manière d'interpréter d'autres mots de l'énoncé dans lequel ils interviennent. La situation indiquée par le syntagme *le repas d'hier soir* ne doit pas être confondue avec celle indiquée par le syntagme *le repas de ce midi*. La présence du mot *avant*, dans le syntagme *le repas avant la réunion*, suggère

que le repas dont il est question a précédé la réunion. Ces indications sont précisément ce qu'apportent les grilles temporelles et ce que ne permettent pas, à eux seuls, les mécanismes associatifs du temps qualitatif. Nous n'avons aucune difficulté pour nous rappeler ce que nous avons fait avant le repas et ce que nous avons fait après. Ce qui permet à notre interlocuteur d'évoquer en nous l'un ou l'autre de ces souvenirs est la différence qui existe entre les effets cognitifs produits par les mots avant et après. Selon notre modèle, ces mots nous donnent accès au souvenir adéquat par l'intermédiaire d'une carte temporelle. La carte fournit non seulement la direction dans laquelle aiguiller nos associations, mais également l'échelle à laquelle effectuer cette recherche associative.

Nous pouvons résumer l'organisation de la procédure de localisation temporelle par le schéma suivant.



Certains mots du lexique sont associés à des instructions pour le déclenchement de grilles temporelles. Ainsi, en français, le mot *pendant*, de même que la marque de l'imparfait, convoquent une époque. Le passé simple, en revanche, convoque un moment. À partir de ces instructions, l'auditeur est capable de former la grille temporelle appropriée. Dans la phrase elle est sortie avant le repas, le mot *avant* convoque deux moments, ce qui donne une grille de séparation $\circ\text{---}\text{---}\circ$. En revanche, dans la phrase elle est absente depuis le repas, le mot *depuis* convoque un moment et une époque, que l'on retrouve dans la grille de début $\circ\text{---}\text{---}$.

La formation de la carte temporelle commence au moment où les mots produisent leur pouvoir d'évocation par association. Dans la phrase elle est sortie avant le repas, le mot *repas* évoque, par exemple, le dernier repas que les interlocuteurs ont pris ensemble. Le mot *avant* pousse l'auditeur non seulement à considérer mentalement l'épisode du repas dans son ensemble, mais également à rechercher une évocation pour la sortie du personnage dans le temps qui a précédé le repas. Cette capacité suppose que nous soyons capables de diriger la recherche au sein de notre mémoire épisodique dans certaines directions. Lorsque nous nous demandons ce que nous avons fait hier, nous essayons de raviver nos souvenirs récents. Ce processus est sujet à l'erreur : nous pouvons prendre un épisode de l'avant-veille pour un épisode de la veille. L'indexation des épisodes est rendue plus aisée par certaines contraintes, comme la continuité des lieux où l'on se trouve, ou la causalité, qui fait qu'un effet suit sa cause. Par le jeu combiné des associations et de ces contraintes, nous parvenons à nous diriger dans notre mémoire épisodique. Ainsi, nous pouvons nous transporter mentalement dans l'année dernière, par exemple en ravivant un souvenir marquant ou typique de cette époque. Nous pouvons également nous projeter dans le mois prochain, en utilisant les connaissances que nous avons sur ce que sera notre situation (JOHNSON 1988 [52]). Il est important de noter que cette exploration de la mémoire épisodique est très peu déterministe, dans la mesure où elle repose largement sur les capacités associatives. Nous verrons comment la précision peut être obtenue par raffinements successifs.

Le schéma précédent décrit un partage des tâches. Certains éléments lexicaux d'un énoncé, comme les mots *repas* ou *sortie*, évoquent des épisodes que l'auditeur connaît ou qu'il peut imaginer. Pour aider à la détermination temporelle de ces épisodes, d'autres mots fournissent des instructions. Ainsi, le syntagme *pendant le repas* pousse le sujet à rechercher

mentalement une association au sein de la durée qualitative du repas⁵. Par exemple, la phrase elle est sortie pendant le repas peut conduire l'auditeur à former ou à rappeler un épisode dans lequel la personne, attablée et en train de dîner, se lève et sort. Les aspects qualitatifs de l'évocation ne sont pas toujours aussi riches. L'important est que l'interprétation de l'énoncé produit une carte temporelle en appliquant les éléments d'une grille temporelle à des épisodes plus ou moins instanciés. Comme nous le verrons, cette carte sert ensuite à la formation d'autres cartes ou à la production d'inférences.

Il importe de commenter la nature des instructions associées à certains mots à connotation temporelle. Par exemple, nous supposons que le mot après convoque une grille de séparation du type $\circ\text{----}\circ$; le mot pendant implique une idée d'inclusion et donc force à considérer une grille du type $\text{--}\circ\text{--}$ ou du type ==== ; de même, le mot quand suggère une simultanéité qui ne peut être exprimée que par les grilles $\text{--}8\text{--}$, $\text{--}\circ\text{--}$, ou ==== . Ce faisant, il semble que nous soyons en train d'assigner de manière rigide des schémas symboliques à des mots. Cela signifierait que ces mots sont, cognitivement, représentés sous la forme de ces schémas. Or, comme nous le verrons, une telle hypothèse pose des problèmes considérables (CF. CHAPITRE 7). Nous évitons cette difficulté en considérant que ces mots sont simplement associés aux instructions qu'ils donnent. Comme toute association, le lien que certains mots entretiennent avec les instructions qu'ils suggèrent est fortuit et modifiable par l'habitude. Simplement, l'éventail extrêmement restreint des instructions possibles fait que ces associations passent difficilement d'un type d'instruction à un autre. Elles sont donc d'une grande stabilité, au point de paraître aussi rigides que des définitions. Il n'en reste pas moins que les instructions temporelles attachées à un mot donné peuvent fluctuer et être ambiguës, comme pour tout lien associatif. Ainsi, l'évolution de la langue et la variété des parlers régionaux ont associé le mot français tantôt à des instructions temporelles équivoques, qui conduisent à considérer un moment soit antérieur, soit postérieur, voire un moment défini de la journée.

Du temps qualitatif vers le langage

Pour être cognitivement plausible, le même modèle doit fonctionner aussi bien en production qu'en compréhension des expressions langagières. Notons que les liens de notre schéma sont bidirectionnels. Les liens marqués évocation et instruction étant de type associatif, sont par nature bidirectionnels. Il est important de comprendre en quoi l'opération correspondant à la constitution d'une carte temporelle par application d'une grille temporelle au temps qualitatif peut se faire dans les deux sens. Dans un sens, celui de la compréhension, la grille est convoquée par certains mots de l'énoncé et contraint les associations que l'on peut faire au niveau du temps qualitatif. L'opération inverse, correspondant à la production d'un énoncé consiste, pour le sujet, à partir d'une situation, mémorisée ou imaginée au niveau qualitatif, qu'il cherche à localiser dans le temps, pour arriver à la mettre en relation avec une autre situation. Au niveau qualitatif, le temps n'est disponible que sous forme implicite. Il existe une perception des durées, tant pour la situation présente que pour les situations mémorisées ou imaginées, mais rien ne permet de faire des associations avec des mots comme avant, déjà, pendant, *et cætera*, qui supposent qu'au moins deux situations soient discernées sur le plan temporel. Nous supposons que, pour ce faire, les sujets possèdent les deux capacités suivantes.

Premièrement, les sujets peuvent considérer une situation en se plaçant mentalement à l'extérieur ou à l'intérieur, transformant respectivement la situation en moment ou en époque. Cette distinction correspond à la séparation traditionnelle entre états et événements, avec la

⁵ Un phénomène analogue existe pour l'espace. Une expression comme dans le stade pousse l'auditeur à explorer, visuellement ou mentalement, l'intérieur de l'espace circonscrit par les limites du stade.

différence notable que la propriété d'être un moment ou une époque est transitoire, et n'est attachée de manière intrinsèque ni aux situations, ni aux mots qui peuvent l'évoquer.

Deuxièmement, les sujets ont la capacité de contraster temporellement deux situations quelconques qu'ils considèrent. Selon qu'ils les considèrent de l'intérieur ou de l'extérieur, ils parviennent aux grilles temporelles indiquées dans le tableau ci-dessous. Ici, les termes thème et repère réfèrent respectivement au premier et au deuxième élément du contraste.

thème repère	moment	Époque
moment	○-----○ --8--	○----- --○-- -----○
époque	○----- --○-- -----○	=====

La première case de ce tableau décrit le cas où les deux situations sont considérées de l'extérieur et forment des moments. Les seules grilles appropriées à ce cas sont la grille de séparation et la grille de superposition. À l'inverse, la dernière case décrit le cas de deux situations considérées de l'intérieur. La seule grille qui puisse s'appliquer dans ce cas est une grille d'extension. Les deux cas croisés peuvent donner lieu à différentes configurations de type inclusion, début, fin. Noter que si les grilles sont les mêmes pour ces deux cas, l'ordre de l'évocation des deux situations mises en relation est inverse. Ainsi, la grille d'inclusion peut être produite lorsqu'un moment est situé au sein d'une époque, ce qui donnera une verbalisation comme la phrase elle est née en cette année-là, ou lorsqu'une époque est située par rapport à un moment, donnant par exemple la phrase l'année où elle est née.

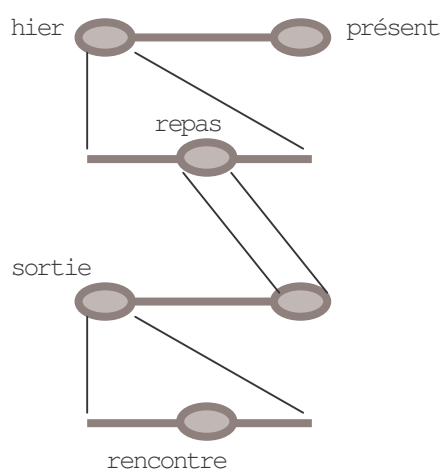
Insistons sur le fait que l'opposition moment/époque est une affaire de point de vue. La conséquence est que le seul sens que l'on peut donner à un constat comme "le verbe éternuer est un verbe d'événement" est que "typiquement, la situation évoquée par le verbe éternuer est considérée de l'extérieur". Il est parfaitement possible de dire et de comprendre une phrase comme pendant qu'il éternuait, il s'est produit plus d'un million de désintégrations dans l'échantillon de plutonium. Dans ce cas, la situation de l'éternuement est considérée de l'intérieur et apparaît comme une époque. Le fait que nous nous placions mentalement à des durées rarement inférieures à la seconde, dans la vie courante, est un fait fortuit. On ne saurait trouver là de raison cognitive fondamentale pour opérer une partition du lexique.

Une autre remarque concerne la pauvreté de l'inventaire des grilles temporelles. Le caractère topologique des relations qu'elles représentent exclut la possibilité de distinguer sur une seule carte deux épisodes légèrement décalés. Les seules distinctions possibles sont celles du tableau précédent. L'un des aspects prédictifs du modèle réside dans le fait que, pour conceptualiser des relations plus fines, nous sommes contraints d'utiliser plusieurs cartes, comme nous allons le voir.

Les opérations qui consistent à considérer et à contraster les épisodes permettent de parvenir à une carte, puis à une verbalisation. Le modèle fonctionne ainsi à la fois en compréhension et en production. Il semble cependant que, tel quel, il ne puisse expliquer les relations temporelles des énoncés complexes. Un aspect fondamental du modèle est que la procédure de localisation peut s'appliquer de manière récursive.

Procédure récursive de localisation

Une partie des productions conversationnelles, notamment dans les narrations, sont consacrées à la localisation temporelle des situations. Cette localisation ne se fait généralement pas en une fois. Le principe est que, par une succession d'inclusions et de séparations, l'auditeur a les moyens de reconstituer, avec la précision requise, la localisation d'une situation qu'on cherche à lui indiquer. Reprenons notre précédent exemple, que nous complétons pour obtenir la phrase hier, avant le repas, quand elle est sortie pour acheter des cigarettes, elle a rencontré son professeur. Notre modèle stipule que, pour traiter ce fragment de discours, l'auditeur forme successivement un certain nombre de cartes temporelles. La première lui permet de séparer hier d'aujourd'hui ; puis il place à l'aide d'une autre carte le repas à l'intérieur d'hier ; une autre carte encore lui permet de séparer la sortie du repas ; une dernière carte lui permet de situer la rencontre à l'intérieur de la sortie. Nous résumons cela dans le schéma suivant.

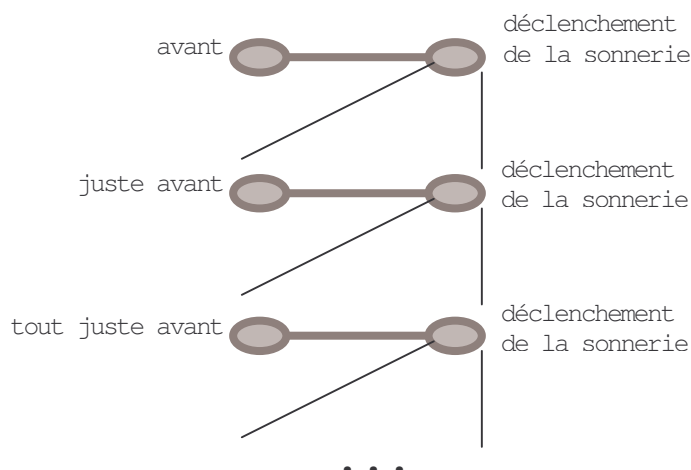


Cette succession de carte est due à une contrainte fondamentale imposée par notre modèle, qui est absente de la plupart des modèles rendant compte de la temporalité dans le langage. Cette contrainte veut que les différentes grilles ne soient pas superposables. En d'autres termes, il n'existe qu'une seule carte à un moment donné du traitement. Pour entrer à l'intérieur d'un moment, il faut le transformer en époque, ce qui exige un changement de carte. De même, pour opérer une séparation plus fine que la séparation courante, il faut également changer de carte. L'exemple précédent illustre bien le phénomène qui veut qu'une carte "chasse" l'autre. D'un point de vue cognitif, cela signifie que nous ne sommes pas capables de considérer (dans le sens particulier que nous avons donné à ce mot) plus de deux situations à la fois. Lorsque nous localisons la sortie avant le repas, la position d'hier par rapport à aujourd'hui n'est pas directement accessible ; de même, lorsque nous plaçons la rencontre avec le professeur pendant le temps de la sortie, le moment du repas n'est plus directement accessible. Cela ne veut pas dire que ces éléments ne laissent pas de trace mnésique. Simplement, ils ne sont accessibles que par voie associative, et non par "lecture directe". Ce phénomène a été étudié dans le cas des images mentales (KOSSLYN 1994 [61]) : lorsque le sujet déclare zoomer dans l'image, il perd la vision de l'ensemble (CF. CHAPITRE 9). Selon notre modèle, une situation analogue prévaut dans le cas des cartes temporelles. Cette contrainte conduit la procédure de localisation à se répéter en enchaînant carte après carte jusqu'à ce que la localisation soit jugée suffisamment précise pour les besoins de la communication.

À chaque changement de carte, les moments et les époques acquièrent un aspect qualitatif, lorsque cela est possible. Ainsi, l'auditeur peut se remémorer le repas de la veille. Il peut imaginer le trajet vers le bureau de tabac. Il peut rappeler à sa mémoire le visage du professeur s'il le connaît. Un aspect fondamental du modèle est que ces associations peuvent évoquer des images peu instanciées, voire aucune image. Cela ne bloque pas, la plupart du temps, le processus de compréhension. Ainsi, le lecteur a pu donner un sens à l'énoncé de notre exemple sans donner de durée précise au repas mentionné. On trouve un phénomène analogue dans l'imagerie spatiale, lorsqu'il s'agit d'interpréter la description d'un itinéraire. Le résultat aura un aspect qualitatif tout différent selon que l'on connaît les lieux ou non. Cependant, si la description est appropriée, celui qui ne connaît pas du tout le secteur peut néanmoins se former à l'avance une image, pauvre certes, mais suffisante, de ce qu'il doit trouver sur sa route.

L'itération de la procédure de localisation temporelle fait qu'on ne peut poser de limite *a priori* à la précision qu'il est possible d'atteindre dans l'évocation des épisodes. Cette propriété du modèle lui permet de résoudre le problème de la localisation temporelle sans tomber dans le dilemme de la granularité. Si nous reprenons la métaphore de l'atelier de montage cinéma, l'itération de la procédure de localisation conduit le monteur à passer de boîte en boîte, de bobine en bobine, de séquence en séquence, en suivant des liens associatifs, jusqu'à parvenir au passage recherché. Ce type de recherche ne requiert pas que l'ensemble des séquences soit totalement ordonné sur un film global unique. Il suffit que le monteur dispose de nombreux liens associatifs et soit capable de comprendre des consignes de direction ("c'était un film plus ancien" ou "la scène d'après") et d'inclusion ("c'est dans l'épisode du voyage"). Par cette technique, le monteur peut localiser, si elle existe, une scène correspondant à une durée arbitrairement courte. L'analogie suggère à la fois qu'il existe un grain pour nos souvenirs : nous ne sommes pas capables de rappeler à notre mémoire tous les détails de ce que nous avons vécu l'an dernier ou même la veille. Ce grain, dans notre analogie, correspond aux images effectivement présente sur la pellicule des reportages. En revanche, lorsque nous imaginons des scènes, ce qui correspond aux images de synthèse dans notre analogie, il n'y a pas de limite à la précision temporelle qui peut être atteinte. La procédure de localisation temporelle peut toujours être continuée un pas de plus, ce qui supprime toute idée de grain dans notre conceptualisation du temps.

L'évitement du dilemme de la granularité, dans notre modèle, est obtenu d'une part par l'abandon de l'idée d'une structure de mémoire temporelle unique, et d'autre part par le fait que la procédure de localisation temporelle est appliquée de manière itérative jusqu'à la précision désirée. Or, ce deuxième point semble problématique. Que se passe-t-il si, comme le suggère Zénon, on itère la procédure à l'infini ? Certes, comme la procédure n'opère pas sur une structure de mémoire globale, on ne peut pas en extrapoler l'existence d'un grain infinitésimal. Mais la possibilité même que la procédure puisse être itérée indéfiniment semble mettre le modèle en péril. En réalité, bien au contraire, le fait que la procédure de localisation puisse être itérée est précisément ce qui permet d'expliquer le fait que les sujets humains soient piégés par l'argument de Zénon. Prenons l'exemple de l'argument du moment inaccessible. L'expérimentateur demande au sujet de se placer mentalement avant le moment où l'horloge va sonner. Une fois cela fait, il demande au sujet d'imaginer un moment encore plus proche du déclenchement de la sonnerie, mais distinct de celui-ci. Cela fait, le sujet est censé recommencer, et ainsi de suite. Le sujet comprend rapidement qu'il s'agit d'un jeu sans fin. La perplexité cognitive qui en résulte est prédite par notre modèle. La figure suivante donne une idée de son comportement.



Lorsqu'il est en mesure, grâce à une grille de séparation, de distinguer un moment situé avant le déclenchement, le sujet ne peut considérer un moment situé "juste avant" le moment du déclenchement sans changer de carte. Le problème est que rien ne distingue la nouvelle carte de la précédente. Ceci est dû à la pauvreté de la représentation qualitative attachée à chaque étape. Comme nous n'avons pas d'expérience directe de durées aussi courtes, nous ne parvenons plus à donner un aspect qualitatif aux moments que nous construisons, si bien que les différentes cartes se retrouvent sans échelle, identiques les unes aux autres. Il en résulte une boucle potentiellement infinie qui ne dure que le temps de la persévérance de l'expérimentateur et de la bonne volonté du sujet. Le fait que le modèle prévoie l'entrée en boucle est donc à porter à son crédit.

Le fait que nous soyons capables de nous repérer dans le temps en l'absence de référence précise à toute expérience qualitative est une propriété remarquable de notre compétence temporelle. Pour le non physicien, une phrase comme il y a quinze milliards d'années, trois picosecondes après le début de l'univers, la symétrie entre la matière et l'antimatière s'est brisée reste compréhensible, bien que les durées mentionnées n'évoquent rien de palpable et restent totalement abstraites. Le sujet est néanmoins capable d'enchaîner deux grilles de séparation, ce qui peut lui suffire, dans de nombreux cas, à effectuer les inférences que l'on attend de lui.

Fonctionnement du modèle

Le modèle que nous venons d'esquisser présente un certain nombre d'avantages, dont le moindre n'est pas qu'il échappe au dilemme de la granularité. Nous pouvons également montrer qu'il fournit une description parcimonieuse d'un certain nombre de phénomènes d'acceptabilité sémantique relatifs à l'expression du temps.

Considérons les exemples suivants.

- (1) Il a mangé pendant le spectacle.
- (2) Il a mangé pendant deux heures.

Dans le premier cas, une interprétation possible est que l'individu s'est rassasié, par exemple en mangeant un sandwich, à un moment donné du déroulement du spectacle. Cette interprétation est impossible dans le deuxième cas, où le repas doit nécessairement durer les deux heures. Pour comprendre cette différence, examinons la nature des cartes construites. Dans les deux cas, le passé composé du verbe *manger* conduit à établir une première carte du type $\circ\text{---}\circ$ pour séparer le moment du repas du moment présent. Dans la phrase (1), la

mention pendant le spectacle conduit à considérer l'épisode du spectacle de l'intérieur, autrement dit à en faire une époque par rapport à laquelle le fait de manger est repéré. Nous avons donc une deuxième carte, qui peut être du type $--\circ--$, auquel cas le repas continue d'être un moment. Il ne s'agit que d'une préférence, car on peut comprendre que l'individu a mangé pendant toute la durée du spectacle, ce qui correspond à une carte du type $=====$. Dans la phrase (2), la deuxième carte correspond nécessairement à une grille d'extension. La raison que nous donnons de ce phénomène est que la mention deux heures est une indication de durée dont la détermination ne peut venir que du repas⁶. Il y a donc adéquation entre cette durée et celle du repas, d'où la grille d'extension. Il est important de noter que, conformément à ce que prévoit le modèle, l'auditeur imagine spontanément des situations simples. Dans le cas de la phrase (1), ce sera un repas ponctuel ou un repas qui occupe toute la durée du spectacle. Il n'imagine pas de configurations complexes comme deux prises d'aliment au début et une au milieu du spectacle. Contraint d'appliquer une grille d'inclusion ou une grille d'extension, il imagine une situation qui s'accorde le plus simplement possible avec ces grilles.

Considérons maintenant les phrases suivantes.

- (3) Le mois dernier, elle a repeint son appartement.
- (4) Le mois dernier, elle repeignait son appartement.

La phrase (3) suggère deux lectures, l'une où l'acte de repeindre n'a pris que quelques jours situés dans le mois précédent, l'autre où l'acte de repeindre a été réparti dans toute la durée du mois. Nous sommes exactement dans le cas de la phrase (1). Dans la phrase (4), seule la deuxième lecture est possible⁷. La marque de l'imparfait convoque une époque, autrement dit l'acte de repeindre est considéré de l'intérieur. L'apposition le mois dernier convoque également une époque. Nous sommes donc dans le cas d'une grille d'extension, ce qui explique que l'acte de repeindre soit perçu comme occupant l'entièreté du mois.

Considérons maintenant les exemples suivants.

- (5) L'an dernier, elle mangeait au café de la Sorbonne.
- (6) L'an dernier, elle mangeait au café de la Sorbonne quand la nouvelle est arrivée.
- (7) L'an dernier, elle mangeait au café de la Sorbonne pendant la coupe du monde.
- (8) * L'an dernier, elle mangeait au café de la Sorbonne pendant trois jours.
- (9) L'an dernier, elle mangeait au café de la Sorbonne pendant deux heures.

L'auditeur de la phrase (5) comprend que la personne a pris régulièrement ses repas au lieu dit pendant l'année précédente, autrement dit qu'il ne s'agit pas d'un seul repas, mais d'une répétition. Ce phénomène est une conséquence de la nécessité d'appliquer la grille d'extension. Comme dans le cas de la phrase (4), la présence de l'imparfait oblige à considérer une époque, ce qui ne laisse comme option que la grille d'extension. Le problème est qu'au niveau qualitatif, un repas typique ne dure pas une année. Dans un tel cas, notre modèle prévoit que pour satisfaire la grille d'extension, le sujet a la possibilité de considérer non un épisode unique, mais la répétition de ce même épisode. À l'échelle d'une année, un repas typique sera forcément considéré de l'extérieur. En revanche, sa répétition, en tant

⁶ Pour la phrase il a mangé entre 14h et 16h, la lecture dans laquelle le repas n'a duré qu'un moment est possible. Elle suppose toutefois que les bornes 14h et 16h aient leur propre détermination, par exemple s'il s'agit d'une enquête de police et que ces bornes sont déduites d'autres faits. Sinon, c'est la deuxième lecture, dans laquelle le repas dure au moins deux heures, qui s'impose.

⁷ La phrase (4) peut être aussi utilisée dans une forme particulière de narration, comme dans le fragment de discours il y a deux ans, elle achetait l'appartement, l'an dernier, elle le repeignait, cette année elle le revend. Ici, la simplification fait que les années sont présentées comme ne pouvant contenir qu'une seule situation. Tout se passe comme si les situations, citées en tant que faits marquants par rapport au contexte, suffisaient chaque fois à remplir la durée correspondante. Les situations et les durées sont donc, là aussi, coextensives, ce qui est cohérent avec notre remarque.

qu'habitude, peut être considérée de l'intérieur. Il peut donc y avoir co-extension entre l'année et le repas itéré, conformément à ce que demande la seule grille applicable dans la phrase (5). La phrase (6), quant à elle, ne requiert pas l'itération de l'épisode de manger. La marque de l'imparfait introduit bien une époque, mais cette époque n'est pas impliquée dans une grille d'extension. L'exemple met en jeu, tout d'abord, une grille de séparation pour situer l'an dernier par rapport au présent⁸, puis une grille d'inclusion pour inclure le moment du repas dans l'année, et enfin une grille d'inclusion pour inclure l'arrivée de la nouvelle dans l'épisode du repas. L'époque associée à la marque de l'imparfait est impliquée dans la dernière carte, non dans celle qui repère le repas par rapport à l'année, comme dans le cas de la phrase (5). Il n'y a donc pas lieu de considérer l'épisode comme répété. La phrase (7) s'analyse comme la phrase (5), sauf que la dernière carte redevient une carte d'extension à cause de la présence du mot *pendant*, si bien que l'on retrouve le phénomène de répétition que l'on avait pour la phrase (5). Le caractère peu acceptable de la phrase (8) entre bien dans le cadre de notre modèle. Comme pour la phrase (2), la durée mentionnée par le syntagme *trois jours* n'a pas d'autre détermination que l'acte de manger, ce qui force une grille d'extension. Le problème est qu'on se retrouve avec deux cartes contradictoires. Toutes deux utilisent une grille d'extension et concernent la version itérée du repas, comme dans le cas de la phrase (5). Or, le repas itéré ne peut pas être à la fois co-extensif avec une année et trois jours. L'acceptabilité de la phrase (9) est due au fait que la première carte d'extension concerne la version itérée du repas, alors que la seconde porte sur un seul exemplaire. Il n'y a plus de contradiction. Cet exemple illustre le fait que la possibilité d'itérer un moment pour en faire une époque, de manière à satisfaire une grille donnée, n'a pas de conséquence sur les autres grilles.

La revue de ces différents exemples avait pour but de montrer que le modèle fournit un cadre adéquat pour la description des phénomènes langagiers liés au temps. L'inventaire linguistique complet des instructions associées aux différents morphèmes à connotation temporelle, exprimées dans les termes de notre modèle, sera un pas utile vers l'élaboration du traitement automatique des aspects temporels de la langue.

Temps et raisonnement

Le modèle que nous venons d'esquisser permet non seulement de mieux comprendre les phénomènes langagiers liés au temps, mais également d'expliquer comment les individus raisonnent par rapport au temps. L'une des fonctions rendues possibles dans le cadre du modèle est la localisation des épisodes évoqués par le locuteur entre eux et par rapport aux souvenirs de l'auditeur. Comme signalé précédemment, on ne peut pas poser de borne *a priori* à la précision de cette localisation. Le positionnement par rapport à des références mémorisées dépend de la précision, souvent grossière, avec laquelle on peut accéder à ces souvenirs. Le positionnement relatif d'épisodes imaginés peut, en revanche, être arbitrairement précis. Ainsi, le locuteur peut nous amener à nous transporter mentalement quinze milliards d'années en arrière et nous faire constater qu'il y a eu un épisode de temps de quelques secondes cruciales pour l'avenir de notre univers. Nous avons la capacité de suivre ce repérage temporel, même si nous n'avons aucune représentation qualitative directe pouvant servir de support. La procédure de localisation temporelle permet ainsi à nos capacités de raisonnement de s'appliquer à des situations qui échappent à notre perception immédiate et mêmes aux limitations de nos souvenirs.

La procédure de localisation temporelle ne fait pas qu'augmenter considérablement le pouvoir de notre capacité à raisonner. Elle en constitue un élément fondamental. Prenons

⁸ Cette grille est facultative, car le syntagme *l'an dernier* peut être analysé par l'auditeur comme une expression figée évoquant une représentation qualitative par simple association.

l'exemple de l'alibi. Le suspect qui parvient à démontrer qu'il avait quitté les lieux avant le moment du crime convainc de son innocence. Les membres du jury doivent, pour cela, réaliser un raisonnement spatial et temporel. Le suspect, pour être coupable, devait se trouver sur les lieux "au moment" du crime. Dans cette version, le moment du crime est inclus dans l'époque de la présence sur les lieux, selon une grille du type --○--. Dans la version où le suspect est innocent, la grille qui s'applique est une grille de séparation ○----○ : le moment de la présence sur les lieux est séparé du moment du crime. Noter que la présence du présumé innocent sur les lieux du crime est considérée de l'extérieur, dans sa globalité, et apparaît cette fois comme un moment sur la carte temporelle. La capacité que nous avons d'exclure la culpabilité du suspect qui produit un tel alibi, est fondée sur le pouvoir des cartes temporelles de représenter les relations topologiques⁹. Les deux cartes étant incompatibles, le jury doit choisir entre elles. Le sort du suspect hérite du caractère binaire de ce choix.

Nous mettons en œuvre, à maintes reprises quotidiennement, cette capacité de considérer des grilles temporelles exclusives, par exemple pour réaliser que nous allons manquer notre train. Il est essentiel de noter que cette capacité ne peut être expliquée par un système purement associatif. Le caractère binaire du choix entre les deux grilles est dû à une exclusion de principe, non à la présence de deux attracteurs potentiels. Un système connexionniste, par exemple, peut basculer vers l'un ou l'autre de deux attracteurs qui ont été forgés par apprentissage. Avant de converger, il peut être considéré comme hésitant. Si le jury de notre anecdote hésite, ce n'est pas parce qu'il se trouve dans une position intermédiaire, considérant par exemple que le temps de présence du suspect recouvre à 30% le temps du crime. Pour hésiter, le jury "commute", de manière binaire, entre les deux grilles. Dans ce système, contrairement aux systèmes purement associatifs, les positions d'équilibre ne sont pas des attracteurs forgés par apprentissage, et les intermédiaires n'existent pas. L'exclusion propre au raisonnement logique est une exclusion de principe. Notre modèle suggère, au moins dans le cas du temps, que cette exclusion repose sur un mécanisme de nature topologique.

La procédure de localisation temporelle, avec ses grilles temporelles, intervient dans d'autres aspects du raisonnement. Elle intervient notamment dans les inférences liées à notre intuition de la causalité. Une même grille de séparation peut s'instancier dans deux cartes différentes, dans lesquelles deux situations E_1 et E_2 sont séparés dans des ordres inverses. Ainsi les deux moments, celui de gauche et celui de droite, d'une grille ○----○ peuvent tour à tour s'appliquer, respectivement, à E_1 et E_2 ou à E_2 et E_1 . Si c'est la deuxième carte qui est prise en compte, alors l'épisode E_1 ne peut être considéré comme la cause de l'épisode E_2 . Si nous savons que le malade a contracté la maladie après avoir été transfusé, nous sommes tentés d'établir un lien causal. Si nous apprenons qu'il a été déclaré malade avant la transfusion, celle-ci ne peut être la cause de la maladie. Là encore, le caractère incompatible des deux cartes est fondamental, et là encore un système purement associatif serait incapable de conclure à la stricte impossibilité du lien causal.

Considérons un autre exemple du rôle que joue la procédure de localisation temporelle dans le raisonnement, en comparant les deux phrases suivantes.

- (1) Quand elle était malade, il est venu la voir.
- (2) Quand elle était malade, il venait la voir.

Un francophone donnera différentes significations à ces deux phrases, notamment celles qui peuvent être caractérisées par les paraphrases suivantes.

⁹ Pour être complet, le raisonnement doit également être spatial, ce qui suppose là aussi un système d'exclusion topologique. Le lecteur aura compris que le modèle que nous développons pour le temps s'applique, sous une forme analogue, à l'espace, comme cela sera proposé dans la suite de notre travail.

- (a) un seul épisode de maladie, une visite unique.
- (b) un seul épisode de maladie, quelques visites.
- (c) un seul épisode de maladie, des visites régulières.
- (d) plusieurs épisodes de maladie, une visite à chaque occurrence.

La phrase (1) s'interprète préférentiellement comme (a) ou (b), bien que (d) soit possible. La phrase (2) s'interprète exclusivement selon (c) ou (d). La compréhension du mécanisme engendrant cette différence est essentielle si l'on veut prédire les différentes inférences que l'auditeur peut tirer à partir de ces phrases. Le traitement de la phrase (1) se termine par une carte d'inclusion --○-- : la maladie, représentée par une époque du fait de l'emploi de l'imparfait, inclut la visite qui apparaît comme un moment du fait du passé composé. Cette représentation ne permet pas de choisir entre (a) et (b). L'incertitude est prédite par la pauvreté de l'inventaire des grilles temporelles. Du point de vue du raisonnement, l'incertitude est sans importance s'il s'agit de réfuter la version négative de notre exemple, c'est-à-dire la phrase il n'est pas venu la voir quand elle était malade.

Dans le cas d'une occurrence unique de la maladie, la phrase (2) s'interprète selon (c). La raison en est que l'imparfait du verbe venir contraint à une grille d'extension =====. Comme la durée typique de l'acte de venir rendre visite est inférieure à celle d'une maladie, l'exigence de co-extension entraîne la répétition de l'épisode de visite, d'où l'interprétation (c). L'interprétation (d) est également provoquée par une question de durée relative. Si le contexte permet de penser que la durée prise en compte est plus grande que celle d'un épisode de maladie, il faut répéter cet épisode. Il en résulte un intéressant phénomène de synchronisation entre les deux séries. La synchronisation de chaque occurrence de la maladie avec une occurrence au moins de la visite est due à la présence du mot quand, alors que la co-extension des deux séries est due à l'emploi des deux imparfaits. Tout se passe comme si (d) était une version itérée de (a) ou de (b). Cette synchronisation des séries est importante du point de vue du raisonnement. L'auditeur peut, par exemple, en inférer la grande fidélité du personnage. Dans un exemple analogue comme la phrase quand les oiseaux passaient, l'alarme se déclenchait, la synchronisation apparaît comme une corrélation systématique entre deux événements *a priori* indépendants. Une telle corrélation, qui ne peut être due au hasard, amène l'auditeur à inférer un lien causal.

Le système des cartes temporelles permet également de réaliser des inférences concernant les durées des épisodes considérés. Selon que l'on prononce la phrase elle a vendu sa maison avant son voyage pour la Chine ou la phrase elle a vendu son livre avant sa visite au médecin, la durée de temps typique entre la vente et le déplacement n'est pas le même dans les deux cas. La raison de cette inférence est qu'au sein d'une carte de séparation ○----○, les ordres de grandeurs des durées, au niveau qualitatif, ne peuvent pas être trop disparates. Rappelons que les moments sont associés, au niveau qualitatif, à des épisodes considérés dans leur globalité. Si l'échelle qui permet d'embrasser les deux moments permet de les séparer, c'est que la durée qui les sépare est perceptible à cette échelle. De ce point de vue, compte tenu des durées typiques, la phrase elle a vendu son livre trois minutes après son voyage en Chine apparaîtra comme sémantiquement mal formée.

Dans certains cas, l'échelle est imposée par le contexte et contraint l'interprétation de la phrase ainsi que les inférences auxquelles elle donne lieu. C'est notamment le cas des énoncés exprimés au présent. L'interprétation de la phrase c'est Pierre qui prépare le repas peut, selon le contexte, peut avoir les formes suivantes.

- (e) Pierre est en train de préparer le repas.
- (f) Pierre va préparer le prochain repas.
- (g) Pierre prépare habituellement les repas.

La marque du présent portée par le verbe est donc ambiguë en français. Si l'on préfixe la phrase ci-dessus par les expressions en ce moment précis, aujourd'hui ou ce mois-ci, on obtient respectivement les interprétations (e), (f) et (g), cette fois-ci de manière non ambiguë. Ce phénomène s'explique, dans notre modèle, par un mécanisme en deux temps, qui implique successivement deux cartes temporelles. L'emploi du présent, en français, oblige à considérer une grille d'inclusion --o-- contenant un moment de référence, en général le moment d'énonciation¹⁰. Appelons *présent* l'époque qui figure dans cette grille. La grille donne lieu à une carte, dont l'échelle détermine si *présent* correspond à une seconde, un jour, un mois ou tout autre durée. Dans un deuxième temps, l'époque *présent* ainsi déterminée est utilisée dans une carte où intervient la situation décrite par le groupe verbal, généralement une grille d'extension ===== qui impose l'extension de *présent* à l'époque attachée à la situation décrite par le groupe verbal. Ainsi, l'ambiguïté des phrases au présent est due, dans notre modèle, au choix de l'échelle de la première carte. Si l'échelle est telle que *présent*, par exemple, est en deçà d'une heure, l'interprétation (e) pourra être préférée. En revanche, si l'échelle est d'un mois, ce sera (g) qui s'imposera, car la co-extension d'un repas et d'un mois impose la répétition de l'événement du repas. Le modèle explique également la possibilité de l'interprétation (f). Si *présent* est de l'ordre de la journée, les interprétations (e) et (g) sont bloquées, car *présent* ne peut être co-extensif, typiquement, ni avec la préparation du repas, ni avec sa répétition. Le présent doit être interprété comme un futur proche, ce qui convoque, non une grille d'extension, mais une grille d'inclusion --o--.

Les explications qui précèdent reposent sur le fait que l'échelle qui permet d'assigner une durée à *présent* est déterminée par le contexte. En d'autres termes, le modèle prédit que le contexte contrôle entièrement le choix entre (e), (f) et (g). Le même phénomène a lieu dans le cas spatial : le mot *ici* peut signifier, selon le contexte, la pièce, la ville ou le pays où l'on se trouve. Le contexte inclut, bien entendu, les durées typiques des éléments mentionnés dans la phrase. Ainsi, dans la phrase *c'est Sydney qui organise les Jeux Olympiques*, les interprétations (e), (f) et (g) concerneront des durées différentes de celles de notre exemple du repas¹¹.

Avantages du modèle

Les exemples qui précèdent montrent comment le modèle fonctionne et illustrent son caractère parcimonieux. Le principe de l'enchaînement récursif des cartes temporelles résout des difficultés qui nécessitent des solutions complexes dans d'autres modèles. Par exemple, certains modèles linguistiques repèrent des conflits dans les instructions données par différents éléments d'un énoncé. Ces instructions peuvent ne pas être cohérentes entre elles ou être contradictoires avec les données du contexte ou les connaissances que l'on a par ailleurs (GOSSELIN 1996 [44]). Cette perte de cohérence, source de conflit, peut porter sur le temps, sur l'aspect ou sur l'un et l'autre.

Les énoncés suivants constituent des exemples d'emplois de formes verbales et de compléments circonstanciels qui suscitent des conflits (GOSSELIN 1996 [44]).

¹⁰ Le temps du présent, en français, peut être employé pour le futur proche. Ainsi, on peut dire *jeudi prochain, c'est Pierre qui prépare le repas*. Le présent est également employé lorsque les interlocuteurs se situent mentalement à un moment donné du passé ou du futur, comme cela est courant dans les narrations.

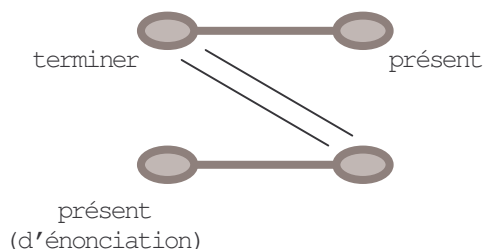
¹¹ Dans ce cas, l'interprétation répétitive risque d'être bloquée pour des raisons pragmatiques, car l'interlocuteur sait que l'organisation des Jeux Olympiques est tournante. Cependant, la phrase depuis vingt ans, *Sydney organise les Jeux Olympiques* sera interprétée avec succès de manière répétitive. C'est parce que cette interprétation réussit au niveau sémantique que l'incohérence peut être décelée au niveau pragmatique.

- (1) J'ai terminé dans une demi-heure.
- (2) Il fit du vélo pendant dix ans.
- (3) Il court le marathon en quatre heures.
- (4) Il découvrit le trésor en trois semaines.
- (5) Il sortit de la pièce où il écoutait de la musique depuis deux heures.

Dans l'énoncé (1), il y a conflit entre la forme du passé composé et le complément circonstanciel qui convoque un futur. En (2), le conflit naît de l'incompatibilité entre une connaissance encyclopédique, la durée typique d'une balade en vélo, et la durée donnée explicitement dans le complément. Dans l'énoncé (3), la morphologie verbale peut suggérer un aspect inaccompli, ce que l'on interprète classiquement par l'inaccessibilité de la borne supérieure de l'intervalle associé au processus, tandis que le complément circonstanciel exige que cette même borne supérieure soit accessible. En (4), l'aspect ponctuel suggéré par la morphologie verbale s'oppose à l'aspect duratif du complément circonstanciel. Dans l'énoncé (5), le sujet désigné par le pronom il semble occuper deux lieux différents à la fois, dans la pièce en train d'écouter et hors de la pièce puisque la sortie porte un aspect accompli.

Ces exemples posent de réels problèmes dans un modèle où les relations temporelles de l'énoncé doivent obligatoirement être représentées sur une structure unique, où les situations occupent des intervalles dont les bornes et les tailles relatives sont précisées. Pour résoudre ces conflits, une solution consiste à envisager que certaines zones de la structure de représentation puissent être sujettes à une déformation (GOSSELIN 1996 [44]). Ainsi, certaines zones temporelles sont censées se dupliquer, se contracter, se dilater, ou se déplacer selon les nécessités posées par le conflit. Bien que ces solutions se rapprochent de notre propre analyse, nous pensons que notre modèle est plus parcimonieux sur le plan cognitif, dans la mesure où il ne postule pas de traitement d'exception pour gérer les conflits. L'opération de la duplication introduite dans l'analyse citée ci-dessus est identique à notre mécanisme de répétition d'un moment au sein d'une époque. Les opérations de la contraction, de la dilatation, et du déplacement correspondent à la reprise d'un élément d'une carte dans une nouvelle carte. Selon le cas, une époque est reprise en tant que moment (contraction), un moment est zoomé pour se transformer en époque (dilatation), ou un moment est repris en tant que moment (déplacement). Notre modèle s'efforce donc d'expliquer les exemples qui précèdent en n'employant que des outils standard.

Le traitement de l'énoncé (1) se heurte effectivement à deux injonctions contradictoires, le passé du verbe et le futur du complément. L'énoncé est compris comme celui de j'aurai terminé dans une demi-heure, ce qui peut s'interpréter comme une priorité donnée au complément (GOSSELIN 1996 [44]). Nous préférons l'interpréter comme une propriété du moment présent qui, comme dans le cas du futur proche indiqué par le temps présent, n'est pas forcément astreint à inclure le moment d'énonciation. Si tel est le cas, l'énoncé (1) se traite par l'enchaînement de deux grilles d'extension, selon le schéma ci-dessous.



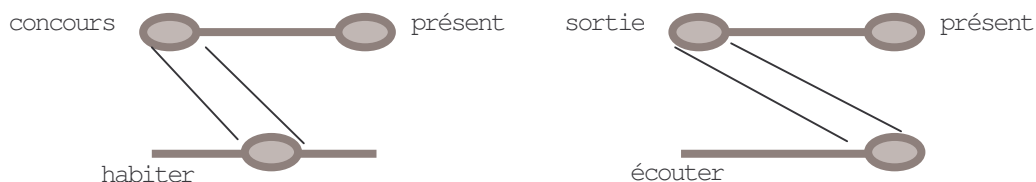
Tout se passe comme si l'époque présent pouvait apparaître de manière décorrélée dans plusieurs cartes. Cette possibilité explique l'emploi du présent narratif (il sort de la maison, il prend sa voiture, il arrive au carrefour, ...).

Pour un énoncé comme (2), nous avons montré comment le conflit entre les durées est résolu par la répétition de l'événement de base. L'événement itéré occupe une époque qui peut entrer dans la grille d'extension avec l'époque convoquée par le mot pendant.

Le traitement de l'énoncé (3), dans notre modèle, implique que l'époque présent soit coextensive avec la situation décrite par le groupe verbal. Par ailleurs, la durée introduite par la préposition en déclenche également une co-extension. Comme l'époque présent est supposée, par défaut, contenir le moment d'énonciation, il peut y avoir incompatibilité entre les deux co-extensions. Dans un contexte où le marathon est en train de se courir, l'incompatibilité peut être évitée (il court le marathon en quatre heures et après il revient pour manger). Si ce n'est pas le cas, l'incompatibilité entre les co-extensions conduit à l'itération de l'événement du marathon, comme pour l'énoncé (2), ce qui offre le degré de liberté manquant.

Dans un énoncé comme (4), notre modèle ne détecte pas de conflit, car les aspects ponctuels ou duratifs ne sont pas attachés à la même carte. Le passé simple convoque une carte de séparation qui distingue un événement, la découverte du trésor, du présent. Le complément circonstanciel introduit par la préposition en convoque une grille d'extension, ce qui nécessite une opération de zoom.

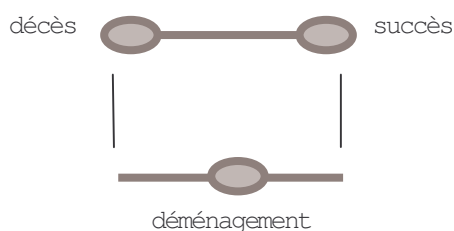
Le mécanisme de changement de carte permet encore à notre modèle d'éviter tout problème en ce qui concerne l'énoncé (5). Comme dans le cas des énoncés (2) et (4), l'emploi du passé simple déclenche une première carte séparant le présent du moment de la sortie. L'imparfait du verbe écoutait impose la convocation, dans une deuxième carte, d'une époque. La spécificité de l'énoncé (5) réside dans la manière dont ces deux cartes sont liées entre elles. Dans un énoncé comme elle passa le concours dans la ville où elle habitait, la deuxième carte comporte une grille d'inclusion, dans laquelle le moment reprend l'événement localisé par la première carte. Le schéma ci-dessous, à gauche, représente le résultat de cette procédure.



Dans l'énoncé (5), cette reprise par défaut est bloquée par la présence du mot depuis qui force une grille de fin, comme indiqué dans le schéma ci-dessus, à droite. En l'absence du complément introduit par le mot depuis, la représentation qualitative des éléments de l'énoncé, *id est* la simulation mentale de la situation, auraient pu également imposer le schéma de droite, seulement dans le cas où la sortie de la pièce signifie l'arrêt de l'écoute. Le conflit suggéré initialement n'existe que si l'on suppose, à tort, que l'imparfait convoque obligatoirement une grille d'inclusion. Or, ce n'est pas le cas. L'imparfait convoque simplement une époque. La carte qui en résulte dépend d'autres facteurs, comme le montre cet exemple.

Mentionnons un dernier exemple, qui semble de prime abord constituer une difficulté pour notre modèle, et qui, *a posteriori*, le conforte. Il s'agit de la préposition entre. Dans une phrase comme elle a déménagé entre le décès de son grand-père et sa réussite au concours, il semble qu'il y ait un double repérage, ce qui nécessiterait une grille temporelle à trois moments. Or, ajouter des grilles de plus en plus complexes risque de saper le modèle, dont la force repose précisément sur un inventaire limité et contraignant de grilles possibles. En réalité, l'idée d'une grille à trois moments provient d'une compréhension erronée du traitement de la préposition entre. L'usage de cette préposition, dans un schéma comme entre

A et B , convoque une grille de séparation $\circ\text{-----}\circ$. L'époque de cette grille est ensuite réutilisée dans une grille d'inclusion, selon le schéma suivant.



Cette description présente l'avantage d'être compatible avec d'autres emplois de la préposition *entre*. Dans la phrase elle a trouvé le temps de la recevoir entre tous ses rendez-vous, la collection des rendez-vous est supposée délimiter une époque qui sert de base à la grille d'inclusion finale¹².

Conclusion

Dans ce chapitre, nous avons défini un modèle permettant d'expliquer la localisation temporelle des situations en évitant de tomber dans le dilemme de la granularité. Nous avons donné, par des exemples, un échantillon des phénomènes langagiers liés au temps que notre modèle est en mesure d'expliquer. Nous avons ensuite essayé de montrer comment notre modèle permet d'expliquer, de manière simple, un certain nombre d'inférences concernant le temps qu'il est possible de faire à partir des énoncés du langage. Nous avons mentionné la localisation relative des épisodes évoqués ou imaginés, les déductions logiques liées à l'inclusion ou l'exclusion temporelle des situations, la causalité, la corrélation des séries répétées et l'estimation des durées. Cette richesse des effets langagiers et inférentiels prévus par le modèle est due au fait que le système des cartes temporelles et la procédure récursive qui le met en œuvre se situent à l'interface entre trois domaines cognitifs : le système du temps qualitatif, le système langagier et le système de raisonnement. Grâce aux cartes temporelles, nous pouvons associer de manière fiable un aspect qualitatif aux éléments d'un énoncé langagier, en tenant compte des indications portées par certains mots et marqueurs spécialisés de la langue et en produisant un certain nombre d'inférences. Ces capacités, dans la plupart des cas, ne peuvent pas être assurées par le seul jeu des associations qui peuvent exister entre ces différentes représentations cognitives. Notre modèle, avec son système d'établissement de cartes temporelles éphémères, est un moyen parcimonieux de rendre compte de ces capacités.

Une caractéristique essentielle de notre modèle est qu'il permet de prédire, comme nous l'avons vu, certains phénomènes systématiques de la temporalité, et ceci sans jamais assigner de manière rigide des durées ou des caractéristiques aspectuelles aux mots du langage. Les liens qui unissent les mots aux durées ne sont que des liens associatifs, présentant des propriétés de typicité. Ainsi, un film de cinéma dure typiquement 90 minutes, mais nous n'avons aucun problème pour comprendre un énoncé du genre le film a duré deux mille cinq cents heures. Si la durée typique d'un film était codée "en dur" dans la signification du mot film, l'énoncé ne pourrait tout simplement pas être compris. De même, comme nous l'avons signalé plus haut, des jugements aspectuels, comme la classification des verbes par leur

¹² Le fait que la préposition *entre* ait pour fonction de construire une interpolation s'observe de manière flagrante dans ses emplois non temporels, comme celui dans la phrase c'est un déjeuner entre amis. Nous reviendrons longuement sur la ressemblance entre les opérations temporelles et les opérations sémantiques en général (CF. CHAPITRE 9).

structure temporelle, ne peuvent être que des jugements de typicité, liées aux échelles de temps typiques qui accompagnent généralement les situations perçues dans le discours quotidien. Le verbe atteindre est souvent catégorisé comme un verbe d'achèvement. Pourtant, l'imparfait, dans l'énoncé alors qu'il atteignait le sommet, une avalanche se produisit, conduit à adopter un point de vue intérieur, si bien que le verbe atteindre se trouve être associé à une époque, perdant ainsi son "trait" d'achèvement.

Une autre caractéristique remarquable du modèle est la pauvreté des cartes temporelles qu'il postule. L'inventaire que nous proposons se limite à six grilles. Il n'existe rien qui ressemble à des grilles comme $\circ=\circ--\circ$ ou $--8--\circ--$. Ces relations plus élaborées sont produites, dans notre modèle, par la procédure récursive de localisation temporelle, qui utilise pour cela plusieurs cartes successives. Cette parcimonie n'est pas habituelle. La plupart des modèles sont contraints de s'arrêter à des schémas plus complexes lorsque, pour prédire les aspects systématiques de la temporalité, ils tentent de réduire la complexité des scènes évoquées par le langage. Par exemple, une distinction de principe entre état et processus semble permettre de distinguer les états stables et les états changeants, susceptibles de conduire à un événement servant d'aboutissement. Ainsi, les trois énoncés *il est assis*, *il court* et *il court à la banque* peuvent apparaître comme décrivant, les uns un état, les autres un processus. Le caractère indécis de l'énoncé *il court* doit nous alerter sur le fait que cette distinction entre état et processus dépend du point de vue adopté. La notion de processus suppose généralement un aspect cinématique, par opposition à la notion d'état qui est considérée comme possédant un caractère statique. Notre modèle n'utilise pas cette distinction statique/cinématique, qui n'est présente qu'au niveau des capacités perceptives. Elle n'entre pas dans les explications que nous donnons des phénomènes systématiques liés à l'expression langagière de la temporalité. En d'autres termes, nous considérons que les processus ne sont que des états. Dans les deux cas, le locuteur s'intéresse à une stabilité de certains paramètres de la situation décrite. Le fait que cette stabilité concerne une position ou une vitesse ne change pas l'analyse que nous faisons de la description langagière des situations concernées. Rappelons, de plus, que l'expression d'un état est due à l'adoption d'un point de vue intérieur. Dans l'énoncé *il a couru à la banque*, le point de vue global sur la situation la fait apparaître comme un événement. Il n'est plus question de suggérer que certains paramètres de la situation sont stables.

L'aspect le plus parcimonieux de notre modèle concerne le renoncement à toute structure temporelle globale, dense et totalement ordonnée. Grâce à la séparation entre grilles temporelles et temps qualitatif, nous pouvons renoncer à la densité et à la linéarité en tant que propriétés globales de la structure temporelle. La densité et la linéarité deviennent des propriétés "procédurales". Elles ne supposent donc pas l'existence d'une infinité d'éléments de mémoire physique, défaut qui disqualifie, d'un point de vue cognitif, la plupart des modèles mathématiques de la temporalité. Ainsi, les structures qui hébergent les bribes de notre mémoire épisodique et les registres qui nous permettent de raisonner sur le temps peuvent présenter une granularité même grossière sans que cela limite nos capacités à imaginer des durées toujours plus petites. C'est ainsi que notre modèle échappe au dilemme de la granularité.

Conclusion

La modélisation que nous venons de proposer demande à être complétée sur le plan linguistique. Il s'agit notamment de réaliser l'inventaire des morphèmes de la langue qui convoquent des instructions pour la mise en place des grilles temporelles. Il s'agit également de vérifier qu'un système à six grilles suffit à reproduire la totalité des effets temporels que nous pouvons exprimer par le langage. Il s'agit enfin de vérifier que l'ensemble des phénomènes liés à l'aspect et à la répétition sont correctement prédits par le modèle. Notre travail est donc susceptible de susciter des recherches en linguistique visant à caractériser de manière parcimonieuse la temporalité telle qu'elle est exprimée par les différentes langues.

Notre projet s'inscrivant dans le cadre de la modélisation cognitive, nous nous sommes efforcée d'étendre notre modèle en essayant d'évaluer de quelle manière notre description de la temporalité pouvait être extrapolée à d'autres aspects de la sémantique. D'emblée, nous avons été frappée, comme d'autres, par l'analogie entre le temps et l'espace. Ainsi, la procédure récursive de localisation s'applique, quasiment sans changement, au repérage spatial. On retrouve les phénomènes de carte, de grain, de récursivité. On retrouve les phénomènes de répétition, provoqués par la co-extension forcée de deux entités de taille différente, comme dans la phrase *la table est couverte de mouches*. Comme dans le cas du temps, certains mots comme *derrière* ou *à droite* de convoquent une grille spatiale.

Confortée par l'extension évidente de notre modèle au domaine spatial, nous nous sommes demandé si le système des cartes et de la procédure de repérage pouvait s'appliquer à l'ensemble des dimensions sémantiques. Or, nous nous sommes heurtée à la richesse des significations associées aux mots de la langue. Un verbe comme *manger*, lorsqu'on le réduit à la dimension temporelle, peut être ramené à un acte qui occupe une certaine durée dans le temps qualitatif. Sur quelles autres dimensions l'acte de prise d'alimentation doit-il être projeté pour que l'on puisse appliquer le système des grilles ? L'acte se décline sur un axe de plaisir/déplaisir, sur un axe d'appétit/rassasiement, sur un axe carnivore/végétarien, sur un axe cannibale/interdit, sur un axe sucré/salé, sur un axe cuisine-au-beurre/cuisine-à-l'huile, sur un axe peu/beaucoup, sur un axe partie/tout, et ainsi de suite à l'infini. Notre modèle de la temporalité, qui repose entre autres sur la possibilité d'imposer une relation d'ordre entre deux situations quelconques, semble incapable de faire face à cette profusion de dimensions.

Pourtant, nous restons convaincue que quelque chose, dans notre système de grilles, avait une portée générale. L'une des raisons de cette confiance est la systématisme de l'opération de négation. Sur la dimension temporelle, la négation est obtenue par une exclusion de nature topologique. Le sentiment qu'il en est de même quel que soit le domaine sémantique concerné nous pousse à mieux comprendre en quoi consistent ces dimensions sémantiques sur lesquelles les concepts exprimés par la langue trouvent toutes leurs nuances. Dans la partie suivante de ce texte, nous allons donc nous intéresser à une analyse critique des travaux visant à caractériser les concepts.

Partie 2 :
**Questions concernant la nature
des concepts**

Introduction

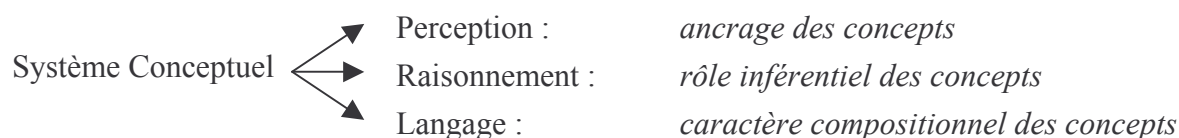
Dans la littérature des sciences cognitives, la notion de “concept” renvoie, de manière non exclusive, à l’une ou l’autre des trois acceptions suivantes.

Représentation : Il s’agit d’une entité mentale qui réfère à, ou désigne, une réalité dans le monde ou une donnée de la perception. Cette acception de la notion de concept, suggérée aussi par la notion d’idée, est sous-jacente aux théories classiques de l’entendement, de la connaissance, et de la signification. Dans ce sens, les concepts garantissent l’interaction du fonctionnement cognitif avec l’environnement, et ils se doivent disposer d’un ancrage dans cette interaction.

Prédicat : Une capacité cognitive largement étudiée par la psychologie est la faculté de raisonnement. Face à un objet perçu particulier, l’être humain ne retient que certaines de ses caractéristiques pour le considérer comme étant un élément d’une classe d’objets. Il peut alors utiliser le prédicat attribué à cette classe pour intégrer sa perception dans un processus de raisonnement. Dans ce cadre, la possession d’un concept garantit la convocation des connaissances associées et permet d’en tirer certaines inférences. Les concepts se voient donc dotés d’un rôle inférentiel.

Mot : La construction du sens exprimé par les langues humaines est un problème central de toute théorie linguistique. Le rôle que jouent les mots dans le processus se révèle incontournable. Dans une analyse cognitive des mécanismes sémantiques, les concepts constituent la projection mentale des mots. Leur manipulation dans des processus de composition serait dictée par la grammaire de la langue concernée. Les concepts acquièrent ainsi un caractère compositionnel.

Ces différentes acceptions sont rarement distinguées dans les textes et sont souvent suggérées simultanément lors de l’emploi du mot “concept”. La confusion qui en résulte parfois peut être évitée si l’on se place dans l’hypothèse d’un système conceptuel autonome. Les trois acceptions ci-dessus correspondent à trois facettes des entités internes au système conceptuel.



Sachant que le système conceptuel possède trois interfaces, avec la perception, avec le raisonnement et avec le langage, il n’est pas étonnant que les concepts soient dotés d’un ancrage, d’un rôle inférentiel et d’un caractère compositionnel. Étant donné que les concepts ne sont pas directement observables, les différents auteurs ont utilisé l’une ou l’autre de ces trois interfaces pour tenter de caractériser les concepts. Nous organiserons donc notre revue sélective des textes selon ces trois interfaces. Nous ne chercherons pas à être exhaustive. Nous retiendrons les théories des concepts susceptibles d’étayer la réflexion que nous mènerons dans les parties suivantes sur le problème de la construction du sens.

Chapitre 3 : **Ancrage des concepts**

Introduction

Les concepts ne sont pas des entités suspendues dans un éther cognitif. En tant que supports mentaux de la signification, ils constituent un intermédiaire entre un mot et l'objet ou la situation que ce mot permet de désigner dans le monde perçu. De ce fait, les concepts doivent être ancrés dans nos expériences. En d'autres termes, le système conceptuel, qui a en charge la construction du sens et héberge les significations, doit posséder une interface avec la perception pour que les significations puissent porter sur les données de nos expériences.

3.1. Les données du problème

À partir du moment où les concepts sont considérés comme des représentations mentales, le problème de leur ancrage se pose. Or, la manière dont cet ancrage peut être réalisé fait l'objet de débats opposant des positions bien distinctes. Ces débats touchent à des questions fondamentales comme la nature du contenu et de la forme des représentations conceptuelles, et sont indissociables du type de modélisation auquel on souscrit. Dans ce chapitre, notre objectif n'est pas d'argumenter et de prendre parti, mais plutôt de présenter certaines positions présentes dans la littérature concernant cette question centrale de l'ancrage. Nous ne chercherons pas à être exhaustive, ne retenant ici que les théories qui interfèrent avec notre problématique. Nous commençons, dans cette section, par clarifier certains points de vocabulaire.

Représentations, modélisation et ancrage

La notion de représentation est avant tout un outil théorique pour l'explication du fonctionnement cognitif. La variété des comportements humains ne saurait être expliquée par une simple correspondance avec des configurations de stimuli. La majorité des modèles de la cognition humaine utilisent la notion de représentation pour médiatiser le lien complexe entre l'expérience du sujet et le comportement observable qu'il produit. Toutefois, ces modèles se divisent en deux groupes, selon le type de calcul qu'ils postulent, si bien que la notion de représentation est utilisée dans deux sens radicalement différents.

Standard explanations of how systems come to exhibit sophisticated cognitive performances advert to internal representations. Computationalists take representations to be static configurations of symbol tokens. Dynamicists conceive representations very differently. They find their representations among the kind of entities that figure in [dynamic system theory] [...] Unlike digital computers, dynamical systems are not inherently representational. A small but influential contingent of dynamicists have found the notion of representation to be dispensable or even a hindrance for their particular purposes. (VAN GELDER 1998 [105] p. 622)

Un modèle computationnel de la cognition repose sur l'application séquentielle de règles explicites à des représentations symboliques. Les représentations jouent le rôle de *tokens*, de simples "jetons" manipulés par le système d'après leur forme. Les règles, dans la mesure où elles sont explicites, sont elles-mêmes constituées de *tokens*. Ces *tokens* sont supposés avoir une existence physique identifiable dans le système (NEWELL 1980 [78]). Cette nécessité de considérer l'implantation matérielle des *tokens* mise à part, un modèle

symbolique constitue une description qui se situe à un niveau relativement abstrait, celui des manipulations symboliques formelles telles que celles qu'effectue une machine de Turing. Cependant, du fait de l'existence matérielle des symboles manipulés, un tel modèle est bien davantage qu'un simple outil théorique permettant de prédire l'évolution du système. Il existe un isomorphisme supposé entre les opérations décrites dans le modèle et les mécanismes physiques qui se déroulent dans le système. Dans un tel schéma, la notion de représentation est donc incontournable. Non seulement les *tokens* constituent les ingrédients fondamentaux du modèle, mais le modèle prédit même la possibilité de caractériser ces *tokens* de manière indépendante, par exemple par une technique de neurophysiologie.

Cette manière de concevoir le traitement cognitif introduit automatiquement une distinction de type syntaxe *versus* sémantique. La manipulation des représentations est une mécanique formelle, une syntaxe, qui dépend exclusivement de la structure de ces représentations. Pour que cette mécanique ne tourne pas à vide, il faut bien que ces représentations soient ancrées, soit dans les mécanismes perceptifs, soit dans les mécanismes effecteurs. Elles possèdent, de ce fait, une sémantique : elles sont interprétables dans le domaine des perceptions ou dans celui des actions.

We need the syntactic or the symbolic level because we must preserve certain interpretations over mental operations [...] This we can do only if we have a semantic function whose definition has access to the generative structure of the symbolic expressions. [...] To count as a computation rather than simply any functionally described physical system, [a machine] must contain symbols that are interpreted. [...] This quality of symbols and of computational states, whereby they can consistently be given a semantic interpretation, is not the only thing that makes useful computation possible; but it is one of the most important characteristics shared by computation and cognition.

(PYLYSHYN 1984 [88] p. 62)

Les propriétés sémantiques doivent être compatibles avec la syntaxe : à partir d'un premier ensemble de représentations directement ancrées, la mécanique syntaxique engendre de nouvelles représentations qui peuvent être, à leur tour, sémantiquement interprétées. Cette contrainte impose que les représentations intermédiaires, celles qui interviennent dans la chaîne de traitement symbolique, soient également interprétables, même si elles ne sont pas directement associées, par le biais de l'expérience, à la perception ou aux actions¹ (HARNAD 1990 [46]). Comme elles possèdent une structure formelle interne, elles peuvent hériter leur ancrage de leurs constituants et du calcul qui les a produites.

What is the representation of a zebra? [In our example] it is just the symbol string "horse & stripes". But because "horse" and "stripes" are grounded in their respective iconic and categorical representations, "zebra" inherits the grounding, through its grounded *symbolic* representation. In principle, someone who had never seen a zebra (but had seen and learned to identify horses and stripes) could identify a zebra on first acquaintance armed with this symbolic representation alone (plus the nonsymbolic representations of horses and stripes that ground it). Once one has the grounded set of elementary symbols provided by a taxonomy of names (and the iconic and categorical representations that give content to the names and allow them to pick out the objects they identify) the rest of the symbol strings of a natural language can be generated by symbol composition alone, and they will all inherit the intrinsic grounded of the elementary set.

(HARNAD 1990 [46] p. 343)

¹ Cela n'empêche pas l'introduction de tokens non interprétables en tant que tels. Par exemple, dans un système logique classique, les symboles de prédicats ou les constantes reçoivent une interprétation directe, au contraire des symboles logiques dont la fonction se limite à diriger l'interprétation des expressions complexes.

La situation est tout autre si l'on adopte une théorie dynamique de la cognition. Les théories dynamiques considèrent le modèle formel comme un outil de prédiction qui simule le système physique, sans postuler l'existence d'un isomorphisme. Ainsi, les équations de Kepler ne sont pas censées posséder la moindre contrepartie matérielle dans le système des planètes. La description formelle à l'aide de symboles désignant la position d'une planète, sa vitesse ou l'aire du rayon vecteur est propre au modèle, et ne s'applique que métaphoriquement au système matériel. La notion de représentation, contrairement aux modèles symboliques, est ici facultative. Une partie des défenseurs des systèmes dynamiques sont même anti-représentationnalistes, considérant que rien, dans le système, n'est interprétable dans les termes d'un domaine extérieur au système.

L'idée fondamentale est donc que les facultés cognitives sont inextricablement liées à l'historique de ce qui est vécu, de la même manière qu'un sentier au préalable inexistant apparaît en marchant. L'image de la cognition qui s'ensuit n'est pas la résolution de problèmes au moyen de représentations, mais plutôt le faire-émerger créateur d'un monde, avec la seule condition d'être opérationnel : elle doit assurer la pérennité du système en jeu.

(VARELA 1988 [106] p. 111)

Cependant, une telle position ne représente pas un consensus parmi les défenseurs de l'approche dynamique. Notamment, certains modèles connexionnistes s'en éloignent en introduisant la notion de représentation distribuée.

While [behaviorist accounts of behavior] do involve simple mechanisms of learning, there is a crucial difference between our models and the radical behaviorism [...]. In our models, we are explicitly concerned with the problem of internal representation and mental processing, whereas the radical behaviorist explicitly denies the scientific utility and even the validity of the consideration of these constructs. [Our connectionist models] all concern internal mechanisms for activating and acquiring the ability to activate appropriate internal representations.

(RUMELHART & MCCLELLAND 1986 [93] p. 121)

Certaines caractéristiques des modèles dynamiques méritent l'étiquette de représentation. On peut appeler représentation l'état global du système, celui qui résume à lui seul tout son passé. Ainsi, dans un système de planètes considéré du point de vue de la mécanique newtonienne, la position du centre de gravité des planètes et leur vitesse suffit à résumer l'état du système et à permettre la prédiction de ses états futurs. Dans un modèle de réseau de neurones, il en est de même de l'ensemble de valeurs constitué par les poids synaptiques et les activités des neurones. Il y a dans ce cas une seule représentation globale, dont les sous-ensembles peuvent être également considérés comme des représentations. Ainsi, les données relatives à une planète constituent la représentation associée à cette planète. De même, on peut appeler représentation l'état d'activité d'un sous-ensemble de neurones.

Should we identify one's conceptual framework with the configuration of synaptic *weights* in one's brain? Or with the *partitions* they effect across the activation vector space of the assembled neurons to which they connect? Or perhaps with the overall input-output *function* that the network comes to instantiate? The weights uniquely dictate both the partitions and the function, but despite the functional primacy of the weights, there are good reasons for identifying the partitions, and the function they serve, as reflecting most directly the antecedent notion of a "conceptual framework". [...] At least for now, therefore, let us adopt the partitions and the functions they serve as the closest available neural analogue of what the philosophical tradition conceives as our "conceptual framework".

(CHURCHLAND 1989 [19] p. 232)

Dans un réseau de neurones sujet à des phénomènes de synchronisation, le choix du support matériel des représentations laisse moins de doute. Tout ensemble synchrone de neurones peut être considéré, de manière fort naturelle comme une représentation, et ceci d'autant plus facilement que la synchronisation est parfois invoquée comme corrélat de la conscience (CRICK & KOCH 1990 [23]).

Les systèmes dynamiques comme les réseaux de neurones présentent la propriété de posséder des attracteurs. Ainsi, un réseau de neurones placé dans certaines conditions initiales, va évoluer spontanément vers un état d'équilibre. Dans la mesure où ils sont façonnables par apprentissage, ces attracteurs sont interprétables par un observateur : ils sont ancrés dans la configuration d'entrée du système qui a permis leur stabilisation par apprentissage (HOPFIELD 1982 [48]). Il est donc intéressant de considérer que les attracteurs d'un tel système dynamique constituent autant de représentations.

Ce type d'ancrage, propre aux systèmes dynamiques, a cependant ses limites. C'est ainsi que pour certains auteurs, la sémantique des attracteurs a peu de choses à voir avec les entrées du système. Le rôle de ces entrées est ramené à celui de simples perturbations. Le système possède un certain nombre d'attracteurs, mais ces attracteurs ne dépendent que de manière fortuite de l'histoire des interactions du système avec l'extérieur. D'où l'idée selon laquelle le système crée des significations pour lui-même (VARELA 1988 [106]). Certains attracteurs peuvent sembler posséder une signification pour la raison qu'ils sont corrélés à des configurations de stimuli, mais ce genre de signification n'existerait que dans l'œil du modélisateur (SKARDA & FREEMAN 1987 [95]). Selon cette manière de concevoir les systèmes cognitifs, il ne saurait donc y avoir d'ancrage proprement dit.

Lorsque l'on s'intéresse au langage, il est difficile de souscrire d'emblée à la position qui vient d'être invoquée. Il semble que les mots perçus dans le message de l'interlocuteur aient un effet moins aléatoire que celui de simples perturbations. Puisque l'interaction langagière est possible, il est permis de supposer qu'elle ne résulte pas de la confrontation entre deux systèmes "autistes" ne connaissant que leurs propres significations internes. Nous considérons donc, dans ce document, que l'hypothèse représentationnaliste mérite d'être explorée, que ce soit sous sa forme computationnelle ou dans sa réalisation dans les systèmes dynamiques. Nous nous intéresserons donc à des systèmes dont certains états peuvent être qualifiés de représentationnels. Nous nous demanderons comment ces états peuvent entretenir un rapport d'ancrage avec d'autres états du système ou avec certaines configurations des entrées et des sorties. Nous constaterons, dans ce chapitre, que cette question de l'ancrage est particulièrement problématique.

Représentations : le contenu

La notion de représentation constitue un outil commode, peut-être incontournable, pour élaborer et exprimer un modèle de la sémantique du langage. Cependant, l'enjeu dépasse la question de la modélisation du langage. La question de la nature des représentations est au centre d'un débat philosophique concernant leur propriété intentionnelle, propriété attribuée à certains phénomènes cognitifs. Certains états mentaux semblent posséder la propriété remarquable d'être obligatoirement à propos de quelque chose. Ainsi, une grande partie des états conscients, par exemple l'état correspondant au fait de voir orange ou d'avoir mal, sont à propos de quelque chose : ils ont un contenu. Ainsi, l'état conscient associé à la couleur orange sera lié à la vision d'un livre ou de la surface d'un mur ; une douleur donnée sera ressentie comme si elle se trouvait dans le pied ou dans une dent. D'autres types d'états mentaux possèdent manifestement cette propriété d'avoir un contenu. Par exemple, une croyance ou un désir n'existe cognitivement, la plupart du temps, que si elles sont à propos d'un état de choses, par exemple le fait de réussir un examen. La question de l'intentionnalité

se pose donc pour plusieurs types d'états mentaux, parmi lesquels les représentations mentales associées aux expressions langagières occupent une place de premier plan. Ainsi, la représentation mentale associée au mot *arbre* semble être à propos de quelque chose : l'image d'un arbre qui se trouve dans la cour ou cet arbre lui-même. Dans le cas des représentations mentales impliquées dans le traitement sémantique du langage, la propriété d'intentionnalité, le fait de posséder un contenu, permet d'expliquer la possibilité même de la communication : si les expressions langagières renvoient à des représentations mentales qui ont un contenu, les expressions langagières servent à communiquer à propos de ce contenu. Comme nous allons le constater, les choses ne sont pas si simples.

Il est important de noter que la notion technique de représentation n'est pas obligatoirement liée à celle de contenu. Il est par exemple courant, dans les modèles de la syntaxe, de postuler l'existence de représentations dont la raison d'être est interne au système et qui n'ont pas de contenu. C'est ainsi que le "pronom" *PRO* est invoqué pour expliquer certains phénomènes syntaxiques, notamment l'existence des impersonnels. Dans la phrase *il pleut*, l'impersonnel *il* donne une réalité phonologique au pronom *PRO*, ce qui n'est pas le cas dans les langues dites *pro-drop* comme le persan. Le pronom *PRO* est une représentation, un *token* qui entre dans le calcul formel proposé par le modèle et dont la présence semble incontournable pour expliquer certains phénomènes syntaxiques. En tant que représentation, le pronom *PRO* est supposé avoir une réalité cognitive. Pourtant, rien ne permet, dans le modèle, de lui allouer un contenu.

À l'autre extrême, on pourrait penser que certaines représentations ont une intentionnalité tellement affirmée que leur contenu, ce sur quoi elles portent, se retrouve dans leur forme. Selon cette idée, la représentation serait une re-présentation, une sorte de copie d'une autre chose qui lui préexisterait. Cette idée n'est pas entièrement sans fondement. On sait par exemple que certaines propriétés topologiques sont conservées entre la rétine, les relais thalamiques et les différentes cartes corticales du système visuel, si bien que l'image d'une pomme qui se forme sur la rétine est, par certains aspects, copiée dans certaines aires du cortex. Si l'on conçoit le fonctionnement cognitif comme un traitement de type purement analogique, ces images corticales peuvent constituer d'authentiques re-présentations qui peuvent déclencher d'autres traitements, par exemple une réponse motrice consistant à cueillir la pomme. Cependant, de nombreux autres types de traitements, notamment les traitements de type computationnel, supposent des niveaux de représentation dans lesquels les différentes parties des images corticales sont intégrées et ne sont pas conservées. Dans ce cas, la représentation n'entretient plus aucun rapport de ressemblance avec l'image rétinienne. Dans la plupart des modèles, l'intentionnalité des représentations ne suppose pas que la forme de la représentation re-présente son contenu. On doit plutôt concevoir le lien intentionnel comme un renvoi : la représentation renvoie à son contenu, il est possible d'obtenir l'un à partir de l'autre, sans pour autant que les deux se ressemblent. Dans ce cas, l'intentionnalité rejoint d'emblée la notion d'ancrage. Il est important d'insister sur le fait qu'il est parfaitement possible de postuler l'existence, dans le système cognitif, de représentations dotées d'une intentionnalité sans que celles-ci soient des re-présentations.

La question que nous essayons de traiter dans ce travail concerne la nature des représentations mentales associées aux expressions langagières. S'agit-il de simples pions manipulés selon une mécanique aveugle ? Dans ce cas, le caractère intentionnel d'une expression langagière pourrait n'être qu'une illusion. Il serait produit *post hoc* par une opération d'interprétation. La phrase *donne-moi le crayon qui est sur la table* exprime un souhait à propos de la possession du crayon en question. Ce souhait pourrait n'être qu'une interprétation produite par l'interlocuteur et n'avoir aucune réalité cognitive interne à l'esprit du locuteur, pas plus que le souhait exprimé par un ordinateur qui affiche la phrase *j'ai faim* sur son écran. Le caractère intentionnel d'une expression langagière serait ainsi dérivé de l'acte

interprétatif de l'observateur et n'exprimerait aucune propriété intrinsèque au processus cognitif qui a produit cette expression (DAVIDSON 1984 [25]). Sur de telles bases, une théorie du langage se contentera de chercher le caractère intentionnel de la communication langagière dans l'acte de traduction dans une autre langue, dans la prédiction des comportements ou dans l'attribution des valeurs de vérités aux énoncés par l'auditeur.

Ce type d'approche peut convenir pour une théorie behavioriste de la communication langagière. Malheureusement, le pouvoir prédictif que l'on peut attendre d'une telle théorie est limité, si on le compare à celui des théories représentationnalistes. Lorsqu'un locuteur énonce la phrase donne-moi le crayon qui est sur la table, sa phrase ne saurait être le produit d'un simple processus associatif. Il est plausible de supposer qu'elle résulte d'un processus impliquant des représentations. Par exemple, le locuteur désire écrire une lettre, se retrouve dans l'embarras de ne pas avoir de crayon, aperçoit le crayon qui est sur la table, et forme le plan qui consiste à demander à l'interlocuteur un petit service. Les éléments de ce processus, tel qu'on peut scientifiquement le reconstituer, se situent au même niveau que la description qui vient d'être esquissée. Il est, de plus, plausible de postuler qu'il s'agit de représentations qui renvoient aux objets perçus comme lettre, crayon, table. Les règles qui dirigent ce calcul peuvent s'exprimer au même niveau. Par exemple, une telle règle stipule de résoudre toute incompatibilité détectée entre attitudes propositionnelles. Ainsi, le locuteur résout une incompatibilité entre le désir d'écrire la lettre et le constat qu'il n'en a pas les moyens matériels. Si l'interlocuteur réagit en prononçant la phrase ce n'est pas un crayon, mais un porte-mine, c'est qu'il conçoit une incompatibilité entre ce qu'on lui dit et ce qu'il croit pour l'avoir perçu. De même, l'interlocuteur qui répond par la phrase il vaut mieux écrire ta lettre au stylo résout l'incompatibilité entre la croyance que la lettre va être écrite au crayon et le désir de conférer à la lettre un caractère plus officiel. De tels calculs opèrent sur des représentations dotées d'une propriété intentionnelle. Renoncer à l'existence cognitive de telles représentations conduit à renoncer au pouvoir prédictif de ce type de calcul.

À partir du moment où l'on recherche un modèle de la génération et de la compréhension des expressions langagières qui met en jeu plusieurs niveaux de représentations, on peut s'interroger sur l'intentionnalité propre à chaque niveau. Comme nous venons de le voir, le niveau où les phrases sont interprétées suppose une forme d'intentionnalité, celle des attitudes propositionnelles. On pourrait envisager d'en rester là, en supposant que toutes les représentations qui se situent en-dessous, les représentations associées aux syntagmes, aux mots, aux syllabes, *et cætera*, sont dépourvues d'intentionnalité. Dans ce texte, nous allons considérer en détail l'hypothèse selon laquelle il existe des représentations mentales associées aux entités lexicales, les concepts, possédant un caractère intentionnel propre, qui peut être qualifié de sémantique². Les entités sub-lexicales, comme la syllabe, peuvent recevoir d'autres types de contenu, par exemple une forme acoustique. Le point de vue que nous adopterons dans la majeure partie de ce texte est que les concepts sont les représentations mentales minimales dotées d'une sémantique, c'est-à-dire qu'ils s'expriment dans le monde perçu.

On all hands, then, concepts serve both as the domains over which the most elementary mental processes are defined, and as the most primitive bearers of semantic properties. Hence their centrality in representational theories of mind. (FODOR 1994 [36] p. 96)

Une position fréquemment adoptée consiste à considérer que le contenu des concepts se situe dans le "monde", considéré comme objectivement accessible. Cette position semble raisonnable, dans la mesure où l'on utilise les concepts pour communiquer à propos des

² Nous reviendrons sur ce point en introduisant la notion plus stricte de concept lexical (CF. CHAPITRE 6).

entités et des situations du monde. Les concepts, en tant que représentations mentales, auraient ainsi cette propriété de renvoyer à des entités indépendantes de l'esprit humain. Le postulat de l'existence de ce type d'entité est qualifié de réaliste. Les candidats pour ce type d'entité sont des objets concrets, des objets abstraits, des classes d'objets, des attributs d'objets, des propriétés, des faits, des états de faits, *et cætera*. La question de l'existence indépendante de ces entités et de leurs qualités est une question métaphysique qui divise les penseurs en deux groupes, selon qu'ils sont réalistes ou non par rapport aux catégories universelles dans le monde. La contrepartie sémantique de la question revient à se demander si l'on établit un lien entre les concepts et de telles entités externes, ce qui constitue le point de vue externaliste, ou si le contenu des concepts est lui-même une entité mentale, ce qui constitue le point de vue internaliste.

Nous n'entrerons pas dans le débat métaphysique du réalisme, car son issue n'a pas de conséquence sur notre thèse. Nous choisissons, dans ce texte, d'éviter l'hypothèse de l'existence d'un monde structuré indépendant de l'esprit. En revanche, nous adopterons une position, empruntée à la tradition kantienne, qualifiée de réalisme empirique. Le réalisme empirique affirme l'existence des entités mentales qui consistent en notre expérience. Nous parlons ici des expériences au sens large, ce qui inclut la perception mais aussi les sensations et les attitudes. En ce qui concerne le monde extérieur, la seule conviction que nous pouvons avoir quant à son existence est liée au sentiment intuitif que quelque chose nous résiste. L'intuition d'une résistance qui contrarie nos souhaits ou notre volonté nous conduit à faire l'hypothèse d'une réalité extérieure, sans que nous ne puissions dire quoi que ce soit sur la cause précise de cette résistance. Les seules informations auxquelles nous avons accès, en tant que sujet connaissant, nous sont fournies par l'état d'activité de nos capteurs sensoriels. C'est à partir de ces informations et à propos de ces informations que nous formons des expressions langagières.

En marge de cette question métaphysique se pose le problème de choisir entre les versions internaliste et externaliste de l'explication de l'intentionnalité des concepts. Les concepts, en tant qu'entités mentales intentionnelles, portent-ils sur d'autres entités mentales ou portent-ils sur des entités qui sont dans le monde ? Il ne s'agit pas de répondre à cette question dans l'absolu. Lorsque l'on aborde cette question sous l'angle de la modélisation, la tâche est d'expliquer l'établissement du lien intentionnel pour un sujet qui est le produit exclusif d'une phylogenèse et d'une ontogenèse, en renonçant à toute autre source d'information. Lorsque l'on parle de l'objet "ballon" ou de la propriété "rouge" ou de la situation "jeu", il faut lier les concepts correspondants, *id est* les représentations mentales associées à ces mots, à des phénomènes vécus par le sujet. Là encore, l'hypothèse d'un monde structuré peut être évitée, puisque le sujet doit l'intentionnalité de ses représentations mentales à ses dispositions biologiques et aux seules données contenues dans ses expériences³. Nous sommes donc contrainte de renoncer, pour des raisons épistémologiques, aux facilités que procure l'externalisme pour adopter une position que nous appelons internalisme méthodologique : les concepts acquièrent leur contenu dans le domaine des entités mentales associées aux expériences du sujet⁴. Cette position ne présuppose aucun jugement concernant la correspondance entre les concepts et le monde. Elle consiste à proposer d'expliquer la mise en place des concepts par le seul fait des dispositions mentales et des expériences du sujet. C'est dans cette perspective que la notion d'intentionnalité rejoint celle d'ancrage. Le caractère intentionnel des concepts résulte ainsi du fait qu'ils sont, en tant que représentations mentales, ancrés dans les données de l'expérience.

³ Nous discutons plus loin la possibilité d'une intentionnalité innée qui serait le résultat de la phylogenèse, pour conclure que cette possibilité ne change pas fondamentalement le point de vue que nous adoptons ici.

⁴ Nous démarquons plus loin notre position de celle qui est habituellement appelée internaliste.

Dans la perspective de l'internalisme méthodologique, l'intentionnalité des concepts se présente comme un phénomène d'évocation. Les concepts sont intentionnels car il existe des expériences qui les évoquent et qu'ils évoquent. La question de l'intentionnalité des concepts se ramène donc, du point de vue de la modélisation, au problème d'expliquer l'interface entre le système conceptuel et d'autres systèmes de la cognition, notamment la perception, qui produisent les entités mentales associées aux expériences. Cette manière de procéder peut sembler en contradiction avec la position externaliste. C'est en partie le cas. Ceux qui adoptent le point de vue réaliste considéreront qu'il existe trois domaines : les choses du monde, les expériences et les concepts. Ils peuvent imaginer deux interfaces, entre le monde et les expériences d'une part, entre les expériences et le système conceptuel d'autre part. Le projet externaliste consiste notamment à établir un lien intentionnel objectif entre les concepts et le monde en composant ces deux interfaces. Ainsi, là où l'internalisme méthodologique explore des mécanismes d'évocation des concepts dans les données de l'expérience, le point de vue externaliste propose des conditions d'application des concepts aux entités du monde extérieur. Cependant, la définition de la première interface, entre le monde et les expériences, suppose que l'on dispose d'un "oracle", un regard externe qui embrasse les deux domaines. Dans une perspective matérialiste, cet oracle est habituellement constitué par la science. Ceci signifie que les catégories attribuées au monde sont fixées par nos connaissances scientifiques actuelles ou potentielles. Il ne reste ensuite qu'à établir un lien entre ces catégories "réifiées" et la cognition. Une telle méthodologie est toutefois problématique, si l'on admet que la science n'est qu'un produit de la cognition humaine. Il y a donc pétition de principe, puisque c'est l'analyse des capacités cognitives qui doit nous amener aux lois qui gouvernent l'apparition de ces catégories du savoir. On ne saurait ancrer ce qui constitue une partie fondamentale de la cognition humaine, le système conceptuel, dans les catégories provisoires produites par l'activité de la communauté scientifique. Lorsque l'on cherche à élaborer un modèle plausible de la compétence sémantique des humains, on ne peut pas accepter que les nouvelles théories que des physiciens des particules peuvent produire dans leur laboratoire puissent avoir une influence quelconque sur cette compétence. Toutefois, les théories externalistes nous intéresseront dans la mesure où, en dédoublant l'interface concept - monde, nous pouvons leur donner une interprétation internaliste, en particulier en ce qui concerne l'acquisition des concepts. Là où une théorie externaliste postule l'existence d'entités indépendantes de la cognition pour établir un lien intentionnel avec les concepts nouvellement acquis, nous ne nous intéresserons qu'au seul lien d'ancrage entre les concepts et les expériences que ces entités indépendantes sont supposées susciter.

On pourrait reprocher à l'internalisme méthodologique de supprimer la notion de référence. Dans un cadre externaliste, un mot du lexique renvoie à une entité du monde qui constitue la référence de ce mot. L'existence des références externes pour les mots est une hypothèse forte, motivée entre autres par le souci d'expliquer que la communication au moyen du langage soit possible. L'efficacité du langage, dans ce cadre, s'explique de la manière la plus simple qui soit : en prononçant un mot, le locuteur désigne une entité du monde à son interlocuteur. Ainsi, le projet externaliste peut se focaliser, d'une part, sur l'établissement du lien référentiel des mots aux entités du monde, et d'autre part, sur la propagation de ce lien à toutes les expressions langagières. En supprimant toute référence directe au monde, l'internalisme méthodologique semble renoncer à expliquer la communication. Si les mots n'évoquent que des représentations mentales, si l'on ne peut pas se référer à des entités du monde, il semble impossible d'établir une corrélation entre les états mentaux des interlocuteurs. Nous considérons cependant que cette conclusion est erronée et que le problème de la communication langagière mérite précisément une analyse d'un point de vue strictement cognitif, sans un recours aux postulats métaphysiques sur l'existence des entités indépendantes de l'esprit humain. Nous tâcherons plus tard de développer notre point

de vue sur ce sujet (CF. CHAPITRE 9). Dans le cadre que nous adoptons, la notion de référence externe est remplacée par une référence interne. Par exemple, la référence du mot pomme pourra être, dans certains contextes, une image ou un goût, plutôt qu'un élément de nomenclature botanique comme l'exigerait sans doute une position externaliste matérialiste.

La question du contenu d'un concept porte donc sur la référence du terme auquel est associé ce concept. Cette référence appartient à un autre domaine que celui des concepts, le domaine général des expériences. Or, le fait que le concept soit ainsi ancré ne nous renseigne pas sur sa propre nature. Après avoir considéré la cible de l'ancrage, nous abordons maintenant la question d'analyser la forme de ce qui est ancré.

Représentations : la forme

Les modèles représentationnalistes postulent l'existence de représentations de manière à expliquer et à prédire le comportement du système cognitif. Dans la mesure où l'on cherche à décrire ce comportement à l'aide de mécanismes généraux, il est indispensable que les représentations soient dotées d'une structure interne. C'est le moyen par lequel ces mécanismes peuvent les discerner. Lorsque des mécanismes généraux agissent sur des représentations, ils le font d'après leur forme, *id est* d'après leur structure interne, ce qui permet de les différencier des autres représentations. Dans un modèle dynamique, la structure de la représentation résume une partie de l'état du système et est nécessaire, au sein du modèle, pour prédire son trajet avenir. Dans un système computationnel, la notion de forme est encore plus importante, puisqu'elle est supposée avoir une contrepartie matérielle. Les règles du modèle, qui ne sont pas limitées à un simple moyen de description mais sont censées être constitutives du système, manipulent des représentations, elles-aussi matériellement constitutives du système, d'après leur forme. On peut, dans ce cas, considérer que la forme des représentations résume leur propriété causale au sein du système.

Il est difficile de concevoir une théorie de la communication langagière qui soit non représentationnaliste, tant on est frappé par l'apparente complexité des processus en jeu. L'introduction de représentations conceptuelles est, en premier lieu, motivée par le désir, assez naturel, de faire figurer dans le modèle une contrepartie de la notion de signification lexicale qui, nous l'avons vu, est assumée par la propriété intentionnelle des concepts. Un deuxième intérêt d'introduire des représentations conceptuelles est de doter le modèle d'un substrat sur lequel les mécanismes de combinaison sémantique peuvent opérer. À l'agencement des mots dans une expression langagière peut correspondre une combinaison de significations lexicales qui permet de doter l'expression d'un sens. Ainsi, les représentations conceptuelles acquièrent un caractère compositionnel : elles entrent dans des combinaisons pour engendrer d'autres représentations plus complexes. La manière dont les représentations conceptuelles se combinent dépend de leur forme, et la forme du résultat est construite à partir de la forme des constituants. Ainsi, les concepts, de par leur forme, possèdent une propriété causale.

Une autre raison d'introduire des représentations conceptuelles pour expliquer le traitement cognitif du langage est liée au souci de reproduire les inférences qui sont faites spontanément par les sujets. Les représentations conceptuelles semblent présenter un rôle inférentiel : la représentation associée à une expression langagière peut déclencher d'autres représentations, par exemple par un processus de déduction. Là encore, il est raisonnable de tenter d'expliquer les inférences de manière causale à partir de la forme des représentations conceptuelles qui les déclenchent.

La forme des concepts peut donc être invoquée pour expliquer l'interface avec le langage et l'interface avec le raisonnement. Il est même possible de faire reposer sur la forme des concepts leur signification elle-même : pour certains auteurs, la caractérisation de la

forme des concepts est suffisamment contrainte par les interfaces avec le langage et le raisonnement, si bien que toute idée d'ancrage devient superflue. Dans ce cas, la signification des concepts, au moins en partie, est ramenée à leur seule propriété causale. Ce courant de pensée est qualifié d'internaliste. Dans sa version radicale, l'internalisme conceptuel ignore toute idée d'interface avec l'expérience, si bien que le système conceptuel apparaît comme un jeu dépourvu d'ancrage. Il existe cependant des versions modérées de l'internalisme conceptuel dans lesquelles la forme des représentations conceptuelles contribue partiellement à l'explication de l'intentionnalité. Si l'on peut encore parler de contenu dans le cas où l'on s'intéresse aux rapports que les concepts entretiennent entre eux, il s'agit d'un contenu étroit.

The idea of the two-factor version is that there are two components to meaning, a conceptual role component that is entirely "in the head" (this is narrow meaning) and an external component that has to do with the relations between the representations in the head (with their conceptual roles) and the referents and/or truth conditions of these representations in the world. [...] The internal factor, conceptual role, is a matter of the causal role of the expression in reasoning and deliberation and, in general, in the way the expression combines and interacts with other expressions so as to mediate between sensory inputs and behavioral outputs. (BLOCK 1986 [7] p. 627)

Dans cette approche, qualifiée de sémantique du rôle conceptuel (*Conceptual Role Semantics*), le concept est déterminé par sa fonction au sein du système conceptuel, et c'est cette fonction qui lui procure, en partie, sa propriété intentionnelle. Il s'agit là d'un changement de perspective par rapport à la manière dont nous avons introduit la notion de concept. Nous avons caractérisé les concepts en tant que représentations dotées d'un contenu dans le domaine de l'expérience, avant de leur conférer une fonction au sein du système conceptuel. L'internalisme conceptuel, quant à lui, procède de manière inverse en caractérisant les concepts par le biais de la fonction qu'ils remplissent dans le comportement de l'individu qui les possède.

A possession condition for a particular concept specifies a role that individuates that concept. The possession condition will mention the role of the concept in certain transitions that the thinker is willing to make. These will be transitions that involve complete propositional thoughts involving the concept. In some cases they are inferential transitions; in others they are transitions from initial states involving perceptual experience. Normally, a possession condition has several clauses, each treating of a different kind of case. (PEACOCKE 1992 [81] p. 107)

Les concepts logiques servent de modèle à cette approche. Par exemple, le concept de conjonction peut être caractérisé par sa fonction dans la réalisation de certaines inférences par les individus qui le possèdent (PEACOCKE 1992 [81]). Cette fonction peut être représentée par l'ensemble de deux règles d'élimination ($P \wedge Q \vdash P$ et $P \wedge Q \vdash Q$) et une règle d'introduction ($P, Q \vdash P \wedge Q$). La possession du concept de conjonction est ainsi caractérisée d'une manière qui garantit la production de toutes les inférences ayant la forme de l'une des trois règles citées. Si l'on généralise cette méthode, il s'agit de déterminer, pour chaque concept, l'ensemble des clauses nécessaires qui cernent sa fonction cognitive. Cette fonction cognitive s'étend au pouvoir d'évocation du concept qui le lie aux mécanismes perceptifs ou effecteurs.

Certes, il nous paraît naturel que le postulat des concepts s'accompagne d'une description exacte de la fonction qu'ils sont censés remplir dans le système cognitif humain. Cependant, nous pensons que, dans une approche représentationnaliste, cette tâche ne peut être entamée qu'en définissant, de manière claire, d'une part le domaine où les concepts

acquièrent leur contenu, et d'autre part les mécanismes qui agissent sur les concepts d'après leur forme. Dans notre schéma de base, le système conceptuel retient trois interfaces avec la perception, le langage et le raisonnement. Notre internalisme méthodologique exige que les concepts acquièrent leur contenu dans le domaine de l'expérience. Ainsi seule l'interface avec la perception fournit aux représentations conceptuelles l'ancrage nécessaire pour qu'ils possèdent une propriété sémantique. Les deux autres interfaces, quant à elles, permettent d'établir des liens associatifs, des liens de déclenchement par appariement ou des liens de contrôle, qui lient les mécanismes à l'œuvre dans le système conceptuel aux mécanismes du langage et du raisonnement. C'est l'explication de ces mécanismes qui nécessite l'analyse de la forme des représentations conceptuelles⁵. Or, selon l'hypothèse de l'internalisme conceptuel, l'ensemble des relations que les représentations conceptuelles peuvent entretenir avec d'autres entités mentales intervient dans la détermination de leur contenu. Notamment, les relations inter-conceptuelles prennent un rôle déterminant dans la définition du contenu étroit des concepts. Nous estimons qu'il s'agit ici d'un emploi métaphorique de la notion de contenu : dans un modèle représentationnaliste du système conceptuel, les relations inter-conceptuelles ne peuvent intervenir que dans la détermination de la structure interne des représentations conceptuelles ; ce n'est que d'une manière indirecte que cette structure peut être pertinente pour l'ancrage des représentations structurées⁶. Répétons que ce que nous appelons internalisme méthodologique n'est pas lié par l'exigence posée par l'internalisme conceptuel qui veut que le contenu des concepts soit déterminé, pour l'essentiel, de manière fonctionnelle. Notre position laisse par exemple la possibilité que chaque concept soit directement ancré dans l'expérience.

La notion de forme, en conférant un caractère fonctionnel aux représentations conceptuelles, réintroduit l'idée du sens des mots. La référence d'un mot est donnée par le contenu du concept correspondant. Son sens est donné, concrètement, par la forme de ce concept. La forme du concept précise la fonction du mot dans la communication langagière, c'est-à-dire les rapports qu'il entretient par son sens avec les autres expressions langagières. Le sens du mot pomme est ainsi représenté par la forme du concept associé qui lui donne la possibilité d'intervenir dans un syntagme comme une pomme rouge, ou dans une assertion comme la pomme est un fruit sucré.

À partir du moment où les concepts possèdent une forme, il faut se poser la question du rapport entre cette forme et celle de la représentation non conceptuelle qui l'évoque et qu'il évoque. On peut imaginer qu'elles soient de même nature, qu'elles soient qualitativement différentes mais que leurs structures se correspondent, ou enfin qu'il n'y ait aucun lien entre les deux. Dans les deux premiers cas, l'ancrage devient trivial. Dans le troisième cas, le problème de l'ancrage se révélera particulièrement délicat à résoudre.

3.2. Percepts intégrés

L'une des fonctions remarquables attribuées aux concepts est celle de la catégorisation. Les concepts nous permettent d'avoir un jugement sur les données de la perception, jugement qui est présenté comme binaire dans certains modèles. Ainsi, le concept LIVRE nous permet de catégoriser les objets du monde perçu de manière à en extraire, de manière plus ou moins précise, tout ce qui mérite d'être qualifié de livre.

⁵ Nous développerons cette analyse en posant des questions sur, d'une part, le rôle inférentiel, et d'autre part, le caractère compositionnel des concepts (CF. CHAPITRES 4 & 5).

⁶ Nous aurons l'occasion de reparler en détail des motifs et des méthodes pour structurer les concepts lexicaux, et des conséquences d'une telle entreprise (CF. CHAPITRES 7).

La manière la plus simple d'expliquer la capacité de catégorisation est sans conteste de l'imaginer comme une activité de comparaison. Ainsi, est une instance du concept LIVRE tout ce qui ressemble au livre typique. La caractéristique fondamentale de cette approche est que les deux éléments de la comparaison, le percept à catégoriser et le concept qui sert d'étalon sont de même nature. Son intérêt est de résoudre de la manière la plus simple qui soit la question de l'intentionnalité : non seulement le percept typique est le meilleur candidat pour constituer le contenu du concept, mais il n'y a plus de raison de distinguer le concept de son contenu. Les concepts, selon cette approche, ne sont rien d'autre que des percepts moyens, issus d'une intégration d'expériences multiples. Il s'agit donc, à la base, d'une théorie purement empirique des concepts.

Avantages de l'approche empiriste des concepts

L'intérêt premier de l'approche empiriste est de résoudre le problème de l'ancrage. Les concepts étant des percepts moyens, *id est* des invariants perceptifs, les concepts sont *ipso facto* ancrés dans l'expérience. Cette forme d'ancrage présente l'avantage d'être économique : contrairement aux théories qui postulent l'existence d'un espace conceptuel qualitativement distinct de celui des expériences, il n'y a pas, dans l'approche empiriste, de duplication des représentations. Ainsi, nous pouvons exprimer le fait qu'un zèbre possède des rayures parce que les rayures sont présentes dans la perception du zèbre. Alors qu'un système symbolique est obligé de dupliquer cette information sous une forme explicite, l'approche empiriste se contente d'exploiter l'information qui est disponible dans les données de la perception. Du point de vue de la modélisation, il s'agit là d'un avantage considérable, car toute duplication pose le problème de la complétude et de la cohérence.

L'ancrage empiriste résout un autre problème délicat, celui de l'acquisition. L'enfant acquiert de nouveaux concepts en généralisant ses percepts. Pour réaliser cette opération inductive, plusieurs mécanismes simples ont été proposés. Le mécanisme empiriste par excellence est un mécanisme statistique d'extraction de régularités. Ainsi, ne sont retenues dans le percept moyen que les caractéristiques qui sont fréquemment présentes dans les instances à partir desquelles il est construit. C'est ainsi que fonctionnent les réseaux de neurones artificiels classiques, qu'ils soient supervisés ou non.

One of the basic principles which drives learning in neural networks is *similarity*. Similar inputs tend to yield similar outputs. Thus, if a network has learned to classify a pattern, say 11110000, in a certain way then it will tend to classify a novel pattern, *e.g.*, 11110001, in a like fashion. Neural networks are thus a kind of analogy engine. The principle of similarity is what lets networks generalize their behaviors beyond the cases they have encountered during training. (ELMAN & BATES & JOHNSON & PARISI 1996 [32] p. 59)

Cette capacité de généralisation est à la base de la formation des percepts moyens. Le système extrait les régularités des exemples qui lui sont soumis. Les aspects contingents, qui varient d'un exemple à l'autre, ne laissent pas de trace, alors que les aspects invariants sont mémorisés ensemble. Une fois que cet apprentissage a eu lieu, le système peut reconnaître un objet qu'il n'a jamais perçu, à condition que cet objet possède un nombre suffisant de caractéristiques en commun avec l'un des ensembles invariants qui ont été mémorisés. Certains problèmes cognitifs ne possèdent pas cette propriété de continuité selon laquelle des entrées proches correspondent à des sorties proches. Un contre-exemple bien connu est celui de l'induction de la grammaticalité des énoncés du langage (CHOMSKY 1975 [15]). Néanmoins, la formation des catégories semble posséder, pour l'essentiel, cette propriété de continuité que les réseaux de neurones permettent de reproduire.

Noter que l'apprentissage statistique par extraction de régularités n'est pas le seul mécanisme permettant la formation de catégories. La théorie de la Gestalt et le constructivisme peuvent être présentés comme des approches fondées sur des mécanismes d'induction de "bonnes formes" à partir d'un ensemble réduit d'exemples (DESSALLES 1998 [29]). Dans ce cas, les concepts sont ces bonnes formes acquises par l'expérience qui servent ensuite à catégoriser les situations.

Le fait de ramener l'ancrage des concepts à un mécanisme de catégorisation possède plusieurs autres mérites. Le moindre n'est pas de pouvoir gérer la gradualité des jugements de catégorisation, ce que bien des systèmes symboliques sont incapables de faire. Ainsi, bien qu'un gland ne soit pas un chêne, le gland qui germe et qui pousse constitue chaque jour une meilleure instance du concept CHÊNE, sans qu'aucune discontinuité ne soit jamais franchie. De même, les liens de ressemblance qui peuvent unir les concepts, les airs de famille, sont naturellement pris en charge par la ressemblance des représentations dans l'espace perceptif. Ainsi, nous n'avons pas besoin d'explication pour percevoir la ressemblance entre une puce électronique et une vraie puce. Un tel appariement sera beaucoup plus délicat à effectuer pour des représentations conceptuelles ne comportant que des descriptions symboliques explicites.

Inconvénients de l'approche empiriste des concepts

Si, malgré le caractère impressionnant de ces points positifs, de très nombreux auteurs ont imaginé un système conceptuel disjoint du système perceptuel, c'est que les inconvénients de l'approche empiriste leur ont semblé rédhibitoires. Une partie de ces inconvénients est liée aux deux autres interfaces, laissées de côté par l'approche purement empiriste : l'interface avec le langage et l'interface avec le raisonnement.

Le principal problème que la formation empirique des concepts ne permet pas de résoudre est celui des phénomènes systématiques. Les concepts sont impliqués dans des opérations systématiques, que ce soit dans la catégorisation, les inférences ou la composition, qui sont incompatibles avec le caractère fondamentalement graduel inhérent aux mécanismes statistiques postulés dans l'approche empiriste. Par exemple, une entité incluse dans le concept VIVANT doit être systématiquement exclue du concept MORT. Or, pour un système empiriste, toute entité perçue est plus ou moins vivante, certaines franchement, d'autres presque pas, mais le lien entre l'entité perçue et le concept VIVANT n'est jamais strictement nul, car la statistique ménage toujours une ressemblance, même très faible, entre deux représentations quelconques. La frontière entre les concepts VIVANT et MORT n'est donc pas systématique comme il se devrait. De la même manière, la systématisme de l'inférence selon laquelle certaines entités perçues peuvent mourir dès lors qu'elles sont vivantes n'est pas assurée, car l'induction statistique dont elle dépend n'offre aucune garantie. Enfin, la systématisme de composition est également absente : rien ne garantit un sens systématique à une construction comme X est malade dès lors que la variable X désigne une entité vivante.

La nature statistique de l'apprentissage empiriste pose un deuxième problème. Un système empirique est peu sensible aux exceptions et ne leur associe pas d'explication. Dans un réseau de neurones classique, les exceptions n'étant pas représentatives de la classe, sont ignorées en tant que points aberrants de l'échantillon statistique. L'apprentissage d'une exception, par exemple une entité vivante immortelle, demande donc un renforcement répété, d'autant plus important que le lien conceptuel VIVANT - MORT a été fortement imprimé, jusqu'à ce que l'exception puisse constituer une classe à elle seule. Cet apprentissage ne produit jamais un résultat strict, et surtout mémorise l'exception sans pouvoir en stocker la raison. Or, un enfant humain qui lit une fiction comprend aisément, en une seule fois, que tel personnage peut traverser les âges sans vieillir parce qu'il possède un don particulier. Si son

fonctionnement conceptuel obéissait aux seules lois de l'apprentissage statistique empiriste, l'enfant serait incapable de comprendre ce type d'histoire.

Un autre grand problème rencontré par la conception empiriste du système conceptuel est la dépendance du système par rapport à son histoire. Pour que deux individus aient des concepts comparables, il faut qu'ils aient connu des expériences similaires en moyenne, qualitativement et quantitativement. Or, cette exigence est problématique dans les domaines où les individus n'ont eu accès qu'à des expériences limitées. Par exemple, un individu ne pourra pas communiquer à propos de ce qui est nouveau pour les autres. Il devra se limiter aux expériences communes, ce qui offre une piètre modélisation de la communication humaine. Pour expliquer que les enfants parviennent à des états de connaissances similaires malgré des expériences diverses et limitées, certains auteurs ont émis l'hypothèse que ces états de connaissances constituaient de "bonnes formes", dotés par exemple de la propriété de clôture opérationnelle, *id est* de fermeture par composition (PIAGET 1945 [83]). L'intérêt des bonnes formes est que leur apprentissage ne nécessite que peu d'exemples. Par exemple, le concept de justice, développé en plusieurs étapes par les enfants, présente de nombreuses symétries. L'enfant passe d'un état égo-centré, dans lequel est juste tout ce qui lui profite, à un état de justice réciproque dans lequel il considère comme injuste vis à vis des autres ce qu'il n'aimerait pas qu'on lui fasse (PIAGET 1932 [82]). Ce concept de justice, que l'on peut noter juste(acteur , patient , acte), est invariant pour toutes les substitutions d'individus aux deux variables acteur et patient. Ceci inclut en particulier la possibilité d'adopter le point de vue de l'autre, ce dont le très jeune enfant est encore incapable. C'est parce que le concept adulte de justice présente de nombreuses symétries qu'il est atteint de manière identique par des enfants différents, soumis à des expériences différentes. Les concepts produits par généralisation statistique à partir de peu d'exemples constituent également des bonnes formes : la catégorie construite à partir de la généralisation des percepts est invariante pour toutes les transformations qui affectent les dimensions qui sont laissées de côté dans la généralisation. Ainsi, un concept empirique POMME peut ignorer la couleur précise de la pomme, sa taille, la présence de taches, *et cætera*. La catégorie créée à partir d'un percept moyen de pomme sera invariante pour toutes les transformations affectant ces aspects "anecdotiques".

Le fait que différents apprenants, exposé à des expériences limitées et différentes, convergent vers des bonnes formes n'est pas fortuit. Les mécanismes d'apprentissage de portée générale, comme les mécanismes empiristes ou constructivistes, permettant d'induire des catégories, ont la propriété d'être isotropes : ils sont insensibles à des changements de repère. Or, de tels mécanismes isotropes produisent des bonnes formes, *id est* des formes invariantes pour de nombreuses transformations (DESSALLES 1998 [29]). Dans le domaine conceptuel, il existe certainement des bonnes formes. Pour autant, on ne peut pas considérer que le système conceptuel se résume à un catalogue de bonnes formes. En particulier, les rapports qu'un concept donné entretient avec d'autres concepts ne présentent pas la belle symétrie qu'on est en droit d'attendre d'un système d'induction général. Par exemple, les relations conceptuelles évoquées au cours d'une conversation déterminent un ensemble de situations possibles de très faible symétrie (DESSALLES 1993 [28]). Si le seul mécanisme d'apprentissage disponible était un mécanisme inductif isotrope, empiriste ou constructiviste, la constitution des ensembles de connaissances que nous mobilisons dans les conversations serait inexplicable. Cette situation reproduit, au niveau conceptuel, la difficulté d'expliquer l'acquisition des connaissances syntaxiques par des moyens empiriques ou constructivistes (CHOMSKY 1975 [15])⁷.

⁷ Nous aurons l'occasion de développer cette critique plus tard (CF. ANNEXE).

Le fait de considérer qu'il n'y a pas de différence de nature entre les percepts et les concepts, et que ceux-ci sont identifiables à des percepts moyens, est une idée à la fois très séduisante et très décevante. À la fin de ce travail, nous essaierons de montrer que cette idée et ses avantages peuvent être récupérés, à condition de compléter le dispositif inductif par un mécanisme dynamique de production de symboles. Auparavant, nous allons considérer en détail l'option opposée, selon laquelle les concepts sont d'une autre nature que les percepts.

3.3. Symboles

L'idée de symbole s'oppose par plusieurs aspects à la notion de percept intégré. Un percept intégré ne peut, au mieux, que représenter une catégorie d'objets perçus, alors qu'un symbole est avant tout un élément combinable ayant, par ailleurs, des propriétés représentationnelles.

[...] Nor can categorical representations yet be interpreted as « meaning » anything. It is true that they pick out the class of objects they « name », but the names do not have all the systematic properties of symbols and symbol systems [...]. They are just an inert taxonomy. For systematicity it must be possible to combine and recombine them rulefully into propositions that can be semantically interpreted. « Horse » is so far just an arbitrary response that is reliably made in the presence of a certain category of objects. There is no justification for interpreting it holophrastically as meaning « This is a (member of the category) horse » when produced in the presence of a horse, because of the other expected systematic properties of « this » and « a » and the all-important « is » of predication are not exhibited by mere passive taxonomizing. (HARNAD 1990 [46] p. 343)

La tradition empiriste postule des représentations conceptuelles qui entretiennent des rapports de ressemblance avec les données de la perception. La topologie, voire la métrique de ressemblance, au sein de l'espace des concepts, est la même que dans l'espace des percepts. Dans beaucoup de modèles, ceci est dû au fait qu'il s'agit du même espace. Cette situation contraste avec les autres modèles, dans lesquels ce lien de ressemblance n'existe pas. Lorsque l'on met en avant le caractère symbolique des concepts, que l'on dote les concepts de propriétés formelles pour permettre leur inclusion dans des calculs dictés par les mécanismes grammaticaux ou déductifs, les représentations obtenues perdent nécessairement tout lien de ressemblance et toute relation topologique ou métrique avec leur contenu. Cette perte, qui se constate dans les différents modèles qui ont pu être proposés, est due au fait que les mécanismes généraux qui conservent les relations de distance et de voisinage possèdent la propriété d'isotropie, alors que les mécanismes compositionnels qui engendrent les représentations conceptuelles complexes en sont dépourvus (DESSALLES 1998 [29]).

Si aucune ressemblance n'existe entre les concepts et les percepts, si l'organisation des concepts ne reproduit en rien l'organisation des percepts, alors l'ancrage des concepts doit être le fruit d'un "apprentissage", soit ontogénétique, soit phylogénétique. Or l'établissement de ce lien d'ancrage pose le problème de la connexion. Par quel mécanisme les concepts, qui sont ici des entités symboliques, peuvent-ils recevoir un contenu dans l'espace des percepts ? On comprend aisément comment un mot M peut se retrouver associé à un percept P . Il suffit que M soit prononcé en présence du sujet lorsque celui-ci perçoit P . Si le mot M doit renvoyer en même temps au concept C , il faut déterminer le mécanisme par lequel le triangle $M - P - C$ se retrouve complété. Or C , en tant qu'entité symbolique propre au système conceptuel, est en quelque sorte hors d'atteinte.

Ce problème de la connexion concerne toutes les théories symboliques des concepts. Certes, certaines théories parviennent à dériver l'ancrage des concepts de l'ancrage de leurs composants. Ainsi, si le concept TULIPE est défini, au sein du système conceptuel, à l'aide du concept FLEUR, il hérite une partie de son ancrage de ce dernier. Cette méthode peut, en principe, résoudre le problème de l'ancrage pour la plupart des concepts. Il n'en reste pas moins que la connexion restera inexplicée pour certains concepts, dits primitifs, à partir desquels tous les autres sont construits.

Nous allons considérer successivement deux types d'explication de l'établissement de l'intentionnalité, invoquant respectivement la communication et l'information. Au départ, ces explications ont été produites dans un contexte externaliste, pour expliquer que les concepts se trouvent reliés à des entités du monde. Dans chaque cas, nous examinerons comment ces cadres théoriques peuvent être transposés dans notre propre cadre, en prenant l'explication de l'intentionnalité comme un mécanisme d'ancrage. Même dans l'hypothèse d'un monde objectif, notre expérience reste le seul lien que nous pouvons avoir avec ce monde. L'intentionnalité des concepts, comprise dans ces théories comme le lien qui unit ces concepts au monde, vient donc nécessairement de la composition de deux mécanismes, l'intégration de l'expérience perceptive et l'ancrage.

3.4. Intentionnalité et communication

Certaines approches n'abordent pas la question de l'intentionnalité comme un problème de connexion entre des représentations mentales, mais comme un problème d'association entre des mots et des catégories du monde. Pour de telles approches, qui adoptent une perspective externaliste, il est essentiel d'expliquer comment le langage parvient à décrire le monde de manière "correcte". Comment un individu donné parvient-il à connaître l'ensemble des liens mot - référence, même pour des entités du monde qu'il n'a jamais rencontrées ? La réponse suppose que le lien causal entre le mot et sa référence a généralement une origine qui dépasse l'histoire de l'individu. Si je désigne un objet précis de mon champ visuel en prononçant le mot *arbre* et que mon interlocuteur saisit mon intention, c'est parce que mon interlocuteur a appris ce mot quand il était enfant et que je l'ai moi-même appris en cours de français quand mon professeur désignait une image dans le livre de cours. Pour que je puisse communiquer le contenu du mot *arbre*, ce contenu a dû m'être communiqué. Il a été aussi communiqué à mon professeur, et ainsi de suite le long d'une chaîne d'usage jusqu'à la première fois où un individu a glosé l'objet par le mot. La communication assure le lien entre le mot et sa référence et le propage.

Le cas typique, qui sert de modèle à cette approche, est celui des noms propres. Les noms propres sont le résultat d'un acte de baptême. En donnant un nom à un individu, la communauté décide de le désigner par une étiquette qui glosera désormais l'enfant pour ceux qui le connaissent. Le nom constitue alors le moyen le plus simple pour pouvoir communiquer verbalement à propos de cet individu. Grâce à cette étiquette conventionnelle, on peut même parler de l'individu avec quelqu'un qui ne le connaît pas. La personne sait que l'étiquette ne fait que remplacer sa référence.

Someone, let's say a baby, is born; his parents call him by a certain name. They talk about him to their friends. Other people meet him. Through various sorts of talk the name is spread from link to link as if by a chain. A speaker who is on the far end of this chain, who has heard about, say Richard Feynman, in the market place or elsewhere, may be referring to Richard Feynman even though he can't remember from whom he first heard about Feynman or from whom he ever heard of Feynman. (KRIPKE 1972 [62] p. 91)

L'élargissement de la métaphore de baptême initial aux noms communs semble assez intuitif quand il s'agit des espèces naturelles ou des artefacts avec lesquelles les individus sont en interaction directe. On peut imaginer que l'on apprend des mots comme chat ou voiture quand ils sont utilisés pour désigner des objets de la vie quotidienne. Grâce à la chaîne communicationnelle, il n'est même pas obligatoire que l'individu ait eu une expérience directe des objets désignés, si bien que la communication par ostension n'est pas nécessaire pour l'apprentissage individuel.

[...] the species-name may be passed from link to link, exactly as in the case of proper names, so that many who have seen little or no gold can still use the term. Their reference is determined by a causal (historical) chain [...]. (KRIPKE 1972 [62] p. 139)

Si les individus ne peuvent pas profiter de l'ostension pour établir le lien mot - référence, il existe un autre moyen, la description, par lequel ils accèdent au moyen d'utiliser le mot dans les circonstances adéquates. Dans le cas des entités abstraites ou imaginaires désignées par des mots comme bourse ou Pégase, le baptême initial est plus difficile à concevoir. Il suffit dans ce cas d'imaginer que la description sert non seulement à contraindre le lien mot - référence pour tous les individus de la chaîne de communication, mais qu'elle sert également à effectuer le "baptême" de l'entité pour la communauté (KRIPKE 1972 [62]). Ainsi, la première description qui a été faite de Pégase en tant que cheval ailé a permis d'établir un lien mot - référence qui s'est ensuite propagée d'individu en individu, si bien que le mot Pégase, sans être équivalent à la description qui lui est associée, possède une référence pour tous les individus de la communauté.

Cette approche de la notion de contenu repose sur deux considérations de base. La première est liée au caractère indexical d'une grande partie des expressions langagières (PUTNAM 1975 [87]). Selon cette idée, la fonction principale du mot arbre est de désigner sa référence. La seconde est d'ordre social : le fait que l'on utilise le terme arbre et non un autre terme pour désigner sa référence est le résultat d'un accord dans la communauté de communication (PUTNAM 1975 [87]).

La question qui nous importe, dans ce travail, est de savoir comment une telle théorie du contenu peut s'interpréter dans une approche cognitive représentationnaliste du langage naturel. Notre but n'est pas de savoir comment les individus parviennent à employer les mots d'une manière objectivement correcte, mais plutôt de rendre compte de la capacité des individus à comprendre les énoncés langagiers de manière à effectuer certaines inférences et à produire d'autres énoncés appropriés. Dans un cadre représentationnaliste, le mot, qui n'est autre qu'une forme phonologique propre à la langue parlée dans la communauté, peut évoquer une représentation mentale chez les locuteurs de cette langue. C'est cette représentation mentale, le concept, qui déclenche les processus compositionnels et inférentiels. Pour que cela soit possible, il faut que le concept soit lié, à la suite d'un apprentissage, au mot qui l'évoque. Si l'on accepte que le concept est lié à un certain type de perception, par exemple la perception d'un lien causal entre un phénomène et ses effets, alors il suffit que le mot soit associé une fois à cette situation causale pour qu'il se retrouve associé au concept pour tous les individus de la communauté linguistique.

They could all use the term “electricity” [...] what they do have in common is this: that each of them is connected by a certain kind of causal chain to a situation in which a *description* of electricity is given, and generally a *causal* description – that is, one which singles out electricity as the physical magnitude *responsible* for certain effects in a certain way. [...] Let us call this event – my acquiring the ability to use the term “electricity” in this way – an *introducing event*. It is clear that each of my later uses will be causally connected to this introducing event, as long as those uses exemplify the ability I required in that introducing event. (PUTNAM 1975 [87] p. 199)

De cette manière, le lien mot - concept peut s'expliquer, au même titre que le lien mot - objet considéré par ces auteurs, par une chaîne historique causale qui associe, à chacun de ses maillons, une représentation phonologique et une représentation conceptuelle. Le point essentiel que l'on peut retenir de cette approche historique et conventionnelle de la signification est que le fait que les mots aient un sens suppose que les individus interagissent de manière causale avec leur environnement et interagissent de manière causale entre eux, par le biais de la communication.

Cette manière de concevoir la création de nouvelles significations et leur propagation laisse une part minimale à l'apprentissage et aux difficultés qu'il présente. Le problème est de rendre plausible le fait que deux individus finissent par accorder la même signification au même mot. Même dans le cas des noms propres, la réponse ne va pas de soi. Lors de l'événement baptismal, les personnes présentes ont l'occasion d'associer le nom non pas à l'enfant lui-même, en tant qu'entité objective, mais à une perception qu'ils en ont. Il faut qu'un mécanisme leur permette de généraliser le nom à l'individu quel que soit son âge, leur évite d'associer le nom à l'enfant seulement lorsqu'il est dans son landau, *et cætera*. La situation est encore plus problématique dans le cas des espèces naturelles. À partir du moment où l'on renonce à la simplification qui consiste à penser que les entités du monde nous apparaissent directement, détachées de toutes les autres, l'établissement des associations mots - concepts, médiatisées par les associations mots - percepts, devient moins évident. Il s'agit d'expliquer ce qui nous permet de réemployer le même mot chat dans des situations différentes dans lesquelles le contexte d'apparition de l'animal peut varier du tout au tout. Il faut postuler l'existence de capacités inductives puissantes qui permettent à deux individus, ou à un même individu dans deux situations différentes, de repérer des éléments communs qui provoquent l'utilisation appropriée d'un même mot. L'explication du lien historique causal de dénomination y perd la simplicité qui faisait son attrait. S'il existe un mécanisme psychologique complexe par lequel les individus parviennent à déterminer la référence d'un mot en contexte, le fait que ce mot soit prononcé en leur présence dans des circonstances appropriées, bien que nécessaire, ne joue qu'un rôle marginal dans l'établissement de la signification.

Il est également difficile de se contenter d'invoquer l'histoire des interactions lorsqu'il s'agit d'expliquer l'établissement du sens des mots, notamment les mots abstraits, dont la signification est obtenue par des descriptions utilisant d'autres mots. L'interprétation des descriptions suppose là encore des mécanismes riches, laissant à l'origine historique des associations mot - objet une part anecdotique. L'importance des mécanismes cognitifs est encore plus flagrante lorsque la description sert à désigner une situation dont l'individu est censé inférer le concept décrit. Ainsi, une description comme ce qui surnage lorsqu'on fait cuire du lait suppose des capacités à constituer un concept à partir d'une situation imaginée ou remémorée. Tout sujet qui apprend un mot nouveau pour lui est confronté à un problème de même difficulté que celui qu'ont connu ses ancêtres lors de la création du lien mot - objet dans la communauté.

La question de l'émergence de significations dans une communauté a été explorée expérimentalement à l'aide de simulations (KAPLAN 2001 [57]). Il apparaît que l'acte initial de dénomination d'un objet perçu dans l'environnement par un agent ne suffit pas à l'acquisition du concept par les autres agents. Il faut que ceux-ci aient une perception comparable, puis affinent leurs catégories lors des utilisations ultérieures du même mot. Les agents parviennent ainsi à former des catégories conceptuelles pertinentes en affinant des divisions selon une procédure dichotomique, par exemple sur la couleur ou la forme des objets, de manière à discriminer les objets présents. Il peut en résulter des dialogues de sourds : tel objet sera signalé par le mot *gorewa* par l'agent 1, ce qui, pour lui, signifie "rouge" ; l'agent 2, entendant *gorewa*, comprendra par exemple "triangle". La communication peut réussir si le seul objet rouge se trouve être triangulaire. Heureusement, confrontés à d'autres situations, l'un des deux agents au moins sera amené à réviser son lexique.

Ce type d'expérience montre que l'histoire de la communication entre agents, qui ressemble plus à un réseau complexe d'interactions qu'à une simple chaîne de communication, se révèle déterminante pour l'adoption d'un mot plutôt qu'un autre. L'histoire des interactions en contexte permet donc, en principe, d'expliquer l'émergence d'un ensemble donné de concepts dans une communauté. En revanche, il serait illusoire de vouloir fonder une théorie des concepts en se limitant à l'historique des interactions. Les concepts que forment les robots préexistent en puissance dans leur capacité à affiner leurs catégories et à les associer avec des mots. L'observateur pourrait être tenté de penser que la signification des mots qu'ils emploient se situe dans le monde simplifié qui s'offre à leur caméra. Ainsi, le mot *gorewa* semble désigner le triangle rouge sur le tableau. Pourtant, ces robots diffèrent suffisamment de nous pour que nous soyons obligés de reconnaître qu'il n'en est rien. La signification de leurs mots se situe dans leur mémoire, dans les structures qu'ils ont affinées pour répondre aux besoins de la communication qui joue davantage le rôle de déclencheur des mécanismes internes qu'un rôle structurant. Ces robots, pas plus que les humains, ne sont télépathes. Ils ne peuvent pas se transmettre leurs catégories conceptuelles. Ils ne peuvent qu'échanger des mots en situation et employer leurs mécanismes internes pour tenter de parvenir à des catégories compatibles avec leur perception.

Il est indéniable que l'association entre les mots et les perceptions ne doit rien, la plupart du temps, à la forme phonologique des premiers. Les associations mots - percepts ne sont donc possibles que si certains individus les créent et les propagent. Cependant, comme les situations ne se présentent jamais deux fois sous le même aspect, le lien à expliquer est celui qui s'instaure entre le mot et le concept, ce dernier permettant de discriminer les percepts. L'étude des interactions entre individus ne permet pas de faire l'économie d'un mécanisme inductif par lequel ces individus parviennent à forger des concepts compatibles avec l'usage des mots par leurs partenaires.

3.5. Intentionnalité et information

L'histoire des interactions langagières ne peut suffire à rendre compte de l'intentionnalité des concepts : elle ne peut qu'expliquer pourquoi ce sont certains percepts qui sont nommés dans la communauté plutôt que d'autres. Si l'on s'intéresse au résultat de ce processus historique pour l'individu, force est de constater que celui-ci a acquis une capacité lui permettant d'employer les mots à bon escient. Une manière minimale de décrire cette capacité est de la présenter comme une disposition comportementale.

[...] a stimulation σ belongs to the affirmative stimulus meaning of a sentence S for a given speaker if and only if there is a stimulation σ' such that if the speaker were given σ' , then were asked S, then were given σ , then were asked S again, he would dissent the first time and assent the second. [...] What now of the strong conditional, the "would" in our definition of stimulus meaning? [...] what the conditional defines is a disposition, in this case a disposition to assent to or to dissent from S when variously stimulated.

(QUINE 1960 [90] p. 32)

La présence de ces dispositions chez les individus est le résultat d'un apprentissage. Dans la doctrine behavioriste, le mécanisme d'apprentissage par lequel on acquiert de nouvelles significations ne peut être qu'un renforcement positif ou négatif.

It remains clear in any event that the child's early learning of a verbal response depends on society's reinforcement of the response in association with the stimulations that merit the response, from the society's point of view, and society's discouragement of it otherwise.

(QUINE 1960 [90] p. 82)

Pour ce genre d'approche, le fait de posséder des concepts n'est rien d'autre qu'une disposition comportementale conditionnée. On ne peut donc pas parler de contenu pour les concepts. Toute notion même de sémantique est exclue d'un tel schéma (QUINE 1960 [90]). Si l'on est représentationnaliste et que l'on fait découler les dispositions comportementales langagières de la possession de représentations conceptuelles, le seul point à retenir d'une telle théorie est le lien statistique, fruit d'un renforcement répété, qu'elle postule entre les concepts et les percepts.

La reconnaissance de ce lien statistique empêche de ne voir dans les concepts que le simple résultat d'un acte de désignation propagé à travers l'histoire des usages. Selon cette hypothèse, les sujets sont capables d'identifier l'objet qu'on leur désigne et de le reconnaître dans les situations ultérieures parce qu'ils ont forgé un concept qui leur permet de le discriminer (DRETSKE 1981 [31]). Cette fonction de discrimination justifie, d'un point de vue théorique, l'existence de concepts dotés d'un contenu. Le mot désigne un objet, mais cet objet n'existe perceptivement que parce qu'un concept le discrimine. Dans ce genre d'approche où, rappelons-le, les concepts ne sont pas de même nature que les percepts, le pouvoir de discrimination des concepts est attribué au fait que le concept co-varie avec l'objet perçu (DRETSKE 1981 [31]). Cette co-variation a la vertu d'expliquer également l'acquisition de ce pouvoir de discrimination.

In teaching someone the concept *red*, we show the pupil variously colored objects at reasonably close range and under normal illumination. That is, we exhibit the colored objects under conditions in which *information* about their color is transmitted, received and (hopefully) perceptually encoded. [...] If the subject is to acquire the concept *red*, he or she must not only be shown red things (and, presumably, non-red things), they must be allowed to receive the information that they are red (and not red). The reason for this should be clear: it is the information that the object is red that is needed to shape the internal structure that will eventually qualify as the subject's concept *red*. We need information to manufacture meaning (the concept) because information is required to crystallize a type of structure with the appropriate semantic content.

(DRETSKE 1981 [31] p. 194)

Non seulement la co-variation entre le concept et l'objet permet au sujet de former le concept puis, grâce à ce concept, de discriminer les occurrences futures de l'objet, mais elle explique également le fait que le contenu d'un concept reste stable au cours du temps dans la communauté. Ainsi, dans ce type de théories, le lien causal qui unit le concept à son contenu

est de nature statistique, attribué à une co-variation repérée par les sujets entre certaines de leurs perceptions et certains de leurs états mentaux.

Le principe de co-variation remet en cause l'idée de contenu purement conventionnel. Une signification qui serait un pur produit d'une convention sociale, propagée d'individu en individu, pourrait être totalement arbitraire. Le lien qui unit le concept à son contenu, dans la mesure où il repose sur une co-variation statistique, n'est pas arbitraire. Il prend un caractère naturel. Le concept, parce qu'il co-varie avec les objets perçus, devient porteur d'information. Pour un externaliste, les conséquences semblent impressionnantes. Les concepts, par leur intentionnalité, apportent une information sur le monde. On rejoint ainsi l'idée selon laquelle le langage nous renseigne sur la réalité qui nous entoure. Malheureusement, la connaissance que l'on peut espérer avoir du monde en étudiant les concepts est limitée en raison du manque de robustesse d'un lien intentionnel de nature statistique. Dans la mesure où l'erreur est possible, où la perception d'un chat dans l'obscurité peut conduire à tort à l'évocation du concept CHIEN, l'activation d'un concept n'offre aucune certitude sur l'état du monde. Si de telles erreurs sont fréquentes, on peut arriver à ce que le concept soit évoqué dans des conditions considérées comme impropres du point de vue externaliste : le concept CHIEN co-variera avec un ensemble hétérogène d'entités et non avec la catégorie "chien", considérée comme objective. Il semble n'y avoir aucune garantie théorique sur ce que l'évocation d'un concept peut signifier sur l'état du monde. Cette situation, qui est sans conséquences si l'on se place dans le cadre de l'internalisme méthodologique, semble inacceptable pour les tenants de l'externalisme.

La question de la validité du lien intentionnel, considéré comme informationnel, a conduit bon nombre d'auteurs à rechercher une distinction formelle entre l'évocation "normales" du concept et les conditions d'évocations considérées comme erronées. Ainsi, ce n'est pas parce que, dans des cas de perception anormale, le concept CHIEN viendra à être évoqué que ce concept recouvrera la signification "chien ou grand chat dans le noir". Même si le concept CHIEN co-varie avec des situations du type "grand chat dans le noir", cette co-variation peut être éliminée du lien intentionnel dès lors qu'on est capable de filtrer les situations anormales d'évocation. La question évidente est alors de savoir selon quel critère on peut distinguer les conditions normales des conditions anormales d'évocation. On peut considérer qu'il existe une asymétrie fondamentale qui alloue une prééminence aux conditions normales.

La nature de cette asymétrie est difficile à expliciter, en dehors du constat *a posteriori* que, comme les concepts sont porteurs d'information, il faut bien que les conditions normales d'évocation l'emportent sur les conditions anormales. On peut imaginer, par exemple, en raison de la nature statistique du lien d'évocation, que les conditions anormales sont plus rares. Cette supposition peut reposer sur une présomption d'efficacité du système conceptuel qui est capable de filtrer, dans la plupart des cas, les conditions adéquates d'évocation des concepts. Les concepts sont supposés avoir une fonction cognitive propre. On peut donc s'appuyer sur cette fonction cognitive pour affirmer que, dans des conditions normales de fonctionnement, les concepts ont un lien intentionnel non ambigu. Certains auteurs justifient cette idée de fonctionnement normal non seulement sur le plan cognitif, mais également dans une perspective phylogénétique (MILLIKAN 1984 [73]). Ils en appellent à la sélection naturelle pour fonder l'idée de fonction propre pour un concept.

Si nos concepts existent, c'est parce qu'ils sont efficaces, et ils sont efficaces parce que nous descendons d'individus qui ont été sélectionnés pour les avoir possédés, ou tout du moins pour avoir possédé les moyens de les acquérir en tant que concepts efficaces. Le contenu de tels concepts, forgés par la sélection naturelle, est celui pour lequel ils ont été sélectionnés dans la phylogenèse ou dans l'ontogenèse, non les configurations qui peuvent par moments les évoquer de manière erronée. Il semble que l'on récupère, par le biais de cette

fonction propre, une notion de contenu objectif pour les concepts. Un exemple souvent cité est celui de la perception d'une mouche par la grenouille. L'état neuronal provoqué par la vue de la mouche déclenche une réaction de prédation. Dans certaines conditions, le même comportement de prédation peut être provoqué par un leurre ou par un petit objet inerte emporté par le vent. Si l'on suppose que la réaction de prédation est adaptée pour la fonction biologique de prise d'aliment, on est en droit de dire que la mouche est la cause normale de la réaction de prédation et que le leurre en est une cause anormale. Dans ces conditions, la réaction de prédation peut être considérée comme l'indice de la présence d'une mouche, même dans le cas d'une erreur.

L'intérêt de ce type de raisonnement téléologique pour la compréhension du système conceptuel humain doit être relativisé. Notons tout d'abord que l'appel à la sélection naturelle pour justifier l'efficacité des concepts est critiquable. On peut l'accepter lorsqu'il s'agit d'une caractéristique biologique, quoique dans certaines limites (GOULD & LEWONTIN 1979 [45]). Par exemple, on peut imaginer qu'il existe un petit nombre de concepts primitifs innés qui sont de bons candidats pour être le fruit d'une évolution biologique. Or, cette éventualité est loin d'aller de soi, en particulier parce que nous parlons ici de l'innéité de concepts et non de l'innéité de sensations, comme nous aurons l'occasion de le souligner plus loin (CF. CHAPITRE 8). Lorsqu'il s'agit d'expliquer la genèse de l'ensemble des concepts, considérés comme appris au cours de l'ontogenèse, nous n'avons plus affaire à des caractéristiques biologiques à proprement parler, et l'argument de la sélection naturelle ne porte pas. Il faut donc se rabattre sur l'idée métaphorique d'une sélection ontogénétique analogue à la sélection darwinienne qui se ramène aux mécanismes d'apprentissage qui ont été proposés.

Par ailleurs, si l'appel à la sélection naturelle, dans le cas de l'exemple de la grenouille, apparaît comme justifié, dans le cas des concepts est moins évident. La fonction propre du comportement de prédation observé chez la grenouille ne fait pas mystère. Comment définir une fonction propre pour le concept CHIEN ? Si l'on dit qu'il s'agit de discriminer les chiens et que le concept CHIEN est biologiquement efficace pour cette fonction et aucune autre, le raisonnement est circulaire. On définit la fonction propre du concept CHIEN à partir de la catégorie "chiens", considérée comme une donnée objective, alors que l'objectif était, à l'inverse, d'apprendre quelque chose sur le monde à partir du concept. La seule chose qui peut être dite est donc que le concept doit avoir un contenu dans le monde, qui correspond à ce qu'il discrimine dans les conditions normales, sans qu'il soit possible d'en dire davantage puisque nous sommes dans l'impossibilité de définir ce qu'est le fonctionnement normal d'un concept donné. C'est d'ailleurs cette impossibilité fondamentale de faire une différence *a priori* entre discrimination réussie et discrimination erronée qui rend le projet externaliste si difficile à transposer à la modélisation cognitive.

Une tentative pour restaurer la cohérence de la vision externaliste est de renoncer à la fonder sur le seul critère informationnel. Si l'on reconnaît l'impossibilité, dans la plupart des cas, de caractériser de manière objective une différence entre les causes idéales d'évocation du concept et les causes erronées, il est préférable de poser que le contenu d'un concept est constitué par ce qui est commun à toutes les situations qui le déclenchent (FODOR 1990 [35]). De cette manière, le concept véhicule de l'information sur toutes ces causes. La conséquence, pour un externaliste, est que l'information portée par le concept ne suffit pas, à elle seule, pour définir son contenu. Si le concept CHIEN est déclenché dans certaines conditions par un chat, alors le chat dans ces conditions fait partie de l'ensemble des situations avec lesquelles le concept CHIEN co-varie, sans faire partie de son contenu. Que peut-on dire, alors, sur le contenu externe du concept ? Écrivons que le concept C peut être causé par une propriété P du monde dans les conditions s : $C = f(a_s(P))$. La fonction $a_s(P)$ désigne par exemple l'apparence de P dans les conditions s . Le même concept peut être causé par une propriété P' dans les

conditions $s' : C = f(a_{s'}(P'))$). Si la symétrie entre ces deux formules est parfaite, on ne peut espérer distinguer les conditions normales, représentées ici par s , des conditions anormales désignées ici par s' . Cependant, si l'on peut écrire que $a_{s'}(P) = a_{s'}(P')$ sans que $a_s(P) = a_s(P')$, alors nous tenons une dissymétrie qui permet de réhabiliter la notion de contenu P pour le concept C . Ainsi, nous pouvons affirmer $C = f(a_{s'}(P)) = f(a_{s'}(P'))$, mais $C = f(a_s(P)) \neq f(a_s(P')) : P$ et P' deviennent discernables, en théorie, par rapport à C . L'écriture $a_{s'}(P) = a_{s'}(P')$ peut signifier que P et P' ont la même apparence lorsque la perception est sujette à un certain niveau d'imprécision s' , ou qu'elles évoquent la même image si on tolère un niveau d'association s' . Ainsi, un chat peut passer pour un chien dans le noir, et le lait peut faire penser à une vache si on se laisse aller à faire des associations. L'asymétrie entre P et P' vis-à-vis de C permet de penser que l'une constitue le contenu de C et l'autre non. Si C porte de l'information sur P' , c'est seulement parce que C porte de l'information sur P .

Cows cause “cow” tokens, and (let's suppose) cats cause “cow” tokens. But “cow” means cow and not cat or cow or cat because *there being cat-caused “cow” tokens depends on there being cow-caused “cow” tokens, but not the other way around.* “Cow” means cow because, [...] non-cow-caused “cow” tokens are *asymmetrically dependent upon cow-caused “cow” tokens.* “Cow” means cow because *but that “cow” tokens carry information about cows, they wouldn't carry information about anything.* (FODOR 1990 [35] p. 91)

Un externaliste, selon cette théorie, peut prolonger dans le monde le caractère asymétrique des évocations pour en conclure que certaines situations, parmi celles qui évoquent le concept, en constituent le contenu, alors que d'autres, qui l'évoquent tout autant, en sont exclues. Cette asymétrie ne doit rien au critère informationnel, qu'il soit phylogénétique ou ontogénétique, puisque toutes ces situations co-varient de la même manière avec le concept.

Dans le cadre de l'internalisme méthodologique au sein duquel nous nous situons, ce débat concernant le caractère robuste du contenu conceptuel perd beaucoup de son enjeu. La question devient de savoir si les concepts, considérés ici comme des états mentaux distincts de la perception, renvoient de manière non ambiguë à certains percepts, autrement dit aux percepts qu'ils sont censés normalement discriminer. L'argument téléologique revient alors à fonder cette normalité sur le fait que chez des êtres qui sont le produit de la sélection naturelle, le fonctionnement normal est celui qui a été sélectionné. Comme dans le cas externaliste, cela ne nous renseigne pas beaucoup sur le contenu du concept, car en général nous n'avons pas de moyen indépendant d'estimer l'incidence de telle ou telle discrimination sur la survie des individus qui en sont capables. L'argument de la dépendance asymétrique, quant à lui, se limite à postuler une distinction direct/indirect dans les conditions d'évocation des concepts, sans qu'il soit possible d'en dire plus sans une information indépendante sur les catégories que le concept est censé discriminer.

Notons que pour l'internalisme méthodologique, le lien entre les percepts et les concepts est nécessairement sujet à l'erreur. On ne peut pas appliquer l'espace continu des percepts sur l'espace symbolique des concepts sans qu'il se produise des distorsions. On rejoint ainsi les notions de détection en présence de bruit développées dans la théorie mathématique de l'information. La qualité de la discrimination d'un percept par un concept peut s'estimer par la mesure d'une certaine probabilité d'erreur. On peut même estimer la qualité d'un ensemble de concepts comme sa capacité à produire des discriminations qui minimise la distorsion des distances entre percepts.

Avec cette manière de concevoir l'opération de discrimination conceptuelle, l'internalisme méthodologique permet d'abandonner sans dommage l'objectif qui consiste à vouloir établir une “bijection” entre les concepts et un ensemble prédéfini de contenus

possibles. Comme suggéré plus haut, on peut définir le contenu d'un concept comme ce qui est commun à toutes les situations qui le déclenchent. De cette façon, on conserve l'idée de lien informationnel entre le concept et son contenu, puisque la co-variation est systématique. Cette définition a également le mérite de supprimer l'essentiel de ce qui constituait des "erreurs" dans le cas bijectif. Cet avantage est dû au fait que, comme dans le cas empiriste où les concepts étaient constitués de percepts généralisés, le lien intentionnel permet de respecter la gradualité inhérente à la perception. Lorsque l'on cherche une bijection entre les concepts et les catégories du monde, il faut être capable de donner avec précision la limite de la catégorie "chiens" hors de laquelle le concept CHIEN cesse d'être activé. Une telle gageure n'est pas exigée si le concept, bien qu'étant considéré comme une entité symbolique distincte des percepts, renvoie à un contenu correspondant à un percept moyen.

Un autre avantage de l'abandon du lien bijectif au profit de concepts à contenu moyen est de permettre d'envisager plus simplement un mécanisme d'apprentissage. Dans le cas externaliste, l'origine de l'adéquation parfaite entre les entités mentales que sont les concepts et les catégories du monde reste inexpliquée. Comme nous l'avons vu, l'appel à la sélection naturelle ne remplit pas le vide théorique dû à l'absence de critère indépendant pour décider ce qui, dans le monde, mérite, d'un point de vue biologique, d'être isolé comme le contenu d'un concept. La conséquence est un raisonnement circulaire dans lequel ce sont précisément les concepts que nous formons qui sont supposés favoriser la survie. Si, dans le cadre de l'internalisme méthodologique, nous considérons que le contenu des concepts se trouve dans le domaine des expériences, alors l'établissement de ce contenu au cours de l'ontogenèse semble aisé à expliquer. Il suffit d'invoquer un mécanisme statistique d'extraction de régularités. Cette facilité n'est malheureusement qu'apparente, comme nous allons le constater maintenant.

3.6. Le paradoxe de la connexion

Les différentes théories du système conceptuel humain ne sont pas égales devant le problème de la connexion du concept à son contenu. Pour les empiristes, le problème n'existe pas, car les concepts sont de la même nature que les percepts et le lien intentionnel est une identité. Les mécanismes de généralisation ou de construction de bonnes formes produisent des concepts qui, restant dans l'espace des percepts, ne cessent jamais d'être identiques à leur contenu.

Le problème de la connexion se pose de manière extrême, au point de conduire à un paradoxe, dès que les concepts sont considérés comme étant d'une autre nature que les percepts.

[...] there is a further constraint that whatever theory of concepts we settle on should satisfy: it must explain why there is so generally a content relation between the experience that eventuates in concept attainment and the concept that the experience eventuates in attaining. (FODOR 1998 [37] p. 132)

La différence de nature entre les représentations perceptuelles et les représentations conceptuelles rend le problème particulièrement délicat. Les concepts, nous l'avons dit, sont supposés être des représentations discrètes et digitales. Ils n'offrent pas la gradualité qui existe dans l'espace des percepts. Ils ne représentent pas non plus les relations de voisinage qui existent dans le domaine des percepts. Le mécanisme qui les connecte aux percepts ne saurait résulter d'un simple processus d'échantillonnage. Le lien qui semble exister entre

certaines concepts et des prototypes perceptuels, loin d'être une évidence, est précisément ce qui requiert une explication.

[...] what doorknobs have in common qua doorknobs is being the kind of thing that our kind of minds (do or would) lock to from experience with instances of the doorknob stereotype. [...] We have the kinds of minds that often acquire the concept *X* from experiences whose intentional objects are properties belonging to the *X*-stereotype. [...] *Stereotype* is a statistical notion. The only theoretically interesting connection between being a doorknob and satisfying the doorknob stereotype is that, contingently, things that do either often do both.
(FODOR 1998 [37] p. 137)

Le fait de modéliser un lien intentionnel entre deux espaces distincts est déjà problématique en soi, car l'espace conceptuel n'est pas accessible en dehors de ce lien intentionnel. Certes, dans certains modèles, bon nombre de concepts peuvent être accessibles à partir d'autres concepts. Ainsi, si l'on accepte que le concept ÉTALON peut être décrit par une expression comme CHEVAL+MÂLE+ENTIER, on peut espérer définir un mécanisme par lequel le concept ÉTALON hérite son lien intentionnel de ceux des concepts CHEVAL, MÂLE et ENTIER. Toutefois, même dans un schéma définitionnel de ce type, il existe toujours des concepts primitifs qui ne font l'objet d'aucune définition. Comment leur lien intentionnel s'établit-il ? L'expérience produit des percepts, l'exposition au langage permet d'isoler des mots, mais rien, dans l'ontogenèse, ne vient de manière évidente connecter les concepts aux percepts et aux mots qui sont censés les évoquer.

Une solution communément considérée consiste à invoquer l'apprentissage inductif. Ainsi, l'enfant aurait l'occasion de cerner, en généralisant après avoir été exposé à une ou plusieurs occurrences du concept, l'étendue des expériences que le concept est censé reconnaître. Cette solution est incorrecte. Elle présuppose ce qu'elle est censée montrer. Pour reconnaître la première occurrence d'un nouveau concept *P*, l'enfant doit déjà posséder *P* ! Au mieux, le système ne peut que combiner des concepts déjà connus pour produire une formule *G* qui sera équivalente à *P*. Mais un système capable de réaliser une telle opération possède déjà *P* en puissance.

But notice that learning that [*Px is true iff x is G*] could be learning *P* (learning what *P* means) only for an organism that already understands *G*. For, and this point is critical, *G* in [the preceding] formula is *used*, not mentioned. [...] an organism can learn *P* only if it is already able to use at least one predicate that is co-extensive with *P*, viz., *G*.
(FODOR 1975 [34] p. 80)

De fait, les systèmes d'apprentissage proposés en intelligence artificielle, dans la mesure où ils produisent des représentations conceptuelles distinctes des représentations perceptuelles, fonctionnent sur le modèle de l'appariement combinatoire (CF. ANNEXE). L'exposition à des données permet de sélectionner des combinaisons conceptuelles, pas de former des concepts *de novo*. De plus, ces systèmes s'appuient tous sur des concepts primitifs dont le lien avec les données est fourni *a priori*. Il semble donc que le lien intentionnel repose, *in fine*, sur une base innée (FODOR 1975 [34]).

Le problème de la connexion des concepts à leur contenu peut-il ainsi se résoudre par l'hypothèse selon laquelle certains concepts, dits primitifs, disposeraient du privilège d'une intentionnalité innée ? Ce n'est pas certain. Il ne faut pas confondre un concept comme ROUGE, qui est souvent cité pour faire partie du cercle fermé des concepts primitifs, avec la sensation de rouge. Affirmer que le concept ROUGE possède une intentionnalité innée signifie que l'évocation de ce concept est liée, de manière non ambiguë et inamovible, à celle de sa perception. Il devrait être possible de déceler, par des tests psychologiques, la frontière

universelle qui est censée exister entre l'extension perceptuelle d'un tel concept et les percepts qui évoquent d'autres concepts ou qui n'évoquent aucun concept. Rien de tel n'a jamais été montré, et pour aucun concept. Si les sensations connaissent une certaine universalité (CLARK 1993 [20]), il n'existe pas de constante sur les segmentations auxquelles elles donnent lieu (PALMER 1999 [80]). Par ailleurs, invoquer l'innéité des concepts revient à renvoyer la charge d'établissement du lien intentionnel de l'apprentissage à la phylogenèse. Or la sélection naturelle produit des formes localement nécessaires (MONOD 1970 [75]). Il faudrait donc être en mesure d'expliquer que les concepts innés, de par leurs propriétés dans l'espace des concepts, possèdent un lien non arbitraire avec leur contenu, ce qui est loin d'aller de soi (CF. CHAPITRE 8).

Une manière d'échapper aux difficultés liées à l'innéité des concepts primitifs consiste à accepter cette innéité en refusant celle de leur lien intentionnel. L'apprentissage d'un concept pourrait ainsi revenir à une sorte de recrutement d'un concept préexistant pour répondre à un ensemble défini d'expériences. De cette manière, l'important est l'existence d'un mécanisme perceptuel capable de grouper certains percepts, par exemple sous la forme d'un percept moyen, et d'un autre mécanisme capable de connecter ce groupe de perceptions à une représentation de type conceptuel. Ainsi, le concept ROUGE ne se trouverait lié à la perception du rouge que de manière non nécessaire, par le biais d'un mécanisme non inductif de recrutement. Notons que cette solution, si elle permet de résoudre une partie des problèmes évoqués ici, n'en résout pas d'autres, au premier rang desquels celui des liens inter-conceptuels. Nous aurons l'occasion de revenir sur cette question fondamentale de la connexion des concepts à leur contenu (CF. CHAPITRE 8). La solution que nous proposerons consistera à renoncer à l'existence des concepts en tant que représentations mentales permanentes (CF. CHAPITRE 9).

Conclusion

La question fondamentale de l'ancrage des concepts concerne la nature du lien entre les concepts et le monde perçu. Cette question concerne également les théories qui s'intéressent au lien direct que les concepts peuvent avoir avec un monde simplifié considéré comme objectif, puisque la seule connexion qu'un système cognitif peut avoir avec un tel monde s'effectue par le biais de la perception.

La question de l'ancrage se révèle soit triviale, soit extraordinairement ardue. Elle est triviale dans les théories qui ne font pas de différence entre concepts et percepts. Dans ce cas, les concepts sont ancrés de manière naturelle dans la perception. En revanche, dès que les concepts sont considérés comme des représentations distinctes de celles que fournissent nos sens, il devient très difficile d'expliquer comment les concepts se retrouvent liés aux "bons" percepts. Certains auteurs, comme nous l'avons vu, supposent que ce lien est établi par le biais de la communication au sein d'une communauté linguistique, ou du fait d'une disposition cognitive qui nous prédispose à élaborer des représentations conceptuelles portant une information sur le monde. Il n'en reste pas moins que l'amorçage de tels processus reste, dans chaque cas, mystérieux. Comme nous l'avons vu, les "solutions" conventionnelles et informationnelles sont loin de résoudre la question de l'établissement du lien intentionnel. Nous aurons l'occasion de revenir sur cette question fondamentale de l'ancrage (CF. CHAPITRE 8).

Chapitre 4 :
Rôle inférentiel des concepts

Introduction

Pour de nombreux auteurs, les représentations conceptuelles sont avant tout le support des inférences. Par rapport aux questions abordées dans le chapitre précédent, il s'agit d'un changement de problématique. Dans une perspective behavioriste, un concept n'a d'autre caractérisation que les comportements observables qui lui sont associés : catégoriser ou raisonner d'une certaine manière dans un certain contexte. Dans une perspective représentationnaliste, la question est de savoir de quelles manières les représentations conceptuelles peuvent avoir un rôle épistémique, comme celui de provoquer des inférences. Dans quelle mesure les concepts, de par leur forme, sont-ils responsables de ces inférences ?

Les pensées évoquées par la perception ou la communication langagière peuvent causer l'évocation d'autres pensées, de manière plus ou moins systématique. Cet enchaînement entre pensées opère sur un médium qui est supposé être celui des concepts. Certaines configurations conceptuelles, par exemple l'idée I_1 que "le chien a mordu le garçon", entraînent d'autres configurations conceptuelles, par exemple l'idée I_2 que "le garçon a mal". Ces enchaînements, le plus souvent, ne doivent rien à la forme des expressions langagières qui évoquent ces pensées. Or, ils ne sont pas dus au hasard. L'idée I_1 a plus de chances d'entraîner l'idée I_2 dans un cerveau humain que, par exemple, une idée I_3 selon laquelle "le nombre 17637 n'est pas premier". Il est donc naturel de rechercher des mécanismes permettant d'expliquer le fait que les inférences se fassent dans certaines directions privilégiées. Il est également naturel de supposer que ce sont, en dernier ressort, les concepts qui, de par leurs propriétés causales, sont responsables des inférences que l'on produit à partir d'eux. Si l'on inverse la perspective, l'étude des inférences que les êtres humains produisent devient un bon moyen de caractériser les concepts.

La question que nous voulons explorer dans ce chapitre concerne l'étendue des tâches inférentielles que le système conceptuel prend en charge. S'il existe un système inférentiel indépendant, doté d'un ensemble de représentations et de mécanismes permettant la production du raisonnement, alors le système conceptuel en tant que tel peut être minimal. Prenons l'exemple d'une inférence classiquement qualifiée de déductive. On peut imaginer qu'une connaissance de la forme *tout A est B* serait disponible dans un module de raisonnement, tandis que le rôle du système conceptuel serait limité à des tâches comme la reconnaissance d'occurrences, par exemple *x est A*. Dans une telle conception, de nouvelles connaissances, comme la connaissance déduite *x est B*, pourraient être produites par des mécanismes propres au module de raisonnement. À l'opposé, on peut concevoir que le système conceptuel prenne en charge l'ensemble des opérations de raisonnement. Ainsi, la déduction de *x est B* à partir de *x est A* pourrait être entièrement due au fait que le concept *B* et le concept *A* entretiennent un rapport structurel. Le rôle inférentiel des concepts devient alors prépondérant. Les concepts, dans un tel schéma, ne servent pas seulement de base aux raisonnements, ils en constituent la cause. Ainsi, un jugement de catégorisation comme *x est B*, qui n'est pas issu d'une reconnaissance directe, serait dû au rôle inférentiel des concepts eux-mêmes.

La motivation principale pour conférer un rôle inférentiel aux concepts est liée à l'idée qu'il existe des inférences nécessaires, dont la validité serait due à la structure des concepts que de telles inférences mettent en jeu. Dans une première section, nous examinons cette position, ainsi que certaines des critiques auxquelles elle a donné lieu, notamment en ce qui concerne la possibilité de prédire cette partie du raisonnement par le seul emploi de définitions. Dans une deuxième section, nous abordons une critique plus fondamentale, concernant l'existence même d'une frontière entre inférences nécessaires et inférences

contingentes. Une telle critique débouche sur une vision globale du système conceptuel dans laquelle le rôle inférentiel d'un concept donné ne vient pas de sa seule structure, mais des liens qu'il entretient avec les autres concepts. Dans cette approche, le système conceptuel est conçu comme assez riche pour héberger l'ensemble des mécanismes de raisonnement. Nous évoquerons la critique principale adressée à un tel système, liée à son caractère holistique. Dans la troisième section du chapitre, nous abordons une manière radicalement différente d'expliquer le pouvoir inférentiel des concepts à l'aide de structures. L'idée n'est plus de doter les concepts d'une structure pour reproduire le raisonnement dans ce qu'il a de nécessaire, mais dans ce qu'il a de typique. La cause des inférences que l'on produit à partir d'un concept n'est pas entièrement contenue dans la structure de ce concept. Elle n'est pas non plus reportée dans la structure de l'ensemble du réseau conceptuel. Elle est due au caractère régulier ou exemplaire de ces inférences, représentées dans la typicité des traits qui caractérisent le concept.

4.1. Concepts et nécessité

Certaines inférences semblent emporter l'adhésion de tout être doté du bon sens cartésien. Ainsi, du constat exprimé par la phrase X a convaincu Y que A, un être rationnel qui possède le concept CONVAINCRE semble obligé de conclure avec certitude la phrase Y croit que A. Une manière d'expliquer le caractère nécessaire de cette conclusion consiste à dire qu'elle existe, en germe, dans le sens de la prémisse. Par contre l'inférence qui conduit à conclure une phrase comme X y a consacré beaucoup de temps présente un aspect contingent et dépend de diverses connaissances que l'on possède sur les individus X et Y. Comment distinguer ces deux types d'inférences et expliquer la nécessité apparente de la première ?

La nécessité de certains énoncés du langage semble être imposée par la seule compréhension du sens des mots qui les constituent. Le problème du sens, dans ce cas, se retrouve intimement lié à la question du jugement épistémique¹. Considérons les trois énoncés suivants et le jugement épistémique qu'on leur associe spontanément.

1. Un individu non marié n'est pas marié.
2. Un célibataire n'est pas marié.
3. Un adolescent n'est pas marié.

On perçoit intuitivement une démarcation entre les deux premières phrases et la troisième. Les deux premières présentent un caractère nécessaire qui manque à la troisième. La manière dont on justifierait ces phrases, dans la plupart des contextes, diffère. Si quelqu'un prétend ne pas accepter les énoncés 1 et 2, on commence par vérifier que cet interlocuteur a bien compris le sens de ces énoncés. Cette vérification, dans le cas du premier énoncé, se borne à tester la bonne compréhension des deux formes de négation. Dans le cas du deuxième énoncé, la seule vérification possible consiste à tester que l'interlocuteur connaît le sens du mot célibataire et son rapport avec le sens du mot marié. Dans les deux cas, le jugement épistémique semble lié, de manière mécanique, au résultat de la production et de la compréhension sémantique. Ce phénomène a donné lieu à des analyses philosophiques et a suscité la conception de formalismes logiques appropriés.

¹ Il est courant, chez de nombreux auteurs, d'employer les termes "vrai" et "vérité" pour désigner ce jugement épistémique, comme si ce jugement était évalué à l'aune d'une référence omnisciente et indiscutable. N'ayant pas besoin d'une telle hypothèse, nous essaierons d'éviter ces termes dans notre propre discours.

'L-true' is meant as an explicatum for what Leibnitz called necessary truth and Kant analytic truth. [...] A sentence \mathcal{F}_i is *L-true* in a semantical system S if and only if \mathcal{F}_i is true in S in such a way that its truth can be established on the basis of the semantical rules of the system S alone, without any reference to (extra-linguistic) facts. (CARNAP 1947 [9] p. 8)

Les deux premiers énoncés de notre exemple font partie de ces jugements analytiques. Leur acceptabilité ne dépendrait que du traitement sémantique du langage et serait du ressort de ce que nous appelons le système conceptuel. Le jugement épistémique concernant le troisième énoncé de l'exemple, en revanche, fait appel à des faits considérés comme conventionnels ou typiques², dépendant de la culture ou de notre expérience. Généralement, ce type de connaissance est considéré comme ne faisant pas partie du système conceptuel. L'emploi correct du concept ADOLESCENT ne requiert pas la connaissance selon laquelle un adolescent, typiquement ou conventionnellement, n'est pas marié. Les deux premiers énoncés de notre exemple sont remarquables car le jugement épistémique que leur compréhension déclenche n'a pas recours à ce type de connaissances considérées comme externes. Pour le premier, la seule capacité requise est la maîtrise du mécanisme de la négation. Pour le deuxième énoncé, le modèle sémantique doit expliquer le lien obligatoire qui existe entre le mot célibataire et l'expression non marié. Pour ce faire, il est naturel d'imaginer qu'il s'agit d'un lien structurel, par exemple en disant que le concept CÉLIBATAIRE, en tant que représentation, contient comme constituant le concept associé à l'expression non marié, par exemple en postulant que le concept CÉLIBATAIRE équivaut strictement à la définition "individu non marié". La particularité des jugements analytiques s'expliquerait alors par le fait que certains concepts contiennent d'autres concepts en tant que composants.

Suppose we understand some predicators in a given language; that is to say, we know which properties they express. Suppose, further, that we have experienced each of these properties; that is to say, we have, for each of them, found some things which, according to our observations, have that property. We can form compound predicators out of the given predicators with the help of logical particles. Then we understand a compound predicator because its meaning is determined by the meanings of the component predicators and the logical structure of the compound expression. It is important to notice that our understanding of a compound predicator is no longer dependent upon observations of any things to which it applies, that is, any things which have the complex property expressed by it. (CARNAP 1947 [9] p. 20)

De cette manière, le rôle inférentiel du concept lui serait conféré par sa définition. Dans le cas des inférences analytiques, la conclusion serait valide parce qu'elle se trouve en germe dans la structure du concept de départ. Il est tentant de rendre analytiques la plupart des inférences en augmentant les définitions, de manière à enfermer dans le concept lui-même, par exemple sous forme d'une conjonction, toutes les inférences qu'il peut produire.

² Si le lecteur considère que le sens du mot adolescent contient, de la même façon que le mot célibataire, l'idée qu'il s'agit d'un individu non marié, il suffit de le remplacer par le mot étudiant.

For example, the most common sense of the English noun “chair” can be decomposed into a set of concepts which might be represented by the [following] semantic markers [...]:

(object), (Physical), (Non-Living), (Artifact), (Furniture), (Portable),
 (Something with legs), (Something with a back), (Something with a seat), (Seat for one).

It is obvious that this analysis leaves out a considerable information. [...] [Another] way of thinking about semantic markers is as symbols that mark the components of senses of expressions on which inferences from sentences containing the expressions depend. They are, in this case, thought of as marking aspects of the logical form of the sentences in whose semantic representation they occur. The [first sentence] entails [the rest]:

There is a chair in the room.
 There is a physical object in the room.
 There is something nonliving in the room.
 There is an artifact in the room.
 There is a piece of furniture in the room.
 There is something portable in the room.
 There is something having legs in the room.
 There is something with a back in the room.
 There is a seat for one person in the room.

The semantic markers [listed above], which comprise a reading for “chair”, mark the elements in the sense of this word on which [these] inferences depend.

(KATZ 1972 [58] p. 40)

Dans cet exemple, le concept CHAISE se trouve doté d'une structure comportant un certain nombre de traits, chacun de ces traits pouvant donner lieu à des inférences dont la validité semble être aussi indiscutable que l'appartenance du trait au concept. On peut, bien entendu, imaginer que ces traits sont à leur tour composés d'autres traits, ce qui augmente la complexité de la structure. Les inférences considérées comme analytiques ne se résument toutefois pas à ce schéma d'extraction de traits. Elles peuvent concerner les arguments des prédicats, par exemple dans le cas des verbes transitifs.

A selection of causatives with circumstantial functions appears [below]:

Sue forced/pressured/tricked/talked Jim into singing.
 Sue got/forced/caused/coerced Jim to sing.
 [CAUSE ([SUE], [GO_{Circ}([JIM]_i, [TO_{Circ}([i SING])])])]
 Sue kept/restrained/prevented Jim from singing.
 [CAUSE ([SUE], [STAY_{Circ}([JIM]_i, [NOT AT_{Circ}([i SING])])])]
 Sue allowed/permitted Jim to sing.
 [LET([SUE], [GO_{Circ}([JIM]_i, [TO_{Circ}([i SING])])])]
 Sue released Jim from singing.
 [LET([SUE], [GO_{Circ}([JIM]_i, [FROM_{Circ}([i SING])])])]
 Sue exempted Jim from singing.
 [LET([SUE], [STAY_{Circ}([JIM]_i, [NOT AT_{Circ}([i SING])])])]

[...] The familiar inference patterns for CAUSE, LET, GO, and STAY appear here as usual. In [the first example], Sue's action results in Jim's coming to sing [...]. In [the second], Sue's action results in Jim's continuing *not* to sing. In [the third], Sue could have prevented Jim from singing, but she didn't, so Jim probably sang. In [the fourth], Sue was forcing Jim to sing, and now allowed him to stop; in [the fifth], Sue could have forced Jim to sing, but chose not to. In these last two, Jim ends up probably not singing.

(JACKENDOFF 1983 [49] p. 200)

Ces exemples illustrent le type de décomposition que l'on peut proposer pour le sens de verbes comme *forcer* ou *permettre*. Il s'agit dans chaque cas d'une "influence" exercée sur le cours d'un événement. L'événement est présenté comme le "passage" d'un individu vers une action, à l'aide de fonctions conceptuelles comme *ALLER* ou *RESTER*. L'influence est présentée comme le résultat de fonctions conceptuelles telles que *CAUSER* ou *LAISSER*. Les structures fonctionnelles des verbes sont appariées avec celles de leurs arguments, ce qui produit des structures plus complètes sur lesquelles peuvent agir des règles d'inférence. On peut ainsi inférer le fait que l'individu a fini par accomplir l'action en question ou non. Les inférences produites de cette manière présentent un caractère nécessaire, car elles sont, formellement, analytiques³.

L'hypothèse de l'existence de structures définitionnelles, internes au système conceptuel, ne saurait épuiser l'explication de l'ensemble des inférences possibles. En général, les auteurs acceptent qu'un certain nombre d'inférences fassent intervenir des mécanismes externes, et ne soient pas strictement liées à la structure des concepts.

It was mentioned earlier that determination of truth presupposes knowledge of meaning (in addition to knowledge of facts); now, cognitive meaning may be roughly characterized as that meaning component which is relevant for the determination of truth. The non-cognitive meaning components, although irrelevant for questions of truth and logic, may still be very important for psychological effects of a sentence on a listener, e.g., by emphasis, emotional associations, motivational effects. (CARNAP 1947 [9] p. 237)

Le corollaire implicite de ce type de concession est que, si on limite l'ambition à la reproduction du seul rôle épistémique des concepts, alors la démarche qui consiste à établir une équivalence entre chaque concept et une définition répond à la question. Le reste des phénomènes inférentiels, les effets d'emphase, les associations, *et cætera*, reste hors du système conceptuel proprement dit.

Le fait de définir chaque concept, de considérer que chaque concept est équivalent, du point de vue des inférences, à sa définition, constitue-t-il une solution acceptable au problème du jugement épistémique ? Il faut noter d'emblée qu'une partie essentielle du jugement épistémique est laissée de côté. Le système conceptuel, grâce aux structures qu'il héberge, serait responsable de tous les jugements épistémiques nécessaires et de ceux-là seulement. Cette explication par les structures n'épuise pas toutefois le jugement épistémique, qui puise le plus souvent ses justifications dans des connaissances factuelles ou des relations contingentes. Ainsi, lorsque nous jugeons que tel individu, étant étudiant, n'est pas marié, nous exploitons une régularité perçue dans notre environnement. Nous ne tirons pas ce jugement de la définition du concept *ÉTUDIANT*. Le caractère contingent de tels jugements épistémiques serait précisément dû au fait qu'il ne dépendrait pas des connaissances inhérentes à la signification. Il faut donc reconnaître l'existence de deux sources de connaissances : d'une part les connaissances proprement conceptuelles, enfermées dans les définitions, et d'autre part les connaissances contingentes, contenues dans une sorte de base de connaissances. Ces deux sources de connaissances seraient exploitées par le même mécanisme logique pour produire le jugement épistémique. Cette limitation du pouvoir des définitions, qui entraîne une complication du modèle, se double d'une difficulté concernant l'existence même de ces définitions censées reproduire le pouvoir inférentiel des concepts.

Nous pouvons schématiser ce qui précède par un ensemble à deux composants : (1) Un système conceptuel, consistant en un ensemble de structures ; (2) Un système de

³ Les deux types de structures définitionnelles cités précédemment sont donnés à titre d'exemples. Nous reviendrons longuement sur les types de structures attribués aux concepts et les mécanismes généraux utilisés pour expliquer les inférences que l'on peut produire à partir d'eux (CF. CHAPITRE 7).

raisonnement, comportant des mécanismes logiques capables d'exploiter les structures conceptuelles du premier système et hébergeant des connaissances contingentes, c'est-à-dire des liens contingents établis entre les concepts grâce à l'expérience. Ensemble, ces deux systèmes sont supposés expliquer le jugement épistémique.

Le point fort de l'approche réside dans l'explication élégante des inférences nécessaires. Grâce au remplacement des concepts par leur définition, ces inférences nécessaires sont rendues triviales. Si la signification du mot célibataire équivaut à la définition "individu et non marié", alors l'inférence Pierre est célibataire \models Pierre est non marié devient trivialement valide, puisqu'elle équivaut à l'inférence Pierre est un individu \wedge Pierre est non marié \models Pierre est non marié. Un mécanisme inférentiel logique standard, comportant notamment le schéma inférentiel $P \wedge Q \models Q$, suffit à expliquer la validité de cette inférence.

Malheureusement, le caractère esthétique de ce schéma n'est qu'apparent. Un premier problème concerne l'exigence de l'existence de définitions adéquates pour n'importe quel concept. Quelle définition substituer au concept ROUGE ? Assurément, le concept COLORÉ doit être inclus dans cette définition, puisque l'inférence X est rouge \models X est coloré est nécessairement valide. Cependant, cet élément de définition est insuffisant. Lorsqu'il s'agit par exemple de traduire l'inférence contingente si le feu est rouge, alors s'arrêter, on s'aperçoit que les concepts ROUGE et COLORÉ ont des comportements inférentiels différents. Une substitution par cette définition partielle, qui correspondrait à l'énoncé si le feu est coloré, alors s'arrêter, risque de produire des inférences erronées. Le problème vient de ce que l'on peut distinguer des concepts comme ROUGE, ORANGE ou VERT par leur comportement épistémique. Or, si ces concepts n'existent, du point de vue inférentiel, que sous la forme de définitions, et que leur définition ne comporte que le seul concept COLORÉ, leur différence, en ce qui concerne leur comportement épistémique, demeure inexplicée. Cette difficulté, évidemment, est liée au caractère incomplet de la définition retenue. Est-on capable de compléter la définition du concept ROUGE jusqu'à la lui rendre substituable ? La réponse est non (FODOR & GARRETT & WALKER & PARKES 1980 [38]). Quelle que soit la précision de la définition proposée⁴, il sera possible de trouver un contre-exemple, comme nous l'avons fait en mentionnant les concepts ORANGE et VERT. On pourra discriminer conceptuellement les concepts ROUGE et ROUGE₁, ce dernier étant le concept couvert par la définition. Dans la mesure où la discrimination est possible au niveau conceptuel, on s'expose à des erreurs si l'on substitue ROUGE₁ à ROUGE dans le calcul inférentiel.

Quite generally, if an informally valid argument turns on a definition, then there will be some clause that we can conjoin to the consequent which will make the corresponding bi-conditional true. Any informally valid argument which does not meet this condition can't be a definitional argument. [...] Even given the apparatus of definitional analysis, it looks as though some informally valid arguments can't be captured within the inferential apparatus of standard logic.

(FODOR & GARRETT & WALKER & PARKES 1980 [38] p. 272)

Si une définition peut être trouvée pour chaque concept, elle doit inclure les conséquents de toute inférence analytique non triviale (*informally valid argument*) effectuée à

⁴ L'exemple a été choisi par les auteurs cités précédemment pour qu'il soit d'emblée difficile de compléter la définition. Noter qu'une définition du concept ROUGE par une plage de longueurs d'onde serait grossièrement inadéquate. Elle est inadéquate d'un point de vue perceptuel puisque la perception de la couleur dépend significativement du contexte visuel, *id est* éclairement et zones voisines. Elle est inadéquate d'un point de vue conceptuel, puisque les délimitations conceptuelles ne sont pas asservies aux discriminations perceptives. Elle est inadéquate, enfin, d'un point de vue cognitif, puisque la valeur de longueur d'onde n'est pas supposée être disponible en tant qu'élément explicite au niveau du système conceptuel des individus.

partir du concept. L'impossibilité de terminer cette procédure, au moins pour certains concepts, fait que certains concepts, comme le concept ROUGE, restent sans définition. En conséquence, le projet de transformer toute inférence analytique non triviale en une inférence triviale en substituant la définition au concept est un échec. Il en résulte une contradiction : la logique standard ne peut expliquer les inférences analytiques non triviales, alors que leur analyticit     tait cens  e reposer sur un m  canisme logique. Dans ces conditions, sur quoi faire reposer l'analyticit   d'inf  rences comme Pierre est c  libataire \models Pierre est non mari   ou la chaise est rouge \models la chaise est color  e ?

L'id  e que l'on pourrait expliquer les inf  rences n  cessaires    l'aide de d  finitions est encore plus probl  matique en ce qui concerne les structures fonctionnelles propos  es pour d  finir les verbes. Selon cette approche, la d  composition des verbes jusqu'au niveau de composants simples comme CAUSE,   TAT,   V  NEMENT ou FORCE est cens  e expliquer la r  alisation de certaines inf  rences en   vitant de recourir    des r  gles contingentes. Dans un tel mod  le, l'inf  rence qui,    partir de la phrase A convainc B que X, conclut la phrase B croit X est valide car, premi  rement, l'assertion de d  part   quivaut, au niveau conceptuel,    la phrase A pousse B    croire X ; et deuxi  mement, il existe un m  canisme g  n  ral de raisonnement qui lie la repr  sentation conceptuelle de la phrase A pousse B    croire X    la repr  sentation conceptuelle de la phrase B croit X. Cette deuxi  me   tape est cens  e   tre une op  ration formelle sur les repr  sentations conceptuelles. Or, on voit bien que cette op  ration d  pend de la pr  sence, dans la repr  sentation conceptuelle de la pr  misse, du concept POUSSER. Elle pr  sente donc un caract  re contingent. Il ne s'agit pas d'une r  gle d'inf  rence g  n  rale qui, comme une r  gle formelle de la logique du type $P \wedge Q \models Q$, est susceptible de s'appliquer    n'importe quelle proposition. Il faut donc accepter l'existence, dans le syst  me de raisonnement, d'un certain nombre de r  gles contingentes comme celle qui prescrit que la r  alisation de l'  v  nement A pousse B    E entra  ne l'existence de l'  tat E. Malheureusement, cette exigence du projet d  finitionnel entra  ne sa ruine, puisque l'objectif poursuivi en tentant de rendre les inf  rences analytiques   tait de supprimer leur caract  re contingent.

[...] the idea that informally valid arguments will prove to be formally valid when couched at the semantic level is equivalent to the idea that only their logical form is relevant to determining the validity of semantic level arguments; and, as things now stand, there is simply no reason to believe that this is true. If it is not, then a reconstruction of informal validity may require an enrichment of the inferential apparatus of standard logic (*e.g.*, the incorporation of rules which govern the behavior of formulae containing words like "cause") even if it *also* assumes the existence of definitions.

(FODOR & GARRETT & WALKER & PARKES 1980 [38] p. 271)

La seule mani  re de sauver le projet consiste    consid  rer que le nombre de r  gles d'inf  rences qu'il faut ajouter au syst  me logique standard reste limit  . Cela revient    dire qu'il faut d  terminer un petit nombre de primitives, comme CAUSE,   TAT,   V  NEMENT, FORCE, susceptibles d'entrer dans des sch  mas inf  rentiels, au m  me titre que les connecteurs logiques comme ET, OU, IMPLIQUE⁵. C'est la t  che que se sont assign  e un certain nombre de s  manticiens (CF. CHAPITRE 5). Nous aurons l'occasion de mentionner de nombreuses autres critiques d'ordre technique    l'encontre de cette approche (CF. CHAPITRE 7). Nous consid  rons dans l'imm  diat une critique fondamentale, portant sur l'existence m  me de jugements analytiques non triviaux.

⁵ Noter que les connecteurs de la logique standard sont redondants. Il est par exemple possible d'exprimer l'ensemble des expressions de la logique des propositions    l'aide de la conjonction et de la n  gation. Il est m  me possible de les exprimer    l'aide du seul op  rateur d'incompatibilit  . L'enrichissement envisag   ici n'est donc pas anodin.

4.2. Concepts et connaissance

Le fait que les concepts puissent être considérés comme équivalents à des définitions semble un fait incontournable dès que l'on prend en compte le caractère nécessaire des jugements analytiques. Ainsi, on conclut qu'un célibataire est nécessairement non marié en raison de l'équivalence entre le concept CÉLIBATAIRE et sa définition "individu non marié". Or, cette équivalence définitionnelle ne va pas de soi.

There are those who find it soothing to say that the analytic statements of the second class reduce to those of the first class, the logical truths, by *definition*; 'bachelor', for example, is *defined* as 'unmarried man'. But how do we find that 'bachelor' is defined as 'unmarried man'? Who defined it thus, and when?
(QUINE 1953 [89] p. 24)

Le problème soulevé ci-dessus vient de ce que la notion d'analyticité repose sur la notion de synonymie, et que cette dernière demeure non définie. Pour imposer l'absolue nécessité d'un jugement analytique, il faut considérer que la définition est substituable au concept, *salva veritate*. Autrement dit, la définition est censée être un exact synonyme du concept. Or, comment établir une telle synonymie ? Un test empirique, quelles que soient son étendue et sa précision, ne saurait faire l'affaire.

For most purposes extensional agreement is the nearest approximation to synonymy we need care about. But the fact remains that extensional agreement falls far short of cognitive synonymy of the type required for explaining analyticity [...]. The type of cognitive synonymy required there is such as to equate the synonymy of 'bachelor' and 'unmarried man' with the analyticity of [the statement 'All and only bachelors are unmarried men'], not merely with [its] truth.
(QUINE 1953 [89] p. 31)

Il ne s'agit pas d'établir l'exactitude d'un jugement considéré comme analytique, mais bien d'établir sa nécessité. Or, aucune vérification empirique ne peut produire une telle nécessité. Si l'analyticité des jugements épistémiques reste la seule source de leur nécessité, on se retrouve confronté à une pétition de principe. Nous avons besoin de l'analyticité pour établir la synonymie.

[...] to say that 'bachelor' and 'unmarried man' are cognitively synonymous is to say no more nor less than the statement: 'All and only bachelors are unmarried men' is analytic.
(QUINE 1953 [89] p. 29)

Comme la synonymie était censée fonder l'analyticité, il en résulte une forte suspicion quant à la validité théorique de ces notions. Les seuls énoncés analytiques qui subsistent sont les tautologies stériles du type un homme non marié est non marié. Tous les autres énoncés, si l'on suit cet argument, doivent être considérés comme synthétiques. Autrement dit, une connaissance factuelle est impliquée dans tout jugement qui n'est pas syntaxiquement tautologique. En conséquence, un concept donné ne peut être responsable, du fait de sa seule structure interne, des inférences que l'on peut produire à partir de lui.

La remise en cause de la distinction analytique - synthétique est parfaitement acceptable si l'on trouve le moyen de doter le système conceptuel d'une structure lui permettant de guider l'ensemble des inférences, quelle que soit leur nécessité. Une telle approche consiste à ramener dans le concept toute connaissance acquise à son propos. Ainsi, le rôle inférentiel des concepts devient complet, puisque leur structure suffit à expliquer toute inférence que l'on produit à partir d'eux. Par exemple, on stockera sous forme de liens entre concepts tout ce qui

était considéré comme des inférences nécessaires dans l'approche précédente. On introduira de même les relations contingentes sous forme de liens. Le système conceptuel, ainsi conçu, prend la forme d'un grand graphe censé résumer l'ensemble des connaissances de l'individu.⁶

Dans ce genre d'approche du système conceptuel, un concept se caractérise par l'ensemble de la théorie dans laquelle il intervient. Ce changement de perspective a d'importantes conséquences. Le rôle de l'expérience dans la formation des connaissances est reconnu. On s'attend toutefois à ce qu'il reste confiné à une part marginale, car l'importance des liens que les connaissances entretiennent entre elles devient prépondérante. Par exemple, bon nombre de connaissances, inscrites dans le réseau des concepts, seraient dictées par l'exigence de simplicité et de cohérence (QUINE 1953 [89]). À ce titre, les connaissances du sens commun fonctionnent comme celles de la science. Dans les deux cas, les individus postulent l'existence d'ensembles d'entités et de relations de manière à projeter une structure efficace sur le flux de l'expérience. L'édifice, comme c'est le cas des mathématiques, est sous-déterminé par l'expérience et a pour principale raison d'être la simplicité et la cohérence.

Cette vision du système conceptuel correspond à celle qui est adoptée par un courant important de la psychologie développementale. Pour expliquer les transitions dans les systèmes de croyance des enfants, il semble utile de présenter les connaissances des plus jeunes, non comme des versions partielles et dégradées de celles de leur aînés, mais comme des théories à part entière. Par exemple, les enfants de cinq ans ont une notion de vivant qui ne les induit pas à grouper les plantes et les animaux en une seule classe. La plupart refusent qu'une plante puisse être malade ou puisse mourir de faim (CAREY 1985 [8] p. 168). Ce n'est qu'avec l'accès à une théorie élémentaire de la biologie, vers l'âge de neuf ans, que les enfants regroupent spontanément les animaux et les plantes dans la catégorie des êtres vivants (CAREY 1985 [8] p. 182). Pourtant, les plus jeunes enfants observent de nombreuses similarités entre animaux et plantes, par exemple le fait que les deux grandissent. Ce qui retient les enfants de cinq ans de former un concept d'entité vivante incluant les plantes est qu'un tel concept ne jouerait pas de rôle dans leurs théories prébiologiques. De fait, leur concept VIVANT trouve sa place dans des théories enfantines où il s'oppose au concept CASSÉ, *id est* non fonctionnel, ou au concept IMAGINAIRE. Ce type d'observation conduit de nombreux auteurs à l'idée que les concepts n'existent que par leur rôle inférentiel.

Concepts must be identified by the roles they play in theories. (CAREY 1985 [8] p. 198)

All representations will be theory-laden. (GOPNIK & MELTZOFF 1997 [43] p. 44)

Selon cette position, la raison d'être des représentations conceptuelles que nous développons se situe au sein de théories élaborées pour rendre compte du monde qui nous entoure. Par exemple, l'enfant âgé de neuf mois semble disposer d'une théorie du genre "un objet réapparaît généralement au même endroit". Lorsqu'un objet disparaît sous un foulard, l'enfant va le chercher sous un autre foulard où il a précédemment eu l'occasion de trouver l'objet. Son aîné, âgé de dix-huit mois, est capable de visualiser des trajectoires qui se situent hors de son champ visuel. Cela lui permet de former une autre théorie, du type "l'objet se trouve sur l'une des trajectoires invisibles possibles". Grâce à cette théorie plus puissante, il peut former, de manière productive, des hypothèses raisonnables sur la localisation d'un objet qui a disparu (GOPNIK & MELTZOFF 1997 [43] p. 101). Le développement ontogénétique du système conceptuel humain apparaît ainsi comme strictement analogue à l'avancement de la connaissance scientifique au cours de l'histoire. Certains changements conceptuels, comme celui qui a conduit Einstein à abandonner la notion de temps absolu, s'accompagnent de

⁶ Nous analyserons des exemples de formalismes de ce type plus tard (CF. CHAPITRE 7).

changements en profondeur et entraînent le remplacement de théories entières, par exemple celui de la théorie classique de Newton au profit de la théorie de la relativité. Il en irait de même du développement des individus (GOPNIK & MELTZOFF 1997 [43] p. 16).

L'idée selon laquelle le système conceptuel des individus fonctionnerait suivant les mêmes principes que la connaissance scientifique, accumulée au cours des siècles, ne va pas sans poser de problème. L'analogie nous laisse dépourvus lorsqu'il s'agit d'énoncer les principes d'évolution du système conceptuel sous l'effet de l'apprentissage. Les changements de paradigmes scientifiques ont été décrits comme dus à des causes exogènes, par exemple sociales. Ces mécanismes, dont la valeur prédictive est faible en ce qui concerne le contenu des théories scientifiques, ne peuvent pas être transposés dans le domaine cognitif. Nous reviendrons sur ce point (CF. CHAPITRE 7).

Le fait de décrire le système conceptuel comme une collection de théories crée également des difficultés lorsqu'il s'agit d'identifier les représentations qui sont évoquées par les mots du langage. Comment isoler ce qui se rapporte à un concept donné, si toute la connaissance se présente sous la forme d'un gigantesque graphe où tout se retrouve, directement ou indirectement, relié ? L'idée fondamentale est que, dans une théorie, les concepts n'existent que les uns par rapport aux autres.

No individual concept can be understood without some understanding of how it relates to other concepts.
(KEIL 1989 [60] p. 1)

Le risque théorique est de parvenir à un holisme conceptuel dans lequel la moindre inférence met en jeu potentiellement l'ensemble du réseau. Les théories dont dispose l'individu sont censées lui permettre de repérer des régularités et de produire des explications causales. Si ces théories sont toutes connectées entre elles, comment procéder, en pratique, pour prédire les inférences que l'individu produira en entendant un énoncé du langage ? Ces inférences peuvent en principe dépendre de l'état de n'importe quel endroit du réseau de connaissances, car tout point du réseau est relié de manière plus ou moins directe à la principale théorie mise en jeu dans l'énoncé. De ce fait, un même individu pourra produire des inférences radicalement différentes au cours du temps, même si, localement, sa théorie est restée inchangée. Comme l'effet cognitif d'un énoncé sur un individu est impossible à prédire, il en sera de même pour la communication. Une solution de ce problème de holisme peut consister à limiter l'interconnexion entre théories.

One of the dangers of proposing that concepts are intrinsically relational is the implication that they really are just the same as encyclopedic knowledge about the world. Everything is related to everything else in a vast network of roughly equal density. I will try to block this implication by arguing that beliefs tend to cluster in highly structured bundles with special properties that distinguish these clusters from the general interconnectedness of knowledge. Thus, lexical concepts are not viewed simply as the more or less arbitrary association of words to nodes in some vast network; instead, as a subset of all concepts, they tend to be associated with only particular configurations that are isolatable from other aspects of knowledge.
(KEIL 1989 [60] p. 4)

Circonscrire chacune des théories prêtées à l'individu est une nécessité dans ce cadre. Sans cela, on ne peut espérer prédire ni la production d'inférences, ni les changements de théorie au cours du développement, ni la possibilité même de communication puisque deux individus n'ont aucune chance de disposer du même réseau de théories. Le problème est que l'on ne dispose d'aucun critère pour circonscrire les théories. Une théorie sur les voitures doit inclure des connaissances sur leur fonction, donc sur la locomotion ; elle doit donc être connectée à une théorie sur la locomotion animale, à une théorie sur le mouvement, *et cætera*.

De même, elle doit être liée à une théorie chimique sur la combustion des hydrocarbures dont on récupère l'énergie dans le moteur. Elle doit être également liée à une théorie économique sur la rentabilité d'un véhicule. Qui peut dire que la théorie relative aux voitures s'arrête à tel point précis, qu'ensuite il s'agit d'une théorie chimique ou économique ? La théorie du cylindre et des bougies relève-t-elle de la voiture, de la chimie ou de la mécanique ? Celle des bougies et de l'allumeur ne serait-elle pas mieux placée au sein d'une théorie électrique ? En l'absence de critère de délimitation, ce type de question reste sans réponse, et la théorie des théories tombe dans un holisme sans pouvoir de prédiction.

Le problème qui vient d'être soulevé n'est pas seulement technique. Il ne s'agit pas de trouver un critère commode permettant de donner une délimitation raisonnable des théories au sein du graphe des connaissances conceptuelles. Il s'agit d'isoler, en suivant une règle générale, le sous-ensemble des connaissances relatives à un concept qui lui confère l'intégralité de son rôle inférentiel. Or, ce rôle inférentiel, que la théorie des théories prétend expliquer de manière exhaustive, s'exerce lorsque le concept intervient dans le sens d'un énoncé. Les connaissances jugées propres au concept doivent donc être délimitées de manière à déterminer la participation du concept dans l'élaboration du sens de l'énoncé. Un simple "rayon d'influence" dans le graphe des connaissances ne saurait répondre à cette exigence.

Dans l'approche qui consiste à conférer des structures définitionnelles aux concepts, ce problème ne se pose pas. Le principe de compositionnalité⁷ permet de caractériser entièrement la structure conceptuelle d'une expression linguistique complexe à partir de celle de ses composants, à l'aide d'un calcul approprié. Le principe de compositionnalité est facile à respecter avec un système de définitions, car il n'y a pas de différence de nature entre la structure conceptuelle des énoncés et celle des concepts eux-mêmes. On peut ainsi, aisément, déduire le comportement inférentiel du syntagme chat noir, si l'on suppose que sa représentation conceptuelle a une structure qui dérive de celle du concept CHAT qui dérive à son tour de celle du concept ANIMAL. En particulier, on peut reproduire le caractère nécessaire de l'inférence allant de la phrase il y a un chat noir sur le toit à la phrase il y a un animal noir sur le toit, en utilisant, par exemple, une simple déduction logique sur les définitions.

À partir du moment où l'on renonce aux définitions et que le concept CHAT est caractérisé par sa place dans un réseau de liens conceptuels, la caractérisation conceptuelle du syntagme chat noir devient, pour le moins, délicate. Si l'on tente de respecter le principe de compositionnalité en restant dans le cadre des théories, on risque de réintroduire le problème de l'analyticité que le système des relations devait permettre de contourner.

For, on the one hand, an inference is compositional iff its validity is determined by the *inferential roles* that its constituents contribute; and, on the other hand, an inference is analytic iff its validity is determined by the *semantic values* of its constituents; but, on the third hand (as it were), [inferential role semantics] says that the meaning of a constituent is the same thing as the inferential role it contributes to the expressions that contain it. So to repeat, the cost of representing an inference as compositional is that you then have to represent it as analytic. (FODOR & LEPORE 1992 [39] p. 181)

Si l'on suppose que les concepts sont censés, par leur seule structure, expliquer l'intégralité du rôle inférentiel d'un composé, on ne peut engendrer que des jugements analytiques. Or, nous avons commenté le caractère indéfendable d'une telle position.

Peut-on renoncer à la compositionnalité des théories ? Dans ce cas, comment définir la théorie qui caractérise un composé ? La théorie des chats noirs est-elle un sous-ensemble de la théorie des chats, ou un sous-ensemble de la théorie des couleurs ? À partir de la phrase il y a un chat noir sur le toit, la conclusion il y a un animal noir sur le toit a-t-elle le même statut qu'une

⁷ Nous reviendrons longuement sur ce sujet (CF. CHAPITRE 6).

conclusion contingente comme il y a un chat méchant sur le toit ? N'a-t-on pas besoin, pour réaliser cette deuxième inférence, de supposer l'existence d'une théorie spécifique aux chats noirs ? Ces difficultés, et d'autres que nous mentionnerons plus tard, conduisent à penser que le système des théories est incapable de rendre compte de la production des inférences adéquates à partir des compositions conceptuelles (CF. CHAPITRE 7).

La théorie des théories était motivée par le rejet des définitions et le souci d'effacer la différence entre jugement analytique et jugement synthétique. Pour ce faire, le système conceptuel était doté de toutes sortes de liens inter-conceptuels, nécessaires ou contingents, de manière à être en charge des jugements épistémiques. On peut imaginer, à l'inverse, de renoncer à l'idée que le système conceptuel possède un quelconque pouvoir inférentiel.

Une manière radicale de remettre en cause la distinction entre inférences analytiques et inférences synthétiques est de considérer que les inférences que l'on peut produire à partir d'un concept n'ont rien à voir avec la façon dont on le caractérise. Dans ce cas, la caractérisation des concepts, avec leurs propriétés intentionnelles et causales, se fait selon des considérations autres, par exemple le lien du système conceptuel avec le langage ou avec la perception.

[...] if Quine is right [...] then what you mean can't be reduced to what inferences you are prepared to accept. [...] Because, Quine argues, what inferences you are prepared to accept [...] depends not only on what you intend your words to mean, but also how you take the (non-linguistic) world. [...] Quine's rejection of analyticity, insofar as it's actually argued for (and insofar as we are proposing to concede it), is a rejection only of the possibility of an *epistemic* criterion for "true in virtue of meaning." In principle, at least, everything else [about semantics] remains wide open. (FODOR & LEPORE 1992 [39] p. 57)

Ainsi, il ne peut pas y avoir adéquation entre la signification d'un mot et les connaissances relatives au concept associé qui nous permettent de produire des inférences. Toutes les connaissances sont en partie *a posteriori*, elles ne sauraient résulter du seul fait de posséder le concept. En conséquence, il faut distinguer la signification d'un mot et son rôle épistémique, ce dernier dépendant de facteurs contingents.

Le constat selon lequel il n'existerait pas de jugement analytique rend illusoire tout espoir de rendre "locale" la confirmation des raisonnements (QUINE 1953 [89]). Une mise en adéquation de connaissances avec l'expérience peut impliquer n'importe quel élément du réseau de connaissances. Pour autant, on ne saurait automatiquement conclure de ce holisme de confirmation que la signification elle-même ne peut être locale (FODOR & LEPORE 1992 [39]). À partir du moment où un concept n'est pas assimilé à son pouvoir inférentiel, il est possible de refuser le holisme sémantique tout en acceptant le holisme de confirmation. Cela suppose d'abandonner la vision "pragmatiste" qui réduit un concept à ses effets sur le jugement épistémique. Dans une perspective cognitive, la signification est relative à la représentation conceptuelle et à son ancrage (CF. CHAPITRE 3). Elle peut être, en principe, divorcée du rôle inférentiel. À la limite, on peut considérer que rien, dans la représentation conceptuelle elle-même, ne contraint les inférences dans lesquelles elle est impliquée. Ces inférences seraient du ressort d'un mécanisme de raisonnement, alimenté par une base de connaissances. Ainsi, les concepts n'auraient aucun pouvoir inférentiel propre. Ce pouvoir leur serait conféré de l'extérieur. Le holisme, dans ce cas, serait celui de la base de connaissances, non celui des significations. Cette position sera discutée et évaluée plus tard (CF. CHAPITRE 8).

Le passage aux théories était une tentative pour remédier à l'incapacité des définitions à reproduire la nécessité des inférences auxquelles un concept peut conduire. Il existe cependant un moyen d'augmenter les structures conceptuelles pour expliquer les inférences, même si

l'on observe que la plupart de ces inférences n'ont aucun caractère nécessaire. En particulier, comme nous allons le voir, il est possible d'expliquer les inférences présentant un aspect régulier ou exemplaire à l'aide de structures appropriées.

4.3. Concepts et typicité

L'une des caractéristiques manifestes de nombreuses inférences que nous faisons spontanément est leur aspect incertain. Lorsque nous supposons qu'un objet, dont nous savons qu'il s'agit d'un oiseau, vole, notre inférence est souvent valide. Ce jugement est cependant incertain, car nous nous exposons à l'erreur. Lorsque nous décidons qu'une boîte de peinture doit être classée parmi les jouets, nous tolérons encore une incertitude, liée cette fois au flou du jugement. L'ambition de l'approche que nous allons présenter dans cette section est de doter les concepts d'une structure permettant la production de ces jugements incertains.

Une façon de comprendre cette incertitude liée aux inférences consiste à la faire dériver des propriétés de la capacité humaine de catégoriser les objets perçus.

To have a concept of X is to know something about the properties of entities that belong to the class of X, and such properties can be used to *categorize* novel objects. Conversely, if you know nothing about a novel object but are told it is an instance of X, you can *infer* that the object has all or many of X's properties; that is, you can run the "categorization device in reverse."
(SMITH & MEDIN 1981 [96] p. 8)

L'opération de catégorisation est, dans certaines approches, à la base de la plupart des jugements inférentiels. Or, la manière dont les sujets catégorisent est incorrectement prédite par les modèles qui dotent les concepts de propriétés nécessaires. De nombreuses expériences ont montré que l'appartenance d'une série d'objets à une catégorie donnée ne montrait pas, le plus souvent, de transition discontinue, contrairement à ce que l'on attendrait si la catégorie était définie par des traits binaires (ROSCH 1975 [92]). Les tableaux suivants indiquent les moyennes des jugements d'appartenance, entre 1 et 7, donnés par un ensemble de 209 sujets (ROSCH 1975 [92]).

<i>Furniture</i>		<i>Vehicle</i>	
Chair	1.04	Automobile	1.02
Table	1.10	Truck	1.17
Dresser	1.37	Jeep	1.35
Desk	1.54	Van	1.95
Bed	1.58	Train	2.15
Divan	1.70	Bicycle	2.51
Buffet	2.89	Yacht	3.76
Telephone	6.68	Elevator	5.90

Pour expliquer ce phénomène, il est naturel de penser que le jugement des sujets se fonde, non sur une série de tests binaires, mais sur une mesure de similarité. La représentation d'un concept est censée permettre cette mesure de similarité.

In this case, cognitive representations of categories clearly contained more of the information needed to respond to category members which had been rated good examples of the category [...] than to respond to category members which had been rated bad examples. In other words, cognitive representations of categories appeared to be more similar to the good examples than the poor examples. (ROSCH 1975 [92] p. 225)

Un concept pourrait être identifié à une moyenne des propriétés des instances de la catégorie correspondante, un prototype.

We can only come up with the following two assumptions [of the probabilistic view]: (1) the representation of a concept is a summary description of an entire class; and (2) the representation of a concept can *not* be restricted to a set of necessary and sufficient conditions; rather, it is some sort of measure of central tendency of the instances' proprieties or patterns. (SMITH & MEDIN 1981 [96] p. 61)

Les concepts ressemblent ainsi à des représentations moyennes. Certaines inférences concernant les objets perçus peuvent être prédites par la distance entre l'objet perçu et cette représentation moyenne. Cette manière de décrire les concepts résout certaines difficultés du jugement épistémique. L'inférence selon laquelle un oiseau vole n'est pas nécessaire, pas plus que celle qui conduit à penser qu'il chante. Ces deux informations seraient exclues du concept OISEAU dans un système à base de définitions nécessaires. Or, le caractère contingent des deux inférences est relatif : la première est généralement considérée comme une caractéristique presque universelle des oiseaux, la seconde comme une propriété fortuite de certains d'entre eux. Le système des prototypes permet de reproduire cette gradation dans la sûreté du jugement.

Il existe une variété de moyen de doter les concepts d'une structure de manière à reproduire les jugements incertains. Le plus naturel est sans doute d'adapter les définitions à base de traits en attachant une probabilité à chacun de ces traits. On ne retient dans la structure des concepts que les traits qui possèdent une descriptivité (*cue validity*) suffisante⁸.

The second assumption of the featural approach is the critical one: the features that represent a concept are salient ones that have a substantial probability of occurring in instances of the concept. More precisely, if F_i is a feature and X_j a concept, F_i will be a feature of X_j to the extent that (1) F_i is salient (either perceptually or conceptually), and (2) the probability of F_i given X_j , $P(F_i | X_j)$, is high (for example, F_i tends to be true of an instance labeled as an X_j). (SMITH & MEDIN 1981 [96] p. 62)

La descriptivité d'une catégorie se déduit de celle des traits qui la composent. Elle mesure la capacité de ces traits à prédire la catégorie, et par là-même la manière dont la catégorie se différencie des autres catégories.

Il est également possible d'incorporer la gradualité dans les traits eux-mêmes, et non seulement dans leur probabilité d'occurrence. Cette deuxième variante caractérise un concept comme par une structure de traits dotés d'une plage de valeurs typique.

⁸ Cette notion correspond à la probabilité conditionnelle $P(F|C)$ d'observer le trait F dans la catégorie C .

The second assumption [of the dimensional approach] has two parts: (1) any dimension used to represent a concept must be a salient one, some of whose values have a substantial probability of occurring in instances of the concept; and (2) the value of a dimension represented in a concept is the (subjective) average of the values of the concept's subsets or instances on this dimension. [...] Part 1 is the dimensional equivalent of the featural assumption [...]. Knowing only that a dimension is represented in a concept does not tell us what particular value of that dimension is represented, so part 2 is needed to specify the unknown value. (SMITH & MEDIN 1981 [96] p. 102)

La validité des inférences que l'on produit à propos d'instances de tels concepts va dépendre de la typicité des traits présents dans la structure du concept, affectée de la probabilité du trait. Cette grandeur peut être comparée à des seuils, ce qui permet de considérer l'inférence tantôt comme valide, tantôt comme invalide, tantôt comme douteuse. Un tel système permet donc d'expliquer une réponse du type "je ne sais pas" lorsque l'on demande si un tapis est une pièce d'ameublement.

Certaines variantes de l'approche fondée sur la typicité confèrent aux représentations conceptuelles une structure plus concrète que celle d'assemblages de traits moyens et fréquents. Ainsi, un concept serait représenté par un ensemble réduit d'exemplaires qui peuvent être des instances ou des sous-catégories (SMITH & MEDIN 1981 [96] p. 145). Ceci permet d'expliquer que certains jugements impliquant des concepts soient sensibles au contexte. Ainsi, les jugements concernant un oiseau domestique et ceux qui portent sur un oiseau rapace seraient produits par des exemplaires différents du concept OISEAU, exemplaires qui font partie intégrante de cette représentation.

Le caractère concret de la représentation conceptuelle peut rejoindre la vision empiriste selon laquelle les concepts seraient représentés par des percepts plus ou moins généralisés. Les concepts pourraient correspondre à des patrons (*templates*) inanalysables en termes de traits, et le jugement de catégorisation se réduirait à une mesure globale de la ressemblance entre l'objet perçu et le patron (SMITH & MEDIN 1981 [96] p. 131). Cette ressemblance serait suffisante pour conférer un rôle inférentiel aux concepts : nous déduirions qu'un pigeon est un oiseau parce que le patron PIGEON peut être superposé, de manière globale, au patron OISEAU.

Cette contiguïté possible de l'approche fondée sur la typicité et d'une théorie empiriste de l'ancrage (CF. CHAPITRE 3) permet de voir la première comme un prolongement de la seconde. De même qu'un concept empiriste est défini comme une moyenne statistique de paramètres perceptuels, un concept, dans la théorie des prototypes, comporte des traits conceptuels dont les valeurs et la fréquence sont le fruit d'une moyennage. Ce rapprochement permet d'expliquer d'une part l'ancrage des concepts, et d'autre part leur capacité à engendrer des inférences incertaines, par le même mécanisme de représentation moyenne.

La nature statistique des catégories typiques ouvre la voie à des mécanismes d'apprentissage particulièrement simples. Les systèmes de regroupement (*clustering*) comme il en existe en statistique (par exemple l'algorithme des nuées dynamiques) ou en intelligence artificielle (par exemple les techniques SBL) permettent de produire des catégories telles que la distance moyenne entre objets soit minimale au sein des catégories et maximale entre elles (CF. ANNEXE).

La théorie des prototypes présente l'intérêt indéniable de rendre compte de toute la gamme des inférences, en incluant les jugements incertains. Toutefois, plusieurs critiques lui ont été adressées. L'une d'elle concerne le lien qui est postulé entre la ressemblance d'un objet perçu avec un prototype et l'appartenance de cet objet perçu à la catégorie correspondante. Par exemple, l'appartenance à la catégorie "nombre pair" n'est pas graduelle, et pourtant certains nombres pairs comme deux ou quatre apparaissent comme plus typiques que les autres (ARMSTRONG & GLEITMAN & GLEITMAN 1983 [1]). Le problème ne se pose

pas seulement pour les concepts scientifiques. Chaque individu a un jugement personnel sur les traits typiques du concept GRAND-MÈRE, qui peut être lié à l'image de sa propre grand-mère. Pourtant, ce même individu statuera sur le fait qu'une personne est grand-mère non d'après son apparence, mais d'après le constat qu'elle a des descendants de deuxième degré. Ce divorce entre prototypicité et jugement d'appartenance rompt le lien théorique qui fondait le caractère graduel de ce jugement d'appartenance. Du même coup, on perd le fondement théorique de bon nombre d'inférences que l'on est censé produire à partir de la possession des concepts. Si les catégories ne peuvent plus être définies par une relation de ressemblance avec un élément moyen ou avec un nombre restreint d'exemplaires, il faut imaginer deux gradients plus ou moins indépendants, un gradient de typicité et un gradient d'appartenance (KAMP & PARTEE 1995 [55]). Cette complication affaiblit cependant la théorie qui doit expliquer la genèse de ces deux gradients.

Une autre limitation de la théorie des prototypes est liée à son caractère statistique. La théorie prévoit que les capacités inférentielles des individus sont sensibles à leur expérience. La première fois qu'un individu observe un oiseau dépourvu de plumes, il est surpris. Si de tels oiseaux deviennent fréquents dans son environnement, il finira par s'habituer et ne sera plus en situation d'étonnement. La théorie standard des prototypes ne permet pas de prédire un quelconque étonnement. Certes, l'utilisation de seuils permettrait à un système à base de catégories typiques de détecter une exception : un objet perçu qui serait atypique par l'un de ses traits seulement. Un tel objet pourrait ainsi être classé comme atypique au sein de sa catégorie. En cas de répétition d'expériences similaires, soit la catégorie se modifie, soit il se forme une nouvelle catégorie. Dans les deux cas, la théorie prévoit correctement la fin de l'étonnement avec l'habitude. Ce que ce genre de théorie ne prévoit pas, cependant, c'est que l'individu peut se trouver en situation d'étonnement permanent. Par exemple, si notre sujet observe des oiseaux dépourvus de bec, il en inférera que ces animaux ne peuvent pas s'alimenter. Cet étonnement épistémique n'est pas sensible à la répétition du stimulus : quel que soit le nombre de ces oiseaux mutants qu'il lui sera donné d'observer, sa déduction restera valide et son étonnement également. Contrairement à l'exemple des plumes, la modification de la catégorie ou la formation d'une nouvelle catégorie ne changent rien. L'étonnement épistémique ne peut cesser que par une explication, par exemple que ces oiseaux se nourrissent par une opération de succion. L'histoire des sciences est féconde en exemples d'étonnements durables. Avant que le mécanisme d'écholocation soit avancé au milieu du siècle dernier, la propriété des chauves-souris de se mouvoir dans l'obscurité totale restait la source d'un étonnement permanent. La théorie des prototypes est incapable de reproduire le type de jugement conduisant à l'étonnement épistémique. Dans un système cognitif où tous les jugements peuvent être faits et défaits par le fait de l'habitude, il ne peut y avoir de contradictions logiques. L'inférence x est un oiseau \rightarrow x possède un bec ne peut, dans la théorie des prototypes, avoir d'autre source que les régularités de l'environnement. Il n'y a pas de place pour les explications causales et pour le trouble épistémique causé par l'observation d'une mise en défaut de la causalité. En ce sens, il est permis d'affirmer que la théorie des prototypes, telle qu'elle est décrite en moyenne, ne suffit pas à rendre compte des capacités inférentielles humaines. Nous reviendrons sur ce point lorsqu'il s'agira de proposer des mécanismes capables d'expliquer les jugements incertains tout en ménageant la possibilité d'un jugement épistémique (CF. CHAPITRE 9).

Une troisième difficulté que l'on rencontre en cherchant à appliquer la théorie des prototypes concerne les composés linguistiques. La prédiction des capacités inférentielles d'un composé linguistique comme pomme rayée ne pose aucun problème dans un système de définitions. Pour obtenir la définition du composé, il suffit d'utiliser les opérations habituelles de la logique classique. Ainsi, le concept associé au syntagme pomme rayée pourra être défini comme la conjonction logique des définitions des concepts POMME et de RAYÉ. Si l'on

abandonne les définitions binaires en faveur de coefficients numériques d'appartenance ou de typicalité, alors il faut donner un nouveau sens à la conjonction qui apparaît dans la définition d'une pomme rayée. Ceci peut être fait dans le cadre de la logique floue, en représentant le jugement de typicalité par une fonction attribuant une valeur numérique entre 0 et 1 aux différents objets. Par exemple, une chaise sera une pièce d'ameublement à 99%, une lampe à 70% et un téléphone à 40%. Dans un tel cadre, on peut dire que si un objet est pomme à 70% et rayé à 60%, alors il est pomme rayée au plus à 60%, selon la règle du minimum⁹. Malheureusement, il est très difficile d'accorder ce type de calcul avec les résultats suggérés par l'intuition. En particulier, pour un objet qui est pomme à 50%, le coefficient alloué à la conjonction "pomme et non pomme" sera 0.5, alors que l'on attend 0 en raison de la contradiction¹⁰ (OSHERSON & SMITH 1981 [79]). De plus, ces calculs rendent impossible qu'un objet soit plus typique pour la propriété combinée que pour la propriété de base (OSHERSON & SMITH 1981 [79]). Or, un perroquet est plus typique en tant qu'oiseau domestique qu'en tant qu'oiseau. Différents moyens ont été proposés pour remédier à ces difficultés. Par exemple, la théorie des super-valuations consiste à conserver des valeurs binaires lorsque l'appartenance est hors de doute, et à ne pas allouer de valeur d'appartenance lorsque ce n'est pas possible. Ainsi, pour un objet J qui serait classé comme pomme à 50%, l'expression $\text{pomme}(J)$ ne reçoit pas de valeur de vérité. En revanche, la conjonction $\text{pomme}(J)$ et $\text{non pomme}(J)$ peut recevoir la valeur 0, car toutes les manières d'instancier la super-valuation fournissent ce résultat (KAMP & PARTEE 1995 [55]). Pour assigner une valeur à un composé comme pomme rayée, il est nécessaire de tenir compte de la dissymétrie entre le nom et l'adjectif. Le coefficient associé à l'expression $\text{rayé}(J)$ doit être "re-calibré" dans le contexte fourni par l'expression $\text{pomme}(J)$ avant d'entrer dans le calcul de la combinaison (KAMP & PARTEE 1995 [55]).

Noter que ce genre de mécanisme échoue lorsqu'il s'agit d'expliquer des combinaisons de type lion en pierre. La représentation moyenne que l'on peut former à propos d'une expression comme lion en pierre ne semble pas déductible de la représentation moyenne de lion et de la représentation moyenne de pierre. Les inférences que l'expression lion en pierre peut occasionner ne sont donc pas déductibles, au sein de la théorie des prototypes, des inférences que les composants lion et pierre peuvent susciter. De plus, la représentation moyenne ne semble même pas nécessaire pour la formation du jugement d'appartenance des concepts composés.

Thus the extension of *male nurse* is determined in the simplest way possible, namely as the standard logical conjunction of the constituent concepts *male* and *nurse*. [...] This establishes that no prototype is needed to fix its extension. But it is not just that no prototype is required for this; it is hard to see how the prototype for *male nurse* could play any significant role in determining membership. (KAMP & PARTEE 1995 [55] p. 168)

La théorie des super-valuations ne suffit pas non plus à expliquer de tels composés. Aucune re-calibration du prédicat pierre ne donne une intersection non vide avec le prédicat lion. Cependant, on peut supposer que c'est précisément cet échec qui provoque l'activation

⁹ Plusieurs règles ont été proposées pour remplacer les opérations binaires de conjonction et de disjonction. Par exemple, le coefficient attaché à la conjonction peut être représenté par le produit des coefficients associés aux opérandes. Le choix le plus courant consiste à prendre le minimum pour la conjonction et le maximum pour la disjonction. Ces définitions redonnent les résultats de la logique classique lorsque les coefficients sont des valeurs binaires 0 et 1.

¹⁰ De même, la disjonction "pomme ou non pomme" se verra accorder un coefficient de 0.5 au lieu de 1. Une version multiplicative de la conjonction donnera 0.25 à celle de "pomme et non pomme", ainsi qu'à celle de "pomme et pomme" !

d'un sens dérivé pour le nom modifié. Le mécanisme de modification peut s'appliquer au sens dérivé, par exemple la forme du lion dans notre exemple. Le choix de ce sens dérivé se révèle cependant difficile à modéliser dans le cadre des théories actuelles (KAMP & PARTEE 1995 [55]).

Ces difficultés concernant la déduction des propriétés des concepts associés aux expressions composées ne concernent pas uniquement la théorie des prototypes. Dans le chapitre suivant, nous allons étudier ce problème en détail en abordant le caractère compositionnel des concepts.

Conclusion

Ce chapitre nous a permis d'étudier l'interface entre le système conceptuel et le raisonnement. La question première, lorsqu'on s'intéresse à cette interface, porte sur le partage, entre le système conceptuel et le système de raisonnement, des connaissances qu'un mot de la langue peut convoquer. Une manière d'effectuer le partage consiste à ne retenir dans le système conceptuel que les connaissances nécessaires, en les incluant dans des définitions. Une autre acception du système conceptuel y inclut l'ensemble des connaissances typiques relatives aux concepts. Les concepts sont caractérisés, dans cette approche, par des prototypes ou des exemplaires. Le système conceptuel héberge ainsi non seulement les connaissances nécessaires, mais également les régularités et les connaissances incertaines associées aux concepts. Une version plus radicale, où le système conceptuel héberge tout ce que l'on peut savoir sur les concepts, conduit à des modèles comme la théorie des théories dans laquelle un concept est caractérisé par l'ensemble des connaissances dans lesquelles il est susceptible d'intervenir. La version opposée est une vision du système conceptuel dans laquelle les connaissances ne jouent aucun rôle dans la caractérisation des concepts.

Chapitre 5 :
Caractère compositionnel des
concepts

Introduction

Toute langue offre la possibilité à ses locuteurs d'exprimer une variété infinie de sens à l'aide de sons. Chaque mot d'une langue permet d'effectuer un couplage entre une forme sonore et une représentation conceptuelle. La grammaire de la langue permet de prolonger ce couplage de manière à ce qu'un sens puisse être associé à des expressions comportant plusieurs mots. Ainsi, la représentation conceptuelle associée à un énoncé doit être élaborée à partir de la combinaison des concepts associés à ses composants, ce qui confère un caractère compositionnel aux concepts. Or, si la combinaison des mots qui forment une phrase apparaît de manière tangible, il n'en est pas de même de la combinaison des concepts. Ces deux combinaisons s'effectuent-elles en parallèle ? Dans quelle mesure servent-elles de support l'une à l'autre ? Un parallélisme étroit supposerait une correspondance formelle entre la structure grammaticale d'un énoncé et sa structure conceptuelle. Si l'on renonce à une telle hypothèse, il faut imaginer deux mécanismes de combinaison indépendants, l'un grammatical, l'autre conceptuel, produisant des structures de nature différente. Cela suppose l'existence d'une interface définie pour que ces deux mécanismes puissent coopérer et se synchroniser.

Parmi les contraintes qui limitent les possibilités de combiner des mots, bon nombre peuvent être considérées comme indépendantes du système conceptuel. Ainsi, la suite de mots la brûle écrouler maison ne sera pas acceptée par un locuteur du français, non parce qu'il est impossible de lui trouver un sens, mais parce qu'elle ne respecte pas certaines règles, de nature syntaxique, qui limitent l'agencement des types de mots. Peut-on imaginer dresser un catalogue de règles, de type grammatical, telles que leur respect garantit l'interprétation de l'énoncé ? Si ces contraintes sont suffisamment strictes tout en restant correctes, on peut imaginer que les énoncés qui respectent ces règles auront, de manière garantie, une interprétation. Certaines approches, notamment en linguistique computationnelle, ne posent pas de différence de nature entre les contraintes de type syntaxique et les contraintes de type sémantique¹. Ainsi, une règle sémantique stipulant que l'agent d'un verbe d'attitude comme croire doit être de type humain serait de même nature qu'une règle syntaxique d'accord en nombre entre le sujet et le verbe. Une telle perspective suppose que le fonctionnement du système conceptuel est analogue à celui d'une langue. De nombreux auteurs, cependant, refusent une telle perspective. Ils admettent que les mécanismes de composition du système conceptuel et les mécanismes grammaticaux puissent être qualitativement différents. Dans ce cas, il est nécessaire de définir une interface entre la combinaison des mots et celle des concepts.

Dans ce chapitre, nous examinons successivement différentes manières de concevoir le lien entre la construction grammaticale et la composition conceptuelle. Dans une première section, nous présentons la manière dont l'interface syntaxe - sémantique est conçue dans l'approche de la grammaire générative. Cette interface est caractérisée par l'existence d'une "forme logique" qui n'est pas directement observable. Dans une deuxième section, nous présentons une acception différente de la notion de la forme logique, celle de la sémantique formelle. La forme logique y apparaît comme une traduction de la forme linguistique propre à permettre la production d'inférences valides. Dans la troisième section, l'interface syntaxe - sémantique sera décrite comme un appariement précis entre la structure syntaxique et une structure conceptuelle. La quatrième section présente une autre manière de calculer le

¹ L'habitude, dans les modèles basés sur des structures de traits, consiste à regrouper dans un même champ *syn-sem* les traits syntaxiques et sémantiques d'un syntagme, en utilisant le même mécanisme d'unification pour tester les deux types de contraintes.

sens à partir de l'ordonnement des mots, en introduisant des procédures associées aux concepts. La dernière section aborde la construction du sens des énoncés dans les théories de la linguistique cognitive à l'aide des représentations schématiques.

5.1. L'apport de la syntaxe

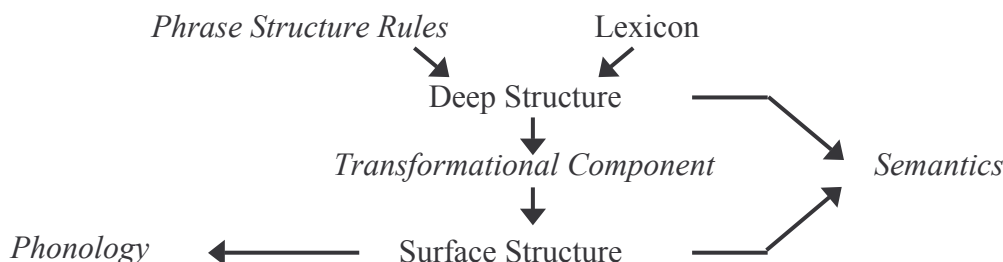
Dans la tradition de la grammaire générative, les mécanismes syntaxiques produisent une structure, appelée forme logique (LF), qui est censée être exploitable par le système conceptuel. La forme logique n'est pas directement observable et ne se confond pas avec la structure grammaticale de l'énoncé tel qu'il est prononcé. Une partie des mécanismes syntaxiques qui engendrent la forme logique est aussi qualifiée de cachée, car ces opérations n'ont pas de conséquences phonologiques (CHOMSKY 1995 [17]). La forme logique est donc l'aboutissement de l'action des mécanismes syntaxiques. Au-delà, d'autres mécanismes, de nature sémantique, sont supposés prendre le relais. Par conséquent, les propriétés prêtées à la forme logique par la théorie syntaxique ont des conséquences sur le fonctionnement du système conceptuel. Celui-ci doit posséder les propriétés suffisantes lui permettant d'intégrer la richesse structurelle de la forme logique.

Considérons par exemple la relation prédicat - argument. La thêta-théorie stipule qu'au niveau de la forme logique, un prédicat doit voir ses arguments satisfaits (CHOMSKY 1981 [16]). Elle proscrit de plus la présence d'arguments surnuméraires. Ces deux contraintes expliquent respectivement le caractère inacceptable de la phrase *lundi, il a dit* et de la phrase *lundi, il a parlé qu'il partait*. Par des règles simples de ce type, la théorie syntaxique guide l'identification des arguments des prédicats. Les marques de cas ou la position relative des syntagmes concourent au même but. Elles permettent de savoir qui est le premier argument du prédicat *frapper* dans chacune des phrases *Paulus Petrum ferit*, *Pierre frappe Paul*, et *Pierre Paul râ mizanad*. Ainsi, la forme logique doit permettre d'exprimer et de distinguer des expressions logiques comme *frapper(x, y)* et *frapper(y, x)*. C'est au niveau de cette forme logique que se révèlent les deux significations d'un syntagme comme le *dépassement de la voiture*, le concept *VOITURE* occupant la première ou la deuxième place du prédicat *dépasser(x, y)*.

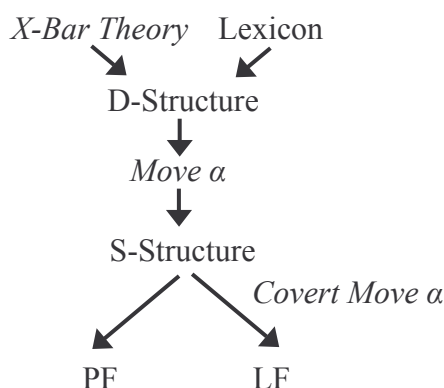
La forme logique est également supposée exprimer la quantification. Par exemple, la phrase *toutes les disquettes ont un virus* possède une structure phonologique non ambiguë qui peut conduire à deux significations. Ces deux significations sont caractérisées, au niveau de la forme logique, par l'ordre dans lequel les quantifications sont positionnées.

La forme logique, définie comme un aboutissement de l'action des mécanismes syntaxiques, nous offre un accès privilégié à la structure du système conceptuel. Il est donc essentiel de savoir en quoi elle consiste et comment elle est engendrée. La réflexion, au sein du courant de la grammaire générative, a été caractérisée par une évolution dans la manière de concevoir l'ordre et la nature des opérations qui engendrent, d'une part la forme logique, et d'autre part la forme phonologique (PF).

La théorie initiale supposait une première phase dans laquelle les mots, guidés par les règles de composition syntaxiques, prennent leur place pour former une première structure, dite profonde, puis une phase de transformations permettant de produire une deuxième structure, dite de surface, qui détermine la forme phonologique de l'énoncé construit. La forme logique, quant à elle, contient des éléments provenant des deux structures profonde et de surface (CHOMSKY 1972 [14]). Le schéma suivant résume cette conception des différents niveaux de composition.



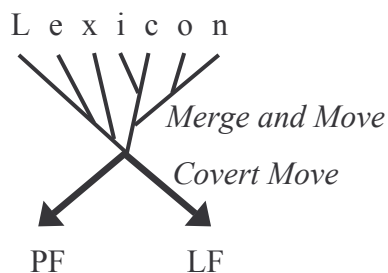
Dans cette perspective, les mécanismes de composition au sein du système conceptuel s’effectuent en deux phases distinctes, en étant soumis, dans chaque phase, aux contraintes fournies par les deux types de structure syntaxique. Cette séparation en deux phases s’est ensuite estompée. Avec l’introduction du mécanisme général du mouvement (*move α*) dans le cadre de la théorie standard de règles syntaxiques (*X-Bar*), l’idée d’une forme logique unique s’est imposée. Dans cette version de la théorie, la structure syntaxique profonde (*D-Structure*) est modifiée par le mécanisme de mouvement pour donner lieu à la structure syntaxique de surface (*S-Structure*), qui assure l’interface avec le système conceptuel (CHOMSKY 1981 [16]). La forme logique est le résultat de l’application, cachée au système phonologique, du même mécanisme général de mouvement à la structure syntaxique de surface. Le schéma suivant résume cette conception du processus de la production de la forme logique.



Dans les versions récentes de la théorie, une phrase est un couple (π, λ) associant une forme phonologique et une forme logique. Elle résulte d’une succession d’insertions d’items lexicaux abstraits conduisant à l’élaboration de π et de λ selon des procédures générales de la langue (CHOMSKY 1995 [17]). Conformément aux principes fondamentaux de la grammaire générative, le lien entre forme phonologique et forme logique n’est pas une correspondance, mais résulte d’un ensemble de dérivations. L’élément nouveau est l’accent mis sur le caractère optimal de ces dérivations.

It seems that a linguistic expression of [the generative procedure] cannot be defined just as a pair (π, λ) formed by a convergent derivation. Rather, its derivation must be *optimal*, satisfying certain natural economy conditions: locality of movement, no “superfluous steps” in derivations, and so on. Less economical computations are blocked even if they converge.
(CHOMSKY 1995 [17] p. 220)

Les niveaux successifs de structure syntaxique disparaissent alors au profit d’une construction directe des deux formes phonologique et logique à partir des entités lexicales. Le schéma suivant résume cette conception.



Ainsi, la forme logique est le fruit d'une construction progressive qui obéit à ses propres lois. Les opérations syntaxiques, comme le mouvement ou la fusion (*merge*) ne dépendent pas du sens, même si elles sont en partie au service du sens. L'opération de fusion confère localement au composé des propriétés calculées à partir de ses composants. Elle réalise notamment la liaison sémantique entre les syntagmes. Le principe de liaison sémantique stipule qu'un syntagme doit partager une variable avec un syntagme qui le domine (DESSALLES 2000 [30]). Ainsi, dans le syntagme le dépassement de la voiture, la fusion provoque l'unification de la variable x dans l'expression *voiture(x)* avec l'une des variables de l'expression *dépasser(y, z)*. L'existence d'un choix concernant la variable à unifier montre que les opérations de fusion, à elles seules, sous-déterminent la construction de la forme logique. Cette sous-détermination reste toutefois limitée, car la structure syntaxique impose de nombreuses contraintes structurelles à la forme logique. Prenons l'exemple de la co-référence. Dans une phrase comme *il dit que Jean est malade*, tout locuteur du français sait que le pronom *il* et le nom propre *Jean* ne peuvent pas être co-référents. Cette contrainte n'existe pas dans la phrase *le fait qu'il soit malade ennuie Jean*, où le pronom *il* et le nom propre *Jean* peuvent désigner la même personne. L'explication proposée pour ce phénomène, qui semble universel, est que la co-référence entre un pronom et une expression référentielle est bloquée lorsque le pronom est en position de c-commander² l'expression référentielle. Le phénomène est spectaculaire, car il peut agir à grande distance. Dans la phrase *il avoue que l'ami de Jean à qui Jacques a emprunté la moto avec laquelle il a renversé le livreur était absent au moment des faits*, nous sommes obligés de comprendre que l'auteur de l'aveu ne peut être ni Jean, ni Jacques, ni l'ami de Jean, ni le livreur (en revanche, l'auteur de l'accident et l'auteur de l'aveu peuvent être la même personne). Nous pouvons en conclure que la syntaxe contraint partiellement certains aspects de la forme logique, comme la co-référence, ce qui a pour effet de rendre la communication plus facile.

Il est légitime de se demander où s'arrêtent les spécifications sémantiques qui résultent automatiquement de la syntaxe. Par exemple, les différentes propositions qui ont été faites pour caractériser un système conceptuel qualitativement distinct du langage, mais capable de s'interfacer avec lui, ne peuvent éviter d'y inclure des aspects propres à la syntaxe du langage.

[...] "conceptual structure" contains not only semantic information: it also expresses thematic relations (concepts like Agent, Goal, Source, and the like) and structural relations (that is, hierarchical relations such as "argument of", "head of", "modifier of", and the like). Thematic relations represent, of course, the core materials from which syntax is made. What conceptual structure does not express [...] includes the linear ordering of words into sentences and the morphophonemic shapes that linguistic concepts must take on if they are going to be used communicatively between two individuals. (BICKERTON 1995 [6] p. 114)

² La relation de c-commande est une relation structurelle au sein de l'arbre syntaxique. Le nœud A c-commande le nœud B si le premier nœud ramifié qui domine A domine B, sans que A et B soient eux-mêmes dans une relation de dominance.

Si l'on pousse ce constat jusqu'à sa limite, il n'y a plus aucune raison de supposer l'existence de structures et de mécanismes propres au système conceptuel qui seraient différents de ceux qui nous permettent de parler.

The same biological mechanisms, developed by and for language, run both distinctively human thought and language. (BICKERTON 1995 [6] p. 105)

Si l'on accepte ce point de vue, il n'est plus question d'interface, et les calculs sémantiques doivent être conçus sur le modèle des transformations syntaxiques. Le sens serait ainsi entièrement exprimé par la forme logique. Une telle position nécessite que la forme logique soit engendrée par des transformations qui, tout en restant par nature syntaxiques, dépassent le stade de ce qui est observable. Par exemple, le mécanisme caché de montée de quantification, qui permet d'expliquer le fait que la phrase *toutes les disquettes ont un virus* puisse avoir deux sens logiquement différents, est calqué sur le mouvement des pronoms relatifs³. En extrapolant, on peut imaginer que tous les calculs conceptuels résultent des mécanismes à l'œuvre dans la syntaxe.

Contre ce point de vue, deux arguments peuvent être invoqués. L'un est que les opérations inférentielles qui portent sur les concepts (inférences déductives, généralisations inductives, détections d'incompatibilités logiques, *et cætera*) ne peuvent pas s'exprimer naturellement avec les outils offerts par le système syntaxique. Il est difficile d'envisager un modèle dans lequel les mécanismes et les structures propres à la langue offrent les moyens d'inférer la phrase *Y est mort à partir de la phrase X a tué Y*. Pour traduire les calculs de nature symbolique qui sous-tendent les processus de raisonnement, il a fallu développer des formalismes logiques, qui entretiennent peu de rapports avec les mécanismes grammaticaux des langues. Inversement, les relations syntaxiques sont, pour l'essentiel, insensibles aux dépendances structurelles entre concepts qui ont été postulées dans le précédent chapitre pour expliquer la formation des inférences. Le plus souvent, seule la catégorie syntaxique importe au niveau grammatical, et cette catégorie est en partie décorrélée des propriétés sémantiques du mot. Il est donc peu plausible de postuler d'emblée que les processus par lesquels nous combinons les concepts sont indiscernables des processus grammaticaux.

Certes, la représentation de type syntaxique s'est vue attribuer un rôle nécessaire pour rendre les représentations conceptuelles accessibles aux autres représentations mentales conscientes, notamment aux attitudes propositionnelles (CARRUTHERS 1996 [11]). En d'autres termes, nous penserions à l'aide de phrases. La représentation syntaxique serait donc le substrat sur lequel opéreraient certains processus conceptuels. L'occurrence d'une attitude propositionnelle à propos d'un contenu de pensée nécessiterait l'activation d'une représentation de type linguistique, autrement dit une phrase exprimant ce contenu de pensée. Si tel est le cas, les structures propres au langage naturel, au premier rang desquelles les structures syntaxiques, doivent jouer un rôle fondamental dans la genèse de la pensée en général, et du raisonnement en particulier, bien au-delà du simple décodage des énoncés entendus. Dans l'état actuel des réflexions sur l'organisation du système conceptuel et de ses mécanismes inférentiels, cependant, les opérations qui sont postulées restent constituées de

³ Dans le syntagme *le livre_(i) dont_(i) elle rêvait* $t_{(i)}$, la variable représentée par le pronom *dont* est supposée provenir de la position complément du verbe *rêver*, par un mouvement qui laisse subsister une trace $t_{(i)}$ co-indexée avec le pronom *dont*. En anglais, le mouvement peut laisser la préposition en place, comme dans le syntagme *the book_(i) she was dreaming of* $t_{(i)}$. La présence de mouvements, et des traces associées, est révélée en français par l'accord du participe passé ou par les clitiques pronominaux dans les interrogatives. De même, les deux formes logiques de la phrase *toutes les disquettes ont un virus* s'obtiennent par le mouvement caché, en tête de structure, des deux quantificateurs, donnant soit [*toutes les disquettes₁ [un virus₂ [t_1 infecte t_2]]*], soit [*un virus₂ [toutes les disquettes₁ [t_1 infecte t_2]]*] suivant l'ordre dans lequel les deux quantificateurs ont migré.

schémas de calculs logiques, de type résolution d'incompatibilité, déduction ou autre. Si les structures syntaxiques ont un rôle dans ces opérations, il ne peut être que partiel. En d'autres termes, nous sommes loin d'une réduction du conceptuel au grammatical. Même si l'on accepte l'importance du rôle de la composition syntaxique pour la genèse de la pensée consciente, la composition du sens doit conserver une partie de son autonomie.

L'autre argument contre une telle réduction est lié au fait que le système conceptuel n'est pas soumis à la même contrainte de sérialité que le canal verbal : les différents concepts d'un énoncé sont simultanément activés pour que l'énoncé soit compris. Or, bon nombre de mécanismes syntaxiques tirent leur justification de la sérialité du canal verbal. Prenons le cas du mouvement. Le mouvement du mot *dont* dans la phrase *le livre_(i) dont_(i) elle rêvait t_(i)*, laissant la trace *t_(i)*, se justifie par la nécessité d'amener le complément du verbe *rêver* sous la dépendance immédiate du syntagme *le livre*, afin que la liaison sémantique puisse opérer. Ce genre de mouvement ne se justifie pas dans le fonctionnement interne du système conceptuel, dès lors que des structures non linéaires comme des arbres ou des graphes peuvent être postulées. La théorie des cas offre un autre exemple de mécanismes motivés seulement au niveau syntaxique. Dans la phrase *Jean semble malade*, le syntagme *Jean* est sujet du verbe *sembler* sans en être argument au niveau sémantique. L'interprétation classique de ce phénomène est que cette phrase provient de la transformation d'une structure comme *il semble Jean malade*. Cette transformation est rendue nécessaire par le fait que l'adjectif *malade* ne peut pas donner de cas à son sujet⁴. Le nom propre *Jean*, dont le rôle sémantique est donné par l'adjectif *malade*, migre en position de sujet du verbe *sembler* afin de recevoir un cas. Nous avons donc affaire à une mécanique possédant ses lois propres, comme l'obligation pour un syntagme nominal de recevoir un cas, même dans une position dépourvue de rôle sémantique. Cette mécanique casuelle serait sans intérêt dans le fonctionnement interne du système conceptuel, puisque les arguments *y* sont connus. Comme le mouvement, la contrainte d'attribution de cas aux syntagmes nominaux se justifie au niveau syntaxique, non au niveau sémantique. La raison d'être du système des cas se justifie par la nécessité d'identifier, pour l'interlocuteur, les arguments d'une relation dans une suite linéaire de mots.

On pourrait passer en revue les différents mécanismes syntaxiques pour montrer que leur existence ne se justifie souvent pas sur le plan des opérations purement conceptuelles. Pour ce genre de raisons, bon nombre d'auteurs préfèrent considérer que syntaxe et sémantique correspondent à deux niveaux distincts de la compétence langagière, et que l'on se doit de définir une interface entre les deux systèmes. Les représentations formelles du sens, que nous évoquons dans la prochaine section, ont été conçues au départ comme un système totalement indépendant de l'organisation du langage naturel. Nous verrons que même dans ce cadre formel, la question de l'indépendance entre syntaxe et sémantique peut se poser.

5.2. La représentation formelle du sens

La combinaison grammaticale présente un caractère manifeste, directement observable dans l'ordonnement des mots et les transformations morphologiques. Une structure grammaticale peut être proposée à partir des unités lexicales de l'énoncé, en utilisant les règles de grammaire de la langue. La combinaison conceptuelle ne repose pas sur une base aussi tangible. Comment inférer les propriétés d'un système que l'on ne peut pas observer directement ? On peut penser que l'une des propriétés fondamentales que le système

⁴ La phrase *il croit Jean malade* n'est pas une exception à cette contrainte. Ici, le nom propre *Jean* a migré en position de complément du verbe *croire* où il peut recevoir un cas accusatif, comme le montre la paraphrase suivante : *il le croit malade*.

conceptuel se doit de posséder est la capacité d'exprimer des états de choses de manière non ambiguë, contrairement au langage qui semble bien imparfait avec ses ambiguïtés, son flou et sa sensibilité au contexte.

Une approche fréquemment adoptée consiste ainsi à prendre le problème de la composition sémantique "par le haut". Dans cette approche, un énoncé possède un sens si cet énoncé représente une situation actuelle ou possible, rencontrée ou imaginée, dans le "monde". De plus, on peut dire que le sens d'un énoncé permet, lorsqu'il est compris, d'effectuer un certain nombre d'inférences. En d'autres termes, les représentations sémantiques attribuées aux énoncés du langage sont supposées s'interpréter dans le "monde" et servir aux raisonnements qui peuvent être effectués à propos du "monde". Or, la réflexion philosophique et scientifique a produit un système de représentations dont c'est précisément la fonction : le formalisme logique.

Le formalisme logique est, en lui-même, un langage, pour lequel on peut distinguer une syntaxe et une sémantique. Dans ce langage, la valeur sémantique d'un énoncé s'obtient de manière récursive à partir de l'interprétation des expressions constituant l'énoncé. Cette évaluation s'effectue par rapport à un domaine d'interprétation, en deux temps. Dans un premier temps, les symboles non logiques reçoivent une interprétation dans le domaine. Par exemple, dans la logique du premier ordre, une constante du langage est associée à un élément précis du domaine, et un symbole de prédicat est associé à une relation définie entre certains éléments du domaine. Dans un deuxième temps, l'évaluation consiste à calculer la valeur sémantique des expressions interprétées⁵, puis à propager ces valeurs à tout l'énoncé en tenant compte des opérateurs logiques⁶. Ce type d'évaluation directe est souvent fastidieux. La raison d'être de la logique est de permettre d'éviter l'évaluation directe des énoncés, grâce à des mécanismes de dérivation qui conduisent à déduire qu'un énoncé est satisfait dans un domaine (ou peut être satisfait pour certaines interprétations de ses variables) dès que certains autres énoncés le sont. Ainsi, la règle du *modus ponens* permet de déduire la satisfaction d'énoncés partiels à partir de celle de l'énoncé englobant⁷. Certains formalismes peuvent être dotés de mécanismes de dérivation⁸ corrects et complets qui permettent de produire la liste de tous les énoncés dont la satisfaction peut être déduite de celle d'un énoncé donné.

Les qualités des formalismes logiques suggèrent une approche de la sémantique du langage naturel par laquelle on cherche à éliminer l'ambiguïté, le flou et le caractère contextuel des énoncés langagiers. Cette approche, qui est celle de la philosophie du langage, consiste à transcrire ce qui est exprimé par le langage dans un autre langage, celui de la logique. C'est au moment de cette transcription que les approximations du langage naturel sont levées. Par l'emploi d'un langage formel idéalisé, on parvient à caractériser de manière parfaite et non ambiguë les propriétés intentionnelles et inférentielles de l'énoncé de départ. En d'autres termes, un énoncé du langage naturel aurait un sens si et seulement si on peut lui associer une représentation logique et que cette représentation logique peut être interprétée dans un domaine donné. De même, une chaîne de raisonnement à partir d'un énoncé du langage naturel est fiable si et seulement si sa transcription en langage formel procède selon les étapes d'une dérivation logiquement bien formée.

⁵ Par exemple, une expression comme $p(a, b, c)$ sera évaluée comme satisfaite si le triplet d'éléments du domaine associé à la configuration de symboles (a, b, c) se retrouve dans le graphe de la relation ternaire associée au prédicat p .

⁶ Par exemple, une expression comme $p(a, b, c) \wedge q(d)$ sera évalué comme satisfaite si les deux expressions $p(a, b, c)$ et $q(d)$ sont, chacune, évaluées comme satisfaites.

⁷ Le *modus ponens* est la règle déductive qui conclut à la satisfaction d'une expression q à partir de celles des expressions p et $(p \supset q)$.

⁸ Ces mécanismes incluent les systèmes axiomatiques et les méthodes de preuve, par exemple les méthodes de preuve par réfutation.

Le fait d'adopter une représentation logique permet de franchir une étape supplémentaire, en considérant qu'un énoncé a un sens s'il peut être évalué. Comme, par hypothèse, un énoncé sensé du langage naturel peut être transcrit dans une expression logique bien formée, et que toute expression logique bien formée peut être évalué par rapport à un domaine, un énoncé sensé du langage naturel doit pouvoir recevoir une valeur de vérité⁹. Plus exactement, si l'énoncé comporte des variables, son sens peut être assimilé à ses conditions de vérité. Selon cette perspective, le sens d'une phrase devient l'ensemble des états possibles du monde dans lesquels la phrase est vraie. Ainsi, le sens de la phrase la voiture de Jean est grise est l'ensemble des états possibles du monde dans lesquels l'individu indiqué par le prénom Jean possède effectivement une voiture grise. On parvient ainsi à donner un critère d'acceptabilité sémantique permettant de distinguer les énoncés sensés du langage, en transformant la question du sens en une question épistémique.

L'idée fondamentale de la sémantique formelle consiste à remplacer la forme imparfaite de l'énoncé du langage naturel par sa transcription dans un langage formel. La forme logique ainsi obtenue a un statut radicalement différent de celle des approches générativistes du langage. Ici, la structure de la forme logique n'est pas nécessairement parallèle à la structure syntaxique. En particulier, la constituance syntaxique peut être perdue. Par exemple, si l'on traduit l'énoncé Jean casse un verre, dans un système de logique du premier ordre, par l'expression $(\exists x) \text{ verre}(x) \ \& \ \text{casser}(\text{Jean}, x)$, on constate que le prédicat *verre* n'est pas un constituant du prédicat *casser*, alors que le nom *verre* est un constituant du syntagme verbal dont la tête est le verbe *casser*. Par ailleurs, les symboles et les structures utilisés dans la forme logique ne correspondent pas nécessairement à des entités psychologiques. Ce sont, avant tout, les éléments d'un système de représentation. Dans la section précédente, la forme logique possédait un statut cognitif, dans la mesure elle était issue d'une forme syntaxique profonde par une suite supposée d'opérations mentales. Elle était censée conserver tous les aspects de la structure syntaxique utiles à l'interprétation sémantique. La forme logique considérée dans la présente section résulte d'une transcription sans prétention cognitive. L'objectif est de représenter le contenu de l'énoncé du langage naturel dans un langage artificiel où les ambiguïtés, le flou et les aspects contextuels peuvent être éliminés. De cette manière, la transcription formelle est un moyen de révéler le sens. La diversité des formalismes logiques a pour conséquence qu'un énoncé donné pourra être associé à plusieurs représentations formelles. Il serait inutile de dire que l'une de ces représentations est plus "cognitive" que les autres.

Si l'on prend à la lettre cette restriction sur le caractère non cognitif de la sémantique formelle, on parvient à un certain scepticisme. La forme logique ne serait qu'un outil commode pour exprimer le sens de manière précise en s'affranchissant des imperfections du langage naturel. Elle ne serait d'aucun secours pour modéliser les mécanismes sémantiques à l'œuvre dans le langage naturel. Ce point de vue est, d'ailleurs, celui de la tradition philosophique qui est à l'origine de la sémantique formelle. La forme logique est incontournable pour comprendre la proposition contenue dans l'énoncé langagier et l'état de fait qu'elle exprime, ce qui n'est pas le cas de la forme grammaticale. C'est donc la forme logique qui doit être retenue pour comprendre la rationalité de la pensée scientifique. Une telle attitude conduit naturellement à accepter un divorce entre la forme grammaticale et la forme logique, et à adopter une vision pessimiste concernant l'espoir de modéliser fidèlement toutes les nuances de sens exprimées par les mots et les grammaires des langues naturelles. Pour certains auteurs, cependant, le défi mérite d'être relevé, car l'adéquation de la forme logique au langage ne peut qu'améliorer son pouvoir d'exprimer la pensée humaine.

⁹ Dans cette tradition philosophique, le mot "vérité" est employé dans un sens objectif. Il n'a donc pas le statut d'attitude propositionnelle qu'on lui donne dans certaines approches cognitives du langage, ni le statut de vérification opérationnelle qu'il peut avoir en intelligence artificielle.

I reject the contention that an important theoretical difference exists between formal and natural languages. (MONTAGUE 1974 [76] p. 188)

Le projet visant à adopter une vision formelle du langage doit d'emblée aborder deux problèmes. Le premier est de concevoir une interprétation qui permette de reproduire la richesse de la sémantique du langage naturel. Le deuxième est d'organiser une correspondance systématique entre la forme grammaticale et la forme logique.

Interprétation

La première difficulté à résoudre pour rapprocher la forme logique du langage naturel réside dans le fait que l'interprétation sémantique est supposée objective, indépendante du point de vue du locuteur, ce qui la rend souvent trop rigide. Le système logique de base utilisé pour traduire les énoncés du langage naturel est la logique du premier ordre. Dans ce système, les phrases Socrate est mortel et tout homme est mortel peuvent être traduites, respectivement, par des formules comme $M(s)$ et $\forall x H(x) \supset M(x)$. Leur interprétation peut s'effectuer dans le domaine des entités "réelles". Dans ce cadre, la constante s est interprétée comme un individu particulier, connu sous le nom de Socrate. Les symboles de prédicat H et M correspondent à des ensembles d'individus, l'ensemble des humains et l'ensemble des mortels. L'évaluation du premier énoncé s'opère par un test d'appartenance, celle du deuxième par un test d'inclusion entre les deux ensembles concernés. Ces procédures sémantiques semblent opérer correctement dans un monde supposé comme objectif. Or, le langage peut exprimer des états de faits qui, bien qu'intuitivement compréhensibles, ne semblent pas entrer dans le schéma interprétatif esquissé ci-dessus. Ainsi, l'énoncé Pégase est blanc pose des problèmes, dans la mesure où le mot Pégase ne renvoie pas à une entité du monde "réel". Le domaine d'interprétation constitué par les objets "réels" ne permet pas de donner un sens à cet énoncé. L'interprétation de cet énoncé est possible à condition de pouvoir considérer un deuxième domaine d'interprétation inhabituel, incluant les entités mythologiques. De cette manière, la phrase Pégase n'existe pas sera interprétée par le fait qu'il existe un domaine, celui des entités "réelles" où aucun élément n'est associé au mot Pégase. Ainsi, en considérant plusieurs domaines d'interprétation, on parvient à donner un sens à une classe beaucoup plus large d'énoncés langagiers.

La rigidité de l'interprétation logique la rend impropre à exprimer la combinaison des significations véhiculées par le langage. Par exemple, lorsque l'adjectif ancien modifie un autre mot, le résultat ne peut pas être représenté par une conjonction logique. Le sens du syntagme ancien soldat n'est pas rendu par la formule $ancien(x) \ \& \ soldat(x)$. Un ancien soldat n'est pas un objet appartenant à l'intersection de l'ensemble des objets anciens et des soldats. Ce problème peut néanmoins recevoir une solution. Il suffit d'accepter, non seulement un changement de point de référence, comme dans l'exemple du mot Pégase, mais la possibilité de considérer tout un ensemble de domaines, chacun constituant un point de référence daté. Ainsi, on recherchera, dans l'ensemble des domaines d'interprétation du passé, ceux qui incluent un individu ayant la propriété d'être soldat, de manière à ce que le syntagme ancien soldat soit avéré dans le domaine correspondant au présent. Le fait de pouvoir adopter plusieurs points de référence augmente le pouvoir d'expression de la logique. La technique consiste à enrichir la structure d'interprétation. Au lieu de s'en tenir à un seul domaine, on considère une famille de domaines, dont chacun constitue un point de référence.

To interpret a pragmatic language L we must specify several things. In the first place, we must determine the set of all [...] complexes of relevant aspects of possible contexts of use; we may call such complexes [...] *points of reference*. [...] In the second place, we should have to specify, for each point of reference i , the set A_i of objects present or existing with respect to i . [...] In the third place, we should have to specify the meaning or *intension* of each predicate and individual constant of L . To do this for a constant c , we should have to determine, for each point of reference i , the denotation or extension of c with respect to i .

(MONTAGUE 1974 [76] p. 121)

Pour augmenter les possibilités de transcription du langage naturel en langage formel, la stratégie consiste, non à complexifier le langage formel, mais à augmenter ses possibilités d'interprétation. La conséquence est qu'un terme n'a plus d'extension en absolu. Son extension dépend du point de référence. L'analyse sémantique classique était cantonnée aux phrases ayant une structure de prédication universelle et atemporelle. Dans le cadre de la sémantique intensionnelle, l'interprétation des expressions logiques ne consiste plus à associer chaque symbole à son extension, puisque l'extension dépend du point de référence. L'interprétation passe par une fonction, appelée intension, qui doit être calculée pour déterminer l'extension. Par exemple, la prise en compte du temps dans la transcription peut être effectuée sans introduire de paramètre temporel explicite dans l'expression logique. Un énoncé comme *Jean est parti* continue à être traduit par un prédicat atemporel *partir(jean)*. Le renvoi de l'événement dans le passé est rendu au niveau de l'interprétation. Pour cela, on considère une famille de domaines indicés par le temps. La modalité temporelle de l'énoncé se traduit par le fait que le prédicat *partir(jean)* était avéré dans l'un de ces mondes passés. Toutefois, le sens de l'énoncé *Jean est parti* ne se limite pas au fait qu'il existe un monde dans lequel le départ de Jean est avéré. La phrase exprime une antériorité par rapport à l'instant d'énonciation. De telles contraintes peuvent facilement être traduites par l'emploi d'opérateurs modaux. Dans le cadre d'une logique temporelle, l'opérateur P , dans l'expression P *partir(jean)*, contraint l'évaluation à rechercher un monde passé dans lequel le prédicat s'avère.

The fourth thing we must provide is an interpretation of operators of L . To do this we associate with each operator of L a relation between points of reference and sets of points of reference.

(MONTAGUE 1974 [76] p. 122)

Dans le cas de la logique temporelle, l'opérateur P associe au monde courant l'ensemble des mondes qui lui sont antérieurs. L'emploi d'opérateurs modaux agissant sur la structure d'interprétation permet de faire figurer des contraintes sémantiques dans la syntaxe du langage logique. Cette méthode augmente considérablement la capacité du langage logique à représenter le sens des énoncés langagiers. Les opérateurs modaux associés à des mondes censés représenter les croyances de divers individus permettent de créer des logiques épistémiques capables de différencier le sens d'un énoncé en tenant compte des locuteurs. Là où la logique classique a accès à une capacité de vérification omnisciente pour évaluer les énoncés, les logiques épistémiques peuvent représenter toute la variété des croyances partielles.

De telles techniques permettent d'aborder le problème du caractère contingent de certaines connaissances nécessaires à la compréhension des énoncés langagiers. Il est habituel, en linguistique et en philosophie du langage, de distinguer les aspects qui dépendent du contexte de l'énonciation et ceux qui ne dépendent que d'une connaissance générale de la

langue et des concepts. La prise en compte du contexte est parfois appelée pragmatique¹⁰. Elle est nécessaire dès que l'on est en présence, par exemple, de termes indexicaux comme les démonstratifs.

It seemed to me desirable that pragmatics should at least initially follow the lead of semantics – or its modern version, model theory – which is primarily concerned with the notions of truth and satisfaction (in a model, or under an interpretation). Pragmatics, then, should employ similar notions, though here we should speak about truth and satisfaction with respect not only to an interpretation but also to a context of use.

(MONTAGUE 1974 [76] p. 120)

Une idée souvent admise est qu'il est possible d'opérer une distinction de principe permettant d'isoler ce qui, dans l'ensemble des connaissances utiles à l'interprétation d'un énoncé donné, dépend du contexte dans lequel l'énoncé est émis.

The interpretation of an utterance depends on the meaning of the expressions used and on various additional facts about the utterance. The truth of an utterance depends, in turn, on whether its interpretation fits the facts. [...] Using situations, one can separate cleanly the two parts of reality that intervene between meaning and truth, the facts about the utterance, and the facts about the *described situation*, i.e., those relevant to the truth and falsity of the utterance as interpreted.

(BARWISE & PERRY 1983 [4] p. 6)

L'interprétation d'un énoncé est ainsi censée se dérouler en deux temps. Dans un premier temps, le contexte de l'énonciation est élaboré en précisant la location spatio-temporelle et les entités auxquelles certaines unités lexicales comme les pronoms ou les démonstratifs font référence. Ensuite intervient la traduction des termes génériques.

Consider the sentence I AM SITTING. Its meaning is, roughly, a relation that holds between an utterance u and a situation e [...] expressed as follows:

u [[I AM SITTING]] e
iff

There is a location l and an individual a , such that
in u : at l : speaks, a ; yes
in e : at l : sits, a ; yes

[...] Now suppose we fix certain facts about the utterance. Fred is the speaker, and the location is a particular spacetime region l . Then *any* situation that has Fred sitting at l will be an interpretation of the utterance.

(BARWISE & PERRY 1983 [4] p. 19)

Dans cet exemple, le sens de l'énoncé est éclaté dans les différents éléments de la traduction. Le pronom "I" est interprété par l'unification de la variable a dans deux situations, caractérisées par les expressions "speaks" et "sits". Le fait que le temps de l'énoncé soit au présent a pour conséquence l'unification des localisations sous le même nom de variable l . Cette variable peut également contribuer à faire le lien entre cet énoncé et le discours dans lequel il se trouve.

Pour de nombreux auteurs, l'amélioration du pouvoir d'interprétation de la logique pourra permettre, à terme, de représenter le sens des énoncés langagiers de manière complète et non ambiguë, en tenant compte des conditions d'énonciation. Ce programme de recherche a produit des avancées indéniables dans l'élaboration d'un langage logique aux capacités de

¹⁰ Dans une autre tradition, issue des travaux de H. P. Grice, le terme pragmatique désigne un authentique niveau de la compétence langagière, celui des inférences et de la prise de parole pertinente.

représentation étendues. Cependant, les moyens mis en œuvre pour représenter le sens peuvent être critiqués sur le plan épistémologique et cognitif.

La première critique concerne le statut des objets utilisés dans la traduction logique. Comme dans le cas des formes profondes utilisées par les syntacticiens, certains éléments n'ont pas de contrepartie explicite dans l'énoncé. La variable *l* de l'exemple précédent n'a aucune réalité phonologique. Cependant, contrairement à l'approche évoquée dans la première section, le souci de rendre la forme logique minimale par rapport à ce qui est exprimé verbalement n'est pas central. La préoccupation est plutôt de rendre le pouvoir d'expression de la logique maximal. D'où l'introduction de variables, d'opérateurs et de structures d'interprétation complexes comprenant souvent une infinité de mondes possibles. Comme l'expression langagière sous-détermine les structures postulées par la sémantique formelle, le statut cognitif de ces structures n'est pas clair. Dans l'exemple précédent, la séparation en deux situations reliées par un partage de variable est due aux choix du formalisme. Ce qui est observable dans l'énoncé *je suis assis* n'offre pas d'indication claire selon laquelle l'auditeur d'un tel énoncé conçoit l'état décrit et l'acte d'énonciation comme deux situations traitées au même niveau. En d'autres termes, les modèles de la sémantique formelle sont conçus pour leur efficacité, non pour représenter les processus cognitifs à l'œuvre.

Une deuxième critique porte sur la possibilité, énoncée par de nombreux auteurs, d'isoler dans l'interprétation sémantique des énoncés ce qui relève du contexte immédiat de ce qui n'en dépend pas. Certaines caractéristiques liées à l'interprétation sont clairement contextuelles. C'est le cas de la localisation, de l'instant d'énonciation, de la référence des pronoms et des déictiques. Peut-on assurer que l'interprétation de certains mots, au moins, peut se faire indépendamment du contexte ? Dans l'expression *un grand vin*, l'adjectif *grand* modifie une dimension de qualité, alors que dans l'expression *un grand arbre*, la qualification opère sur la taille. L'action sémantique de l'adjectif *grand* dépend donc du contexte des autres mots de la phrase. Le mot *arbre*, prononcé par un enfant ou un botaniste expert, en Amazonie ou en Islande, recouvrera des objets différents. Il est difficile de trouver un élément lexical exempt de polysémie. Un mot comme *livre* peut désigner l'objet physique, l'œuvre intellectuelle, l'opération commerciale, l'ensemble des exemplaires, *et cætera*. La polysémie concerne également les mots grammaticaux. Dans les expressions *un livre sur une table*, *un tableau sur le mur*, *un insecte sur le plafond*, la relation spatiale désignée par le mot *sur* s'obtient par un calcul qui dépend des entités mises en relation. En résumé, il semble que le contexte sémantique puisse intervenir dans le sens de tous les éléments lexicaux. Aucun principe formel permettant de distinguer des cas d'immunité au contexte n'a été proposé, hormis ceux qui peuvent être déduits de la structure syntaxique, comme les contraintes de co-référence mentionnées dans la section précédente. Si le contexte est toujours susceptible d'intervenir, quelle sont les lois qui permettent de prédire son influence ? En l'absence d'influence du contexte, on peut espérer calculer le sens des énoncés en utilisant un dictionnaire élaboré contenant des schémas sémantiques pour chaque entrée lexicale. L'omniprésence du contexte nécessite des calculs additionnels pour lesquels la sémantique formelle n'offre pas de description générale et systématique. Il est toujours possible, au cas par cas, de postuler l'existence d'un point de référence ou d'un trait permettant de lever telle ou telle ambiguïté. Certes, la prise en compte du temps, de la localisation et du locuteur dans la structure d'interprétation contribuent significativement à préciser les situations dans lesquelles un énoncé donné est interprétable. Cependant, on ne peut exclure *a priori* aucun paramètre comme non pertinent pour l'interprétation. Ainsi, la température ambiante ou le degré d'hygrométrie peuvent modifier l'ensemble des situations auxquelles un énoncé s'applique (par exemple l'énoncé *l'eau est froide* peut désigner des situations différentes en été et en hiver). Faut-il prévoir une modalité de température, et donc multiplier l'univers des mondes

possibles, ou des situations, pour distinguer tous les cas qui ne diffèrent que par la température ? Le danger est de voir l'ensemble des dimensions concevables par l'esprit humain venir encombrer la structure d'interprétation.

Une troisième critique concerne précisément la structure d'interprétation. Le principe de la théorie des modèles consiste à faire reposer le sens sur un jugement épistémique objectif. Ainsi, le sens de l'énoncé Jean est assis est constitué par l'ensemble des situations, choisies dans les mondes accessibles de la structure d'interprétation, qui rendent l'expression *assis*(Jean) vraie. Si l'on renonce à la doctrine platonicienne intégrale, selon laquelle de telles relations seraient extérieures aux esprits tout en leur étant accessibles¹¹, le problème vient de ce que l'on renvoie toute la richesse sémantique imaginée par les sémanticiens formels dans les capacités perceptives des individus. Autrement dit, si l'on cherche à transformer les théories formelles de la sémantique en théories cognitives, les capacités perceptives humaines deviennent d'une complexité peu plausible. Pour comprendre le moindre énoncé, l'auditeur devrait effectuer des calculs sur des ensembles infinis de mondes, de manière à dresser la liste exhaustive des situations dans lesquelles l'énoncé est avéré. Il faut, de plus, que l'auditeur dispose d'une panoplie de relations non ambiguës portant sur ces mondes. Si l'on accepte une telle hypothèse, il est légitime de s'interroger sur les raisons qui font que les êtres humains utilisent un langage si ambigu et polysémique, sachant qu'ils ont à leur disposition un catalogue de relations parfaites¹².

Une théorie normative du sens, qui viserait à déterminer le sens des énoncés du langage de manière objective, indépendamment des locuteurs, pourrait en principe se satisfaire d'un système d'interprétation en forme d'oracle, quelle qu'en soit la complexité. Une telle théorie aura peu à nous apprendre sur le comportement langagier humain et sera incapable de le prédire, car l'oracle censé dispenser la vérité est inaccessible au scientifique et encore plus aux locuteurs. La situation est différente lorsque l'on utilise les formalismes logiques pour modéliser le raisonnement. Dans ce cas, on obtient de véritables mécanismes auxquels on peut assigner une contrepartie cognitive. Par exemple, il n'est pas choquant d'imaginer un mécanisme cognitif capable de transformer une incompatibilité en disjonction. En revanche, il est difficile d'imaginer un mécanisme cognitif dont la fonction serait de produire des conditions de vérité après avoir parcouru une infinité de domaines constitués d'entités et de relations qui, pour la plupart, sont extérieures au sujet et indépendantes de son fonctionnement cognitif.

Correspondance

Les critiques précédentes, qui portaient sur la nature de l'interprétation, ne sont qu'un aspect des difficultés présentées par le projet de la sémantique formelle. Une autre source de complications est liée au fait que l'on postule l'existence de deux formes concurrentes d'expression. Si une même situation, observée dans le monde, peut s'exprimer spontanément par le langage naturel et scientifiquement sous la forme d'un langage formel, il faut préciser quel lien existe entre ces deux formes d'expression. Or, dans la plupart des formalismes, ce lien n'est pas étroit. Par exemple, dans la traduction $\forall x \text{ homme}(x) \supset \text{mortel}(x)$ de la

¹¹ Une telle doctrine est incompatible avec une approche matérialiste des problèmes cognitifs. L'idée de confier à la science, par exemple aux sciences physiques, la charge de vérifier la vérité des énoncés est illusoire pour deux raisons. D'une part, la connaissance scientifique est toujours provisoire, ce qui rend la notion de vérité également provisoire. D'autre part, la plupart de ces tests de validité scientifique sont inaccessibles aux locuteurs qui, pourtant, parviennent à communiquer.

¹² L'idée qui consiste à introduire des logiques floues ou multi-valuées ne résout pas le problème. La difficulté signalée ici réside, non dans le caractère graduel et flou des catégories désignées par les mots du langage, mais dans l'ambiguïté de ces mots. Or, l'interprétation, dans ces logiques, sans être binaire, doit néanmoins être non ambiguë.

phrase tout homme est mortel, le groupe nominal tout homme se retrouve éclaté entre le quantificateur $\forall x$ et le prédicat $\text{homme}(x)$. Inversement, le prédicat $\text{mortel}(x)$ se retrouve sous la portée du quantificateur, au même niveau que le prédicat $\text{homme}(x)$, alors que l'adjectif mortel est extérieur au syntagme nominal tout homme dans la phrase. Le passage de la représentation grammaticale à la représentation formelle apparaît ainsi davantage comme une adaptation que comme une correspondance. Or, un calcul formel du sens est censé expliquer de manière exhaustive la transformation qui permet de passer de l'expression grammaticale à l'expression formelle. Si la traduction du langage sous une forme logique donne lieu à interprétation, alors l'adoption d'un système formel peut certes aider à la détermination des valeurs de vérité et à la levée des ambiguïtés, mais elle ne peut pas prétendre reproduire le calcul du sens.

Une manière de s'affranchir de cette difficulté est de postuler une correspondance stricte entre la forme grammaticale de l'énoncé et l'expression formelle. Il s'agit là d'une exigence forte, qui rend le choix du système formel particulièrement crucial. Dans ce cas, la compositionnalité devient une préoccupation centrale, puisque la structure des expressions du langage formel doivent être déduites de la syntaxe du langage naturel. Avant de profiter de la précision que permet le langage formel, il faut passer par une phase de transformation algorithmique qui produit l'expression formelle à partir de l'expression verbale. Une manière simple de formaliser cette idée est d'imaginer l'existence d'une correspondance de structures : la syntaxe présente une structure algébrique, la sémantique se caractérise, elle aussi, par une deuxième structure algébrique, et il existe un homomorphisme¹³ qui projette les éléments de la première dans la deuxième (MONTAGUE 1974 [76]).

Le mérite de cette idée est d'astreindre le système sémantique à respecter la compositionnalité grammaticale. Techniquement, on part d'un lexique organisé selon les catégories grammaticales (nom commun, verbe intransitif, adverbe, *et cætera*). On suppose donnée l'interprétation sémantique de chaque terme de ce lexique dans la structure d'interprétation considérée. Une première étape consiste à trouver des règles de compositions syntaxiques qui, à partir du lexique, engendrent de manière récursive des expressions syntaxiquement bien formées (syntagmes nominaux, phrases, syntagmes adverbiaux, *et cætera*). Une deuxième étape consiste à appliquer un ensemble d'opérations de composition sémantique qui engendrent l'interprétation d'une structure syntaxique en fonction de l'interprétation de ses composants. C'est à ce point que s'opère la mise en œuvre de l'homomorphisme. À l'application d'une règle syntaxique S_i à l'ensemble d'expressions langagières $\{\delta_j\}$ correspond, sur le plan sémantique, l'application d'une opération F_i à l'ensemble de significations $\{g_j\}$.

Corresponding to the syntactic rules S_2, \dots, S_{17} (and to given sets A and I) we introduce semantic operations $F_{2,IA}, \dots, F_{17,IA}$ with the following intuitive significance: if the expression μ is « obtained » by the n^{th} syntactic rule from the expressions $\delta_1, \dots, \delta_k [\dots]$ and if g_1, \dots, g_k are possible denotation functions of $\delta_1, \dots, \delta_k$ respectively, then $F_{n,IA}(g_1, \dots, g_k)$ is to be the corresponding denotation function of μ .

(MONTAGUE 1974 [76] p. 201)

¹³ En mathématique, on étudie les fonctions établissant des liens entre ensembles munis de structures analogues. Ces fonctions appliquent la structure d'un ensemble sur celle de l'autre. Dans le cas des structures algébriques, on peut qualifier certaines de ces fonctions d'homomorphisme. Par exemple, étant donnés deux groupes $(G, *)$ et (H, \bullet) , l'application $\varphi : G \rightarrow H$ est un homomorphisme si pour tout $g_1, g_2 \in G$ on obtient $\varphi(g_1 * g_2) = \varphi(g_1) \bullet \varphi(g_2)$. Deux ensembles structurés sont considérés comme algébriquement équivalents lorsqu'il existe un homomorphisme bijectif entre les deux. Dans ce cas, ils sont dits isomorphes.

Les lettres I et A représentent respectivement l'ensemble des mondes possibles et le domaine donnés par le modèle. La compositionnalité est obtenue par le fait que la composition sémantique qui conduit à l'interprétation de l'expression μ est entièrement calculable à partir de l'agencement de ses composants syntaxiques et de leur signification. Ainsi, la structure syntaxique de la phrase *a woman loves every man* est construite, dans la description qui suit, par l'application successive des règles S3, S10, et S9¹⁴.

[a woman loves every man]			(S9)
[a woman loves v_1]	v_1	[man]	(S10)
[v_0 loves v_1]	v_0	[woman]	(S3)
[loves]	v_0	v_1	

Les règles S3, S10 et S9 ont respectivement pour rôle la construction d'un syntagme verbal à partir d'un verbe transitif, l'introduction d'un syntagme nominal déterminé par l'article *a*, et l'introduction d'un syntagme nominal déterminé par le mot *every*. L'interprétation de cette construction à l'aide des opérations sémantiques F_3 , F_{10} et F_9 s'effectue de manière suivante.

$$F_9(F_{10}(F_3(G([\text{loves}]), G(v_0), G(v_1)), G(v_0), G([\text{woman}]), G(v_1), G([\text{man}])))) \quad (F_9)$$

$$F_{10}(F_3(G([\text{loves}]), G(v_0), G(v_1)), G(v_0), G([\text{woman}])) \quad (F_{10})$$

$$F_3(G([\text{loves}]), G(v_0), G(v_1)) \quad (F_3)$$

La fonction G donne l'interprétation d'une entité lexicale ou d'une variable. Les fonctions F_3 , F_{10} et F_9 représentent respectivement l'interprétation d'une relation binaire, d'une quantification existentielle et d'une quantification universelle. Le résultat est équivalent à la formule de la logique du premier ordre $\forall x(\text{man}(x) \supset \exists y(\text{woman}(y) \wedge \text{love}(y, x)))$. Le point important est que l'interprétation alternative de l'énoncé, représentée par la formule $\exists x(\text{woman}(x) \wedge (\forall y(\text{man}(y) \supset \text{love}(x, y)))$ provient d'une autre structure syntaxique, construite en appliquant d'abord la règle S9 et ensuite la règle S10¹⁵. Cette construction dirige en parallèle l'application de la fonction F_9 puis de la fonction F_{10} , ce qui change la portée des quantificateurs.

La volonté de mener la construction sémantique et la construction syntaxique en parallèle conduit de manière naturelle à une conception fonctionnelle des structures sémantiques. La dépendance syntaxique d'un syntagme par rapport à un autre se traduit, au niveau des significations, par le fait que l'un devient argument de l'autre. Le développement de cette idée tire profit d'un certain nombre d'outils mathématiques puissants, en particulier le calcul des lambdas.

[lambda calculus] will provide a formalism in which we can assign meanings to the basic expressions of a language and all of the larger constituents constructible from the basic expressions. In other words, it forms the basis of the compositional method of defining the meanings of expressions in terms of the meanings of their parts. This can be contrasted with first order logic, which is rather impoverished in terms of being able to assign meanings to the relevant syntactic constituents of an expression. (CARPENTER 1997 [10] p. 39)

Le choix d'un formalisme comme le lambda-calcul est motivé par le caractère artificiel de la représentation déclarative. Un syntagme comme *ballon rouge* est traduit classiquement comme une conjonction logique $\text{ballon}(x) \ \& \ \text{rouge}(x)$ dans laquelle la variable x est

¹⁴ Nous rapportons ici, en la simplifiant, l'analyse faite dans (MONTAGUE 1974 [76]) de cet exemple, qui est analogue à la phrase *toutes les disquettes ont un virus* mentionnée dans la section précédente.

¹⁵ Dans le cas du présent exemple, l'alternative syntaxique suppose un syntagme verbal ternaire où le sujet et le complément jouent des rôles symétriques. Cette possibilité est refusée par de nombreux syntacticiens.

répétée pour signaler qu'elle désigne le même objet. En revanche, une traduction fonctionnelle comme (**rouge**(**ballon**)) fait apparaître une dépendance entre les deux termes, dépendance qui est parallèle à la subordination syntaxique de l'adjectif au nom. La même idée est présente dans le formalisme de la logique combinatoire¹⁶. Ce type de langage formel fait intervenir explicitement des termes qui sont des combinateurs. La combinaison résulte alors de la juxtaposition des termes. Le combinateur se comporte comme une fonction qui se combine avec un terme voisin en le prenant comme argument. Dans ce cadre, contrairement au cas des formalismes déclaratifs, les variables n'interviennent qu'en tant que variables "muettes" pour signaler les fonctions.

Le processus de combinaison sémantique résulte ainsi de l'application récursive, dirigée par la syntaxe, de fonctions à leurs paramètres. Au lieu de constituer des arbres complexes, le décodage de la phrase procède, sous l'action des combinateurs, en instanciant progressivement les arguments des fonctions par d'autres fonctions. Pour enclencher ce processus, les catégories syntaxiques élémentaires sont décrites comme des fonctions dont le type d'argument est précisé. Par exemple, dans le cadre d'une grammaire catégorielle, un verbe peut être noté par la catégorie $(np \setminus s) / np$, ce qui signifie qu'il accepte un syntagme nominal de catégorie np à droite, le complément, puis un syntagme nominal à gauche, le sujet, pour donner une phrase de catégorie s . De la même manière, un déterminant pourra être décrit par la catégorie np / n , puisqu'il prend un nom de catégorie n à droite pour donner un syntagme nominal. De même, les structures sémantiques peuvent être construites comme des combinaisons de termes qui représentent les significations de base. L'exemple suivant montre comment les deux structures, syntaxique et sémantique, sont construites en parallèle dans un formalisme de grammaire catégorielle¹⁷. Le signe ":" sépare dans chaque cas l'expression sémantique, à gauche, de l'expression syntaxique, à droite.

<i>Someone</i>	<i>breaks</i>	<i>everything</i>	
some: $np \uparrow s$	break: $(np \setminus s) / np$	every: $np \uparrow s$	(1)
$x:$ np		$y:$ np	(2)
	break (y): $np \setminus s$		(3)
	break (y)(x): s		(4)
	every ($\lambda y.$ break (y)(x)): s		(5)
	some ($\lambda x.$ every ($\lambda y.$ break (y)(x))): s		(6)

Cet exemple permet d'illustrer à nouveau la manière dont sont gérées les quantifications. Dans cette dérivation, le verbe accepte son complément en phase (3) et son sujet en phase (4). Noter que la dérivation sémantique suit fidèlement la progression syntaxique : chaque réduction syntaxique se traduit par une nouvelle relation fonction - argument. Le traitement sémantique des quantificateurs intercale des phases supplémentaires. La phase (2) consiste à créer des variables qui jouent le rôle d'hypothèses. Sous ces hypothèses, on conclut à la dérivation de la forme (4). La réécriture du lien

¹⁶ Le formalisme de la logique combinatoire est, en tant qu'outil de génération de dérivations, équivalent au calcul des lambdas.

¹⁷ Cet exemple est adapté de (CARPENTER 1997 [10]).

hypothèse - conclusion conduit successivement aux phases (5) et (6). L'ordre respectif de ces deux réécritures explique l'ambiguïté de la double quantification¹⁸.

Cette approche de la sémantique formelle dans laquelle les processus de construction du sens sont parallèles aux processus syntaxiques présente un intérêt indéniable d'un point de vue computationnel. Sur le plan de la modélisation cognitive, elle peut cependant être critiquée à plusieurs égards.

La première critique porte sur la vraisemblance de l'homomorphisme entre syntaxe et sémantique. L'existence d'une correspondance stricte entre la syntaxe de la langue et un système formel chargé de représenter la composition sémantique semble contredire toute possibilité d'autonomie relative de la syntaxe et de ce système formel. À la limite, pourquoi parler de deux systèmes combinatoires différents ? Or, d'un point de vue cognitif, il semble bien exister deux jugements d'acceptabilité différents, l'un syntaxique, portant sur l'agencement des mots, et l'autre sémantique, portant sur l'existence d'un sens. Comment expliquer, s'il existe une correspondance stricte, que le sens d'une phrase syntaxiquement incorrecte comme *Pierre pas aimer à Jacques* puisse, dans une situation donnée, être compris, alors que dans certains contextes, la phrase syntaxiquement correcte le carrelage approuve pourra rester incomprise ?

D'un point de vue technique, la thèse de l'homomorphisme strict peut signifier qu'il n'existe, cognitivement, qu'un seul système combinatoire. Ce qui nous apparaît comme la syntaxe ne serait alors que le versant structurel visible de la combinatoire sémantique. Peut-on analyser l'ensemble des phénomènes grammaticaux comme des conséquences de contraintes liées à l'articulation du sens ? La sémantique formelle va clairement dans cette direction. En confiant au lexique la quasi-totalité des instructions de combinaison, elle ne laisse plus de place à des phénomènes de combinaison déconnectés du sens. Dans le formalisme des grammaires catégorielles, par exemple, chaque entrée du lexique comporte un élément d'algorithme qui dirige sa propre insertion dans la phrase. L'expression $(np \setminus s) / np$, qui caractérise un verbe transitif, peut être lue comme un algorithme qui attend des données de type np sur sa droite puis sur sa gauche pour produire une expression de type s . Selon le même principe, les instructions contenues dans les entrées du lexique contrôlent leur propre comportement phonologique et sémantique, si bien que le lexique est, idéalement, censé diriger l'ensemble des processus linguistiques. Une telle position extrême est cependant difficile à tenir en raison de la présence de phénomènes syntaxiques globaux sans contrepartie sémantique. Par exemple, aucune contrainte d'ordre lexical, phonologique ou sémantique n'explique que certaines langues comme le français ou l'anglais imposent systématiquement la présence d'un sujet pour les verbes conjugués, au point d'introduire des sujets explétifs comme le pronom *il* dans la phrase *il pleut*, alors que des langues comme le persan se satisfont de verbes sans sujets. De même, le mode de branchement des syntagmes diffère d'une langue à l'autre, ce qui conduit par exemple le persan à être de type SOV alors que le français est SVO. Parmi les contraintes structurelles systématiques qui dépassent le lexique, on peut également citer la contrainte de liage qui agit par la relation de c-commande, que nous avons mentionné dans la section précédente. Ces phénomènes de structure requièrent l'adjonction de

¹⁸ Cet exemple utilise les principes du calcul de Lambek. La forme $P : np \setminus (s)$ induit une dérivation conditionnelle. Si, sous l'hypothèse $x : np$, posée lorsque P se présente, on parvient à dériver syntaxiquement $Q(x) : s$, alors on a dérivé $P(\lambda x. Q(x)) : s$. Ainsi, le quantificateur est traité comme une fonction dont l'argument est une fonction de la variable quantifiée. Noter que, contrairement à l'exemple précédent, l'ambiguïté liée à la portée respective des quantificateurs n'est pas d'origine syntaxique. La production de **every** $(\lambda y. \text{some } (\lambda x. \text{break } (y) (x)))$ n'est pas liée à l'ordre dans lequel les arguments du verbe sont intégrés, mais provient d'un ordre différent dans le "déchargement" des hypothèses.

mécanismes généraux à la grammaire, par exemple le mécanisme de montée de type¹⁹ utilisé dans les grammaires catégorielles. De telles techniques conduisent toutefois à réintroduire des opérations purement syntaxiques, au détriment du parallélisme syntaxe - sémantique qui est censé motiver l'ensemble de la démarche. Ainsi, il est difficile de concevoir techniquement un homomorphisme parfait. Or, toute différence de traitement laisse entrevoir la nécessité de deux mécanismes distincts, l'un chargé des opérations syntaxiques et l'autre des opérations sémantiques, ainsi que d'une interface entre les deux.

Une deuxième critique que l'on peut adresser à la sémantique formelle d'un point de vue cognitif est, paradoxalement peut-être, la trop grande puissance de ses formalismes. Le formalisme de la logique combinatoire ou le lambda-calcul permettent d'exprimer l'ensemble des fonctions calculables²⁰. Il faut donc imaginer que les catégories admises pour représenter la compétence langagière appartiennent à un sous-ensemble restreint de toutes les expressions formelles possibles. Or, aucun critère permettant de délimiter ces dernières ne se dégage clairement. Quel principe interdit l'existence d'une catégorie comme $(np \setminus np) / np$? Si ce principe existe, comment peut-il être justifié sur le plan de l'organisation cognitive? Il est en effet plus facile de justifier, par exemple par des explications fonctionnelles, l'existence d'une capacité que sa non existence. Si l'esprit humain possède une capacité de calcul générale capable de gérer des expressions formelles comme celles des grammaires catégorielles, pourquoi est-il restreint à ne manipuler que certaines d'entre elles? Si, en revanche, cette capacité générale de calcul n'existe pas, alors les modèles de la sémantique formelle ne peuvent être que des modèles de description, non des modèles cognitifs.

Une troisième critique peut être adressée au présupposé principal des approches formelles de la sémantique, l'idée que la combinaison conceptuelle est réalisée par une relation fonction - argument. Par exemple, la modification du nom *ballon* par l'adjectif *rouge* est rendue par le fait que le terme *ballon* devient argument du terme *rouge*. Or, le fait de traduire les constructions grammaticales par des structures fonctionnelles ne suffit pas à nous faire progresser sur le plan cognitif. La signification du syntagme *ballon rouge* n'est pas élucidée par une expression comme *rouge(ballon)*. Comme dans toutes les approches formelles, la charge de l'interprétation est repoussée dans la procédure de vérification, seule capable d'établir un lien entre les symboles *rouge* et *ballon*. La dérivation donnée plus haut pour la phrase *someone breaks everything* illustre le fait que le décodage se contente de transférer, étape par étape, la combinatoire de l'énoncé langagier dans la structure d'une expression formelle. Ce décodage ne nous renseigne pas sur l'interprétation de l'énoncé. Par exemple, il ne rend pas compte du lien systématique entre les expressions *rouge(ballon)* et *vert(ballon)*. Pour cela, il faut conférer une structure interne aux symboles utilisés dans le formalisme, de manière à ce que les fonctions *rouge* et *vert* puissent partager le fait d'être des indications de couleur. C'est ce que proposent les approches décrites dans la section suivante.

Dans la présente section, nous avons considéré l'apport des systèmes de sémantique formelle pour la compréhension de la construction du sens des énoncés. Si de tels systèmes présentent un intérêt technique parfois incontournable, ils n'en constituent pas pour autant des modèles cognitifs plausibles des processus en jeu dans la construction du sens. En particulier, ils repoussent la difficulté de l'interprétation en postulant l'existence d'une procédure générale de vérification, dont les propriétés sont cognitivement non plausibles (tests binaires dans des ensembles infinis de mondes eux-mêmes infinis). Par ailleurs, la structure

¹⁹ Le mécanisme de montée de type transforme par exemple un terme de catégorie *np* en catégorie $s / (np \setminus s)$ (*id est* une phrase attendant un groupe verbal). La catégorie *np* se voit ainsi imposer la contrainte d'apparaître en position sujet.

²⁰ Ces systèmes ont la puissance d'une machine de Turing.

sémantique utilisée dans les approches formelles, en associant un symbole sémantique différent à chaque entrée lexicale, n'explique pas les liens systématiques qui peuvent exister entre les combinaisons conceptuelles. Il est donc légitime de s'intéresser à une approche plus cognitive de la sémantique, dans laquelle le formalisme utilisé pour représenter les significations est censé modéliser les représentations telles qu'elles sont supposées être évoquées par le langage et ancrées dans la perception.

5.3. Concepts et structures

Une approche cognitive de la sémantique se doit de déterminer les représentations et les mécanismes mentaux à l'œuvre dans la procédure de construction du sens, non des opérations supposées se produire dans des espaces abstraits servant de substitut à une réalité placée à l'extérieur des locuteurs. Certaines approches cognitives de la sémantique sont ainsi guidées par l'idée que le sens d'un énoncé est une représentation mentale structurée, et que l'objectif de la recherche en sémantique est de déterminer cette structure.

[...] the act of understanding a sentence S – recovering its meaning – is to be regarded as standing S in correspondence with a mentally encoded concept C which has internal structure derivable from the syntactic structure and lexical items of S.

(JACKENDOFF 1990 [50] p. 11)

La construction du sens d'un énoncé s'explique par la composition mentale des représentations conceptuelles évoquées par les mots de l'énoncé. L'objectif est donc de décrire une interface entre deux systèmes compositionnels, le système syntaxique et le système conceptuel. L'explication de cette interface commence par le fait d'associer une structure conceptuelle à chaque entrée lexicale (LCS).

An example like [the following one] provides a good first illustration [for the relation of syntactic and conceptual argument structure].

Syntactic structure

[S [NP John] [VP ran [PP into [NP the room]]]]

Conceptual structure

[Event GO ([Thing JOHN], [Path TO ([Place IN ([Thing ROOM])]])])]

[...] In order to see how [it] is put together from its parts, let us examine the lexical entries for the two words *into* and *run* [...].

into	
P	
__ NP _j	
[Path TO ([Place IN ([Thing]j)])]	

run	
V	
__ <PP _j >	
[Event GO ([Thing]i, [Path]j)]	

(JACKENDOFF 1990 [50] p. 45)

La LCS peut être simple, comme les concepts JOHN ou ROOM. Elle peut comporter des indications de type, comme *Thing* ou *Path*, et des fonctions comme GO ou TO qui

transforment leurs arguments en une nouvelle structure. L'énoncé *John ran into the room*, du fait de la fonction GO amenée par la LCS du verbe *run*, est interprété comme un événement. La correspondance avec la syntaxe permet de placer le concept JOHN, qui est l'interprétation du sujet grammatical, en premier argument de la fonction GO. De la même façon, la structure associée au groupe prépositionnel prend la place du deuxième argument de la fonction GO. Ces inclusions de structures conceptuelles au sein d'autres structures conceptuelles s'effectuent en respectant les types. Ainsi, le concept ROOM est du bon type *Thing* pour être argument de la fonction IN. Cette fonction transforme le concept ROOM en type *Place*, ce qui lui permet d'être l'argument de la fonction TO. L'ensemble a le bon type, *Path*, pour venir occuper le deuxième argument de la fonction GO.

La structure conceptuelle ainsi proposée est donc plus riche que celle que requiert la thêta-théorie de la grammaire générative, qui se réduit à préciser une structure argumentale, notamment pour les verbes et les prépositions. Elle est également plus complexe que la forme logique des approches formelles qui associent une fonction à chaque entrée lexicale, sans souci des liens systématiques que ces fonctions conceptuelles peuvent avoir entre elles.

Instead of a division of formal entities into such familiar logical types as constants, variables, predicates, and quantifiers, each of which has nothing in common with the others, it is argued that the essential units of conceptual structure are *conceptual constituents*, each of which belongs to one of a small set of major ontological categories (or conceptual « parts of speech ») such as Thing, Event, State, Action, Place, Path, Property, and Amount. [...] Each conceptual category has some realizations in which it is decomposed into a function argument structure; each argument is in turn a conceptual constituent of some major category. (JACKENDOFF 1990 [50] p. 22)

Cette approche de la représentation conceptuelle conserve l'aspect fonctionnel des approches formelles. Cependant, les fonctions ne sont plus l'ensemble des symboles associés aux entités lexicales de la langue, mais sont limitées à un petit nombre de primitives. Ces fonctions primitives sont supposées représenter des composantes fondamentales de notre cognition. Le fait de représenter le sens en appliquant ces fonctions aux arguments qui leur sont présentés devient cognitivement plausible, dans la mesure où grâce au typage, l'ensemble des configurations possibles reste limité. Le point fort de l'approche est ainsi que les structures conceptuelles tirent leur justification du fonctionnement cognitif, non de la structure de la langue. Le lien avec le langage prend donc la forme d'une véritable interface, dont la loi générale peut s'exprimer de la manière suivante.

[...] the fact that major syntactic phrases correspond to major conceptual constituents [...] and the basic correspondence of syntactic and conceptual structure [...] [can be formalized as the following general correspondence rules].

XP corresponds to [Entity].

$$\left[\begin{array}{c} X^0 \\ _ _ < YP < ZP >> \end{array} \right] \text{ corresponds to } \left[\begin{array}{c} \text{Entity} \\ F (< E_1 >, < E_2, < E_3 >>) \end{array} \right]$$

(JACKENDOFF 1990 [50] p. 25)

XP désigne un syntagme dont la tête est l'item lexical X^0 . Les syntagmes YP et ZP sont les compléments du syntagme global. À cette structure syntaxique correspond la structure sémantique $F(< E_1 >, < E_2, < E_3 >>)$ où F est une fonction sémantique²¹, E_2 et E_3 sont les structures conceptuelles associées respectivement aux expressions YP et ZP, E_1 est associée au spécifieur du syntagme, s'il est présent. Ainsi, la structure conceptuelle de la phrase *John*

²¹ Les crochets <> indiquent un caractère optionnel.

ran into the room conduit à l'expression fonctionnelle GO(JOHN, TO(IN(ROOM))) dont les deux arguments sont occupés par les structures conceptuelles respectives du sujet John et du complément into the room. De même, la structure conceptuelle du syntagme into the room est une expression fonctionnelle qui s'exprime comme la composée de deux fonctions primitives IN et TO et dont le seul argument est occupé par le groupe nominal the room.

L'intérêt premier de ce genre de modèle est qu'il ne fait pas de différence qualitative entre les structures lexicales et supra-lexicales. En d'autres termes, il s'agit d'un système intégralement compositionnel : les structures conceptuelles lexicales, comme les structures conceptuelles des énoncés, résultent de compositions de concepts primitifs. Grâce à ce principe, il est possible de réduire considérablement l'ensemble des éléments primitifs, ceux pour lesquels on doit spécifier directement les propriétés, notamment leur ancrage et leur rôle inférentiel. Le sens de toutes les autres structures peut, en principe, être déduit par composition. La tâche de la sémantique lexicale se ramène ainsi, pour l'essentiel, à déterminer la structure sémantique que l'on attribue aux unités du lexique. De cette manière, contrairement à l'approche formelle, il n'est pas nécessaire de postuler l'existence d'une procédure de vérification externe capable de décider de l'interprétation de toute expression. Le sens des mots, comme celui des phrases, s'obtient par une combinaison récursive de significations de base.

Le principe d'une interface systématique entre le système conceptuel et le système grammatical semble cependant mis en défaut par un certain nombre de difficultés. La structuration des entités sémantiques exige que la syntaxe fournisse des arguments convenables à ces structures, c'est-à-dire des objets sémantiques présentant le type adapté à la place à laquelle ils sont destinés. De ce point de vue, la sémantique formelle pouvait se montrer beaucoup plus permissive, en ne posant pas de contraintes dans la relation fonction - argument autre qu'une spécification de la catégorie syntaxique attendue. Or, dans l'approche de la sémantique conceptuelle, les contraintes de typage sémantique semblent souvent enfreintes. Par exemple, dans la phrase la sérénité entre dans la pièce, la définition sémantique du verbe entrer suppose que le sujet vienne occuper un argument typé comme chose, or le concept associé au mot sérénité est plutôt de type propriété. Ainsi le concept SÉRÉNITÉ ne peut pas être "substitué" à la structure du premier argument de la fonction associée au verbe entrer. Ce problème peut être résolu par l'introduction d'un mécanisme de fusion, qui remplace celui de substitution (JACKENDOFF 1990 [50]). L'idée est qu'en cas de conflit de structures, la structure conceptuelle associée au syntagme en position de sujet doit être "fusionnée" avec la structure du premier argument attendu par le syntagme verbal. Cette opération autorise un assouplissement local des contraintes d'appariement, ce qui permet en principe d'expliquer les tournures non littérales, jusqu'à certaines métaphores²².

Une autre difficulté se présente avec des phrases comme la vitre casse. La définition du verbe casser, qui permet d'interpréter une phrase comme Jean casse la vitre, peut être décrite par la structure suivante.

[Event CAUSE([Entity _],[Event GO([Entity _],[Path TO([State CASSÉ])])])]

Or, dans la phrase la vitre casse, on s'attend à ce que le sujet grammatical vienne occuper la place du premier argument de la fonction GO, contrairement à ce que prévoit l'interface standard. On peut imaginer trois manières de résoudre ces problèmes d'interface (JACKENDOFF 1990 [50]). (1) Il est possible de traiter le problème de manière purement syntaxique. Le syntagme la vitre occuperait la position sujet uniquement dans la structure de

²² Techniquement, la substitution et une partie de la fusion peuvent être assurées par le mécanisme standard d'unification. Ce même mécanisme peut également permettre de contrôler jusqu'à quel niveau les contraintes peuvent être enfreintes. Certaines métaphores, comme l'entrée de la sérénité dans une pièce, peuvent ainsi être expliquées et prédites.

surface, à la suite d'un mouvement provoqué par la vacuité de cette position dans la structure profonde. Dans ce cas, l'interface syntaxe - sémantique reste valide si elle s'effectue au niveau de la structure profonde. Cette solution est naturelle en génération, lorsque la structure profonde est élaborée avant la structure de surface. En compréhension, cependant, la reconstitution d'une structure profonde capable de recevoir un sens passerait par un échec, lorsque l'argument de la fonction GO, le thème, ne peut être apparié. Cet échec serait suivi d'un retour en arrière²³. Or il n'y a pas d'indication claire selon laquelle la compréhension de la phrase la vitre casse demanderait un temps de traitement supérieur à celle de la phrase il casse la vitre. (2) Une deuxième solution consiste à complexifier la règle de correspondance syntaxe - sémantique en introduisant une hiérarchie entre les rôles sémantiques. Par exemple, le premier argument de la fonction CAUSE, que l'on peut dénommer agent, serait prioritaire pour s'interfacer avec le sujet grammatical. Dans le cas où il est non spécifié, c'est le thème, le premier argument de la fonction GO qui pourrait se saisir du sujet. Noter qu'encore une fois, l'explication est naturelle en génération, mais l'est beaucoup moins en compréhension où elle nécessite un effet de retour en arrière. (3) Une troisième solution est de considérer des verbes comme *casser* comme polysémiques. En particulier, ils comporteraient deux définitions, l'une causale et l'autre non. Ainsi, le verbe *casser* aurait une deuxième définition décrite par la structure suivante.

[Event GO([Entity _],[Path TO([State CASSÉ]])])]

La sélection de la bonne définition dépendrait ainsi de la présence d'un complément. Une solution équivalente est de marquer la fonction CAUSE comme optionnelle dans la définition initiale du verbe *casser*. Il est important de noter que la mention d'une ambiguïté dans l'entrée lexicale est une mention purement sémantique. Rien ne distingue le verbe *casser* d'un verbe comme *manger* sur le plan syntaxique, puisque les deux verbes peuvent être utilisés indifféremment avec ou sans complément. Alors que les deux premières solutions se traduisaient par une complication des mécanismes d'interface, la présente solution consiste à augmenter la complexité des représentations sémantiques de manière à prédire correctement les conditions d'emploi des mots associés.

Si l'on systématisait la logique de cette dernière approche, on en vient à classifier de manière toujours plus fine les structures lexicales. La méthodologie linguistique correspondant à ce courant part de l'observation des configurations syntaxiques où une entité lexicale apparaît, puis tente de doter cette entité lexicale des traits sémantiques minimaux qui permettent de prédire ses conditions d'emploi. La structure conceptuelle est ainsi justifiée par la seule prise en compte du caractère compositionnel des mots.

[...] knowing the meaning of a verb can be a key to knowing its behavior. Presumably, predictions about verb behavior are feasible because particular syntactic properties are associated with verbs of a certain semantic type. [...] [Our examples] suggest that general principles of grammar are at work, allowing the syntactic behavior of a verb to be predicted from its meaning. (LEVIN 1993 [66] p. 5)

L'inclusion de traits appropriés dans la structure conceptuelle permet par exemple de distinguer les conditions d'emploi de verbes comme *break*, *cut* et *hit*²⁴.

²³ Il s'agit du phénomène connu sous le nom de *garden path*. Le sujet interprète une phrase comme *the horse ran past the barn fell* de manière directe, en traitant, dans un premier temps, *ran* comme un verbe conjugué. Emporté ainsi par son élan, il se heurte au verbe *fell* qui n'entre pas dans la trame syntaxique élaborée jusque-là. Dans cette situation, il effectue un retour en arrière pour réinterpréter *ran* comme le participe passé d'un verbe transitif.

²⁴ Exemples adaptés de (LEVIN 1993 [66]).

Janet broke the cup / the cup broke
Janet cut the bread / * the bread cut

Carla cuts the meat / the meat cuts easily
Carla hits the door / * the door hits easily

L'idée est non seulement d'attribuer ces contrastes à des propriétés structurelles des mots, mais également de classer les mots selon leur comportement. De cette manière, les mots dérivent de classes plus ou moins générales.

By defining the functional behavior of lexical items at different levels of representation we hope to arrive at a characterization of the lexicon as an active and integral component in the composition of sentence meanings. Because of the more expressive mechanisms involved in composition, this approach will enable us to conflate different word senses into a single *meta-entry*, encoding regularities of word behavior dependent on context, and as a result, greatly reducing the size of the lexicon. (PUSTEJOVSKY 1995 [86] p. 62)

Les différentes approches génératives du lexique voient donc dans la sémantique des mots une combinaison de composants conceptuels en nombre limité. Les versions radicales se donnent comme objectif de prédire, à l'aide de ces structures, non seulement les combinaisons grammaticales standard, mais également les possibilités d'insertion dépendantes du contexte sémantique. Dans l'exemple suivant, on constate que les combinaisons dans lesquelles le verbe *began* est admis dépendent de la nature sémantique des autres mots²⁵.

Mary began to bleed / her bleeding began
Mary began to draw / * her drawing began

Mary began the book / * the book began
Mary began the movie / the movie began

Le syntagme *begin to* + GV peut recevoir deux interprétations différentes : il peut s'agir d'un effet causal, exercé par le sujet, sur l'amorce d'un événement ; il est aussi possible que seul le début de l'événement soit pertinent, sans qu'il y ait une référence à l'origine causale de son départ²⁶. Avec le verbe *bleed*, l'interprétation non causale est préférentiellement sélectionnée, alors que l'interprétation causale est induite par le verbe *draw*. Cela se vérifie par le fait que la paraphrase *her drawing began* est peu acceptable. De même, dans les exemples de la forme grammaticale *begin to* + GN c'est la définition causale qui est préférentiellement choisie. Ceci peut être expliqué par un mécanisme de coercition. Par exemple, dans le cas du mot *book*, on peut supposer deux coercitions différentes, celle où le livre est lu et celle où le livre est écrit. Par contre aucune coercition associée à la structure conceptuelle du mot *book* ne peut donner du sens à la construction *the book began*. En revanche, parmi les coercitions acceptées par l'entrée lexicale *movie*, se trouve une interprétation qui lui permet de se trouver en position du premier argument du verbe *begin*, celle où la projection du film commence. On constate donc que le sens d'un composé s'élabore par un processus de sélection mutuelle entre les traits sémantiques des composants.

L'ambition de la sémantique lexicale est de parvenir à discriminer les combinaisons licites de mots, de la même manière que les syntacticiens tentent de discriminer les combinaisons licites des catégories syntaxiques. Pour cela, le moyen est de constituer des

²⁵ Exemples adaptés de (PUSTEJOVSKY 1995 [86]).

²⁶ Cette alternative est analysée, dans certaines théories syntaxiques, par une distinction entre deux catégories grammaticales différentes, verbe de contrôle et verbe de montée, qui peuvent être attribuées à deux entrées lexicales différentes du verbe *begin*.

structures toujours plus riches, en ajoutant des traits restrictifs lorsque la discrimination échoue et en introduisant des disjonctions ou des traits facultatifs lorsque des configurations permises ne sont pas prédites. Cette ambition se justifie sur un plan technique. Peut-on espérer parvenir ainsi à un modèle cognitif du système conceptuel ? Il faut pour cela que la méthode lexicale parvienne à enfermer l'intégralité des conditions d'emploi des mots dans des structures fixes. Nous évoquerons plus tard le danger que l'inventaire des traits introduits ne soit pas borné, et que l'origine, ontogénétique ou phylogénétique, de ces traits soit difficile à justifier (CF. CHAPITRE 7). Pour le moment, nous pouvons nous interroger sur les limites de la méthodologie, à la lumière des exemples précédents.

La définition que nous avons donnée du verbe *casser* utilise la primitive *CASSÉ*. Si on la remplace par une mention plus générale comme *NON FONCTIONNEL*, le risque est qu'il existe des objets cassés qui restent fonctionnels (par exemple un couteau) et des objets qui cessent d'être fonctionnels sans être cassés (par exemple un ordinateur). Il n'est donc pas évident que l'inventaire des traits soit significativement plus concis que le lexique lui-même. De plus, les emplois considérés comme illicites le sont dans des situations standard ou prototypiques. Ainsi, la construction *le livre commença* est un énoncé sémantiquement acceptable dans un conte où un livre magique raconte son contenu ou s'il s'agit d'un livre électronique en défilement automatique. Un énoncé comme *il commence le tunnel*, semble exclure l'interprétation *il commence à parcourir le tunnel*, au profit de celle de *il commence à creuser le tunnel*. Cependant cette exclusion présente un caractère relatif. La phrase peut très bien figurer, avec ce sens, dans le jargon de l'équipe de surveillance du tunnel. Dans ce cas, l'exclusion est plus dans les habitudes que dans le concept lui-même, si les habitudes peuvent changer sans que le sens en soit affecté. L'exclusion sémantique n'est donc pas de même nature que l'exclusion syntaxique. Aucune modification de contexte ni aucune habitude ne peuvent rendre la phrase *les frères espèrent que Jean aime l'un l'autre* syntaxiquement acceptable²⁷. Si l'acceptabilité des combinaisons sémantiques peut toujours être remise en question par le contexte ou l'habitude, le fait d'enfermer l'acceptabilité dans des structures lexicales ne peut atteindre que les conditions d'emploi standard et prototypiques. Dans ce cas, la méthode ne peut pas conduire, seule, à un modèle de la compétence sémantique humaine.

Une manière de rendre les structures adaptables au contexte est de leur permettre de s'établir de manière dynamique, en y incluant des procédures. C'est cette solution que proposent les approches examinées dans la section suivante.

5.4. Concepts et procédures

La construction du sens peut être conçue non comme un assemblage d'unités prédéfinies, mais comme un calcul pouvant inclure divers aspects contextuels, y compris des faits contingents ou des données de la perception. L'une des idées principales de la sémantique procédurale consiste à élaborer le sens des composés de manière dynamique. Cette idée part de la métaphore des programmes informatiques, qui sont compilés puis exécutés. De la même manière, un énoncé langagier doit être analysé et compilé pour prendre la forme d'un plan d'action exécutable. Le plan peut être comparé à la forme intensionnelle dans la sémantique formelle, tandis que le résultat de l'exécution correspond à la valeur de vérité allouée à une formule dans un certain domaine (JOHNSON-LAIRD 1977 [54]). Ainsi, un

²⁷ Cet exemple, inspiré de (CHOMSKY 1975 [15]), illustre le fait que la présence du prénom Jean en position sujet bloque le lien entre les deux syntagmes *les hommes* et *l'un l'autre*, contrairement au cas de la phrase *les frères espèrent qu'ils s'aiment l'un l'autre*, qui est syntaxiquement bien formée.

énoncé possède un sens s'il peut conduire à un plan, et si, dans un contexte donné, il donne lieu à une exécution réussie.

Pour illustrer ce point de vue, considérons des énoncés simples à propos du jeu d'échecs. L'énoncé mets le fou à droite de la dame peut conduire, lorsque la dame se trouve sur la case c3, à un plan qui consiste à déplacer le fou approprié sur la case d3. Si ce plan peut être exécuté dans le contexte courant, on peut dire que l'énoncé possède un sens. Dans le cas où le fou est bloqué ou s'il n'y a plus de fou de la couleur appropriée, l'énoncé peut être considéré comme dénué de sens²⁸. Il en est de même si la case désignée n'existe pas, comme dans le syntagme à droite de h3, où la case h3 est au bord de l'échiquier. Dans ces exemples, il n'est pas besoin de considérer une procédure de vérification valable en tout temps, en tout lieu et dans tous les mondes possibles. Il suffit d'exécuter le plan suggéré par l'énoncé dans le contexte courant.

Pour que l'exécution soit possible, il faut qu'aux mots du lexique soient associées des informations nécessaires, qui peuvent prendre la forme de schémas.

We have mentioned that the compiler must have access to a lexicon, but we have yet to consider what information the lexicon should contain, or how it should be organized. [...] We took the view that a lexical concept interrelates a word, rules governing its syntactic behavior, and a schema. A schema is made up from both functional and perceptual information, and may well include information that has no direct perceptual consequence. Moreover, lexical concepts are interrelated to one another. They are organized into semantic fields that have a *conceptual core* which reflects a deeper conceptualization of the world and integrates the different concepts with the semantic field.

(JOHNSON-LAIRD 1977 [54] p. 203)

Les schémas sont des routines élémentaires qui permettent de tester la présence d'objets ou de configurations, et exécuter des actions simples. Le joueur d'échec dispose de schémas pour reconnaître un roi, pour repérer une double menace ou pour effectuer une prise en passant. Pour calculer le sens des énoncés, le sujet doit combiner certains schémas sélectionnés parmi ceux qui sont associés aux mots de l'énoncé. Lors de cette combinaison, les schémas associés aux mots se sélectionnent mutuellement de manière à s'adapter les uns aux autres, en tenant compte des rapports syntaxiques. Ainsi, pour exécuter le syntagme le fou à droite de la dame, l'auditeur commence par traiter le complément le plus interne, le groupe nominal la dame, en déclenchant un schéma par lequel il localise la dame sur l'échiquier. Le traitement du syntagme englobant, à droite de __, l'amène à enchaîner le schéma précédent avec un schéma effectuant un déplacement d'une case vers la droite. Cette case étant blanche, l'auditeur sélectionne le schéma associé au mot fou qui consiste à localiser son fou "blanc", puis, pour terminer le traitement du verbe sous-entendu, il enchaîne un schéma de déplacement de la pièce pour amener celle-ci à droite de la dame.

Les schémas ne sont pas des routines indépendantes. Par exemple, les schémas associés aux verbes mettre, déplacer, jouer, dans le contexte du jeu d'échecs, partagent un noyau de sens qui inclut l'action de déplacement d'une pièce. Il faut donc imaginer l'ensemble des schémas comme un réseau d'interrelations. Ce réseau permet de reproduire, sous une forme émergente, l'organisation en champs sémantiques qui était obtenue, dans la section précédente, par la structuration des concepts à partir de primitives. L'intérêt du réseau d'interrelations va cependant au-delà, en permettant au système de s'adapter au contexte par la sélection dynamique des schémas appropriés.

Une manière de représenter les schémas et leur sélection dynamique consiste à utiliser des trames (*frames*). Les trames sont couramment utilisées en intelligence artificielle et sont à

²⁸ Dans une perspective pragmatique, il serait plus juste de considérer l'énoncé comme non pertinent.

l'origine des langages à objets. Elles constituent un moyen puissant et souple de représenter dans un même format les données factuelles, les connaissances procédurales et les informations valables par défaut. Ainsi, une trame associée à l'individu Jean peut stocker le fait que son anniversaire est mercredi. Par héritage, la trame permet d'accéder à l'ensemble des propriétés par défaut que possède un homme adulte. Grâce à un pointeur, la trame permet de mémoriser que la fille de Jean est une jeune fille du nom de Maud dont les caractéristiques sont stockées dans une autre trame. L'utilisation de pointeurs et leur modification dynamique permet de constituer le sens d'un énoncé. Ainsi, pour interpréter un énoncé comme la fille de Jean possède une Ferrari, un tel système interprète le syntagme la fille de Jean en accédant à la trame associée à l'individu Maud, puis établit dynamiquement un pointeur entre le champ *possession*, dans cette trame, et une nouvelle trame, créée par héritage à partir de la trame générale associée au mot *voiture* et portant la mention *Ferrari* dans le champ *marque*. L'interprétation de l'énoncé passe par le succès de cette procédure. Celle-ci peut échouer, par exemple si l'on considère le syntagme *fils de Jean* sans que le pointeur nécessaire existe, où si l'on cherche à interpréter le syntagme *le fils de la voiture*, sachant que la trame associée au mot *voiture* ne peut recevoir de champ *fils*.

La notion de schéma, instanciée entre autres par les trames, présente certes une ressemblance avec les structures mentionnées dans la section précédente : il s'agit ici aussi d'expliquer le caractère compositionnel d'un concept en y enfermant des contraintes qui doivent diriger son comportement dans les compositions possibles. Cependant, les deux types de représentations se distinguent par la portée des contraintes qu'elles posent sur la composition.

In various implementations, schemata correspond to Ceccato's *constellations* (1961), Minsky's *frames* (1975), and Schank and Abelson's *scripts* (1977). [...] By enforcing selectional constraints, canonical graphs rule out anomalies like green ideas sleeping, but they allow such unlikely combinations as purple cows. [...] By incorporating more knowledge about the world, schemata favor plausible combinations and avoid less likely possibilities. In short, canonical graphs represent everything that is *conceivable*, and schemata represent everything that is *plausible*. (SOWA 1984 [99] p. 128)

Les graphes canoniques dont il est question dans cette citation déterminent, sous la forme d'un ensemble de liens inter-conceptuels, l'ensemble des combinaisons dans lesquelles un concept peut intervenir sans que le résultat soit contradictoire avec le noyau dur de nos connaissances sur le monde perçu. Les schémas, quant à eux, rendent compte de la plausibilité des combinaisons grâce à la prise en compte dynamique de l'ensemble des connaissances conceptuelles, y compris les connaissances contingentes ou celles qui relèvent de l'intégration des situations typiques. Cette alternative rejoint le débat présenté dans le chapitre précédent à propos du rôle inférentiel des concepts. Un courant important en intelligence artificielle défend la thèse selon laquelle le caractère compositionnel d'un concept ne peut pas être cerné complètement sans considérer toutes les connaissances qui le font intervenir.

La question abordée ici relève à la fois de la sémantique lexicale (c'est l'ensemble des usages du mot "livre" qu'il s'agit de définir) et de la représentation des connaissances (on veut exprimer les propriétés du concept "livre"). Contrairement aux linguistes, qui cherchent à établir une opposition entre "le linguistique" et "l'extra-linguistique", nous ne voyons pas de véritable frontière entre ces deux domaines, bien qu'il y ait des phénomènes qui se situent clairement d'un côté ou de l'autre. (KAYSER 1997 [59] p. 225)

Ainsi, pour cerner le caractère compositionnel d'un concept, *id est* les combinaisons sensées dans lesquelles le mot associé peut apparaître, il faut auparavant analyser l'ensemble des connaissances impliquant ce concept. Cette analyse produit, pour chaque concept, un ensemble de structures distinctes.

À supposer ces problèmes résolus, c'est-à-dire l'existence d'un ensemble complet de prédicats Livre, Livre¹, ..., Livreⁿ, il resterait à *déterminer* en toute occurrence le prédicat précis qui correspond à une connaissance donnée. [...] Un livre s'exprimera dans tous les cas comme un objet x vérifiant le prédicat livre(x). C'est un raisonnement qui permettra de mieux cerner, si besoin est, ce que dénote un objet vérifiant ce prédicat ; ce raisonnement pourra être mené plus ou moins loin, et il arrivera qu'il ouvre plusieurs pistes.

(KAYSER 1997 [59] p. 228)

La recherche de la structure adéquate, dans un contexte donné, peut s'effectuer grâce à une procédure qui est déclenchée systématiquement, à chaque occurrence du mot, mais dont le déroulement et le résultat peuvent être, chaque fois, différents. Dans un certain contexte, la procédure traitera le mot livre comme un prédicat concernant l'objet, dans un autre cas le même mot conduira à un prédicat relatif à l'œuvre, *et cætera*. La procédure recherche, dans l'ensemble des structures, une structure qui soit compatible avec les connaissances contextuelles déjà construites. Cette recherche de structure est une recherche à "profondeur variable", qui s'arrête à la première structure exempte de contradictions (KAYSER 1997 [59]).

Les modèles regroupés sous le nom de sémantique procédurale ont ceci en commun qu'ils contiennent, d'une part, un réseau d'interrelations de concepts d'où sont extraits les schémas et, d'autre part, un ensemble de procédures qui s'exécutent pour produire le sens. Ces modèles sont utilisés en traitement du langage naturel, mais également dans certaines autres applications où l'idée de composition conceptuelle est pertinente. Un exemple nous est fourni par le programme Copycat (HOFSTADTER 1995 [47]). Ce programme s'efforce de reproduire les analogies que des sujets humains produisent sans effort sur des suites de lettres. Par exemple, en suivant le modèle $abc \rightarrow abd$, il s'agit de produire, à partir de $ppqqrrr$, la suite $ppqqss$. La configuration d'une suite de lettres peut être décrite par un nombre limité de relations de type successeur, prédécesseur, gauche, droit, dernière milieu *et cætera*. L'architecture de ce programme, dont de nombreux aspects sont hors de propos ici, repose sur l'idée d'un partage de la connaissance entre un réseau de concepts relativement stable et une procédure chargée d'activer des schémas temporairement.

The basic image for the Slipnet [Copycat's network of platonic concepts] is that of a network of interrelated concepts, each concept being represented by a *node* [...], and each conceptual relationship by a *link* having a numerical length, representing the "conceptual distance" between the two nodes involved. (HOFSTADTER 1995 [47] p. 212)

Le système peut activer de nouveaux liens en propageant les activations dues au contexte. Par exemple, les deux concepts SUCCESSEUR et PRÉDÉCESSEUR sont liés par la relation OPPOSÉ. Ceci fait partie du noyau conceptuel de ces nœuds. D'autres couples possèdent le même lien, par exemple le couple DROITE/GAUCHE, peut-être avec une intensité moindre. L'idée est que quand la relation OPPOSÉ est activée par le contexte, toutes les paires de concepts qui sont unies par ce lien s'activent, plus ou moins fortement. Le programme parvient à réaliser une analogie lorsque l'une de ces structures, celle qui est la plus fortement activée s'il y en a plusieurs, s'applique au motif à traiter. L'activation de ces schémas est gérée à l'aide d'un grand nombre de petits programmes qui, dans Copycat, prennent le nom de "codelet".

Typical *effector codelets* do such things as: attaching a description to an object (e.g., attaching the descriptor *middle* to the *b* in *abc*); bounding two objects together (e.g., a *successor* bond between the *b* and *c* in *abc*); making a group out of two or more adjacent objects that are bonded together in a uniform manner; making a bridge that joins similar objects in distinct strings (similarity being measured by proximity of description in the Slipnet); destroying groups or bonds, and so on. (HOFSTADTER 1995 [47] p. 221)

Ce principe des codelets peut être utilisé pour rendre compte de l'interprétation des énoncés du langage. L'activation de structures conceptuelles peut être déclenchée par l'observation des structures langagières, et ceci de manière non déterministe. Dans un tel modèle, l'aspect procédural réside dans une multitude de routines dont l'action concurrente produit un résultat émergent.

L'intérêt des méthodes de la sémantique procédurale est d'offrir un modèle du système conceptuel qui autorise la "fusion conceptuelle". Dans de nombreux systèmes sémantiques, interpréter un syntagme consiste à agréger les éléments structuraux empruntés aux composants pour constituer une structure plus complexe, censée représenter le sens du syntagme. Aucune opération ne vient diminuer la complexité de la structure ainsi obtenue. Cette monotonie de la construction du sens sera critiquée plus loin (CF. CHAPITRE 8). Sans moyen de remplacer une structure complexe par une structure simple, l'interprétation d'un énoncé complexe, comme la présente phrase, diverge en construisant un échafaudage sans plausibilité théorique et sans intérêt technique. La sémantique procédurale offre un moyen d'échapper à ce problème de monotonie. En exécutant ses procédures, elle interprète un composé comme une nouvelle structure dont la complexité n'est pas corrélée à celle de l'énoncé. Dans l'exemple illustrant le formalisme des trames, le syntagme la voiture que possède la fille de Jean était interprétée par une trame aussi simple que celle associée au prénom Jean. Il existe donc, dans les systèmes procéduraux, un mécanisme de fusion permettant de fondre deux concepts en un nouveau concept, en évitant de conserver une paire conceptuelle.

L'approche de la sémantique procédurale présente ainsi deux atouts importants : (1) contrairement aux systèmes formels, elle évite de renvoyer la question du sens dans une procédure de vérification externe au système ; (2) elle offre un mécanisme de fusion conceptuelle permettant d'attribuer un sens simple aux composés langagiers.

Sur le plan de la théorie cognitive, on peut toutefois lui adresser les reproches liés à ses qualités. L'aspect dynamique et fluide des connexions entre schémas n'offre pas de critère théorique permettant de prédire quelles sont les combinaisons qui auront un sens et celles qui n'en auront pas. L'ensemble est laissé au hasard de ce qui existe à un moment donné dans la mémoire du système. Or, la puissance d'un tel système dépend de la richesse du réseau des concepts. Pour dépasser le cadre limité des applications d'intelligence artificielle, il faut imaginer un réseau de taille colossale pour que le schéma approprié puisse être, dans chaque cas, sélectionné. Certes, toutes les parties de ce réseau ne sont pas supposées actives en même temps, mais la souplesse du système exige que toutes soient accessibles. En d'autres termes, les structures rigides des sections précédentes sont remplacées par un ensemble gigantesque et redondant de concepts reliés entre eux. Lorsqu'il s'agit d'expliquer la faculté sémantique, la sémantique procédurale se contente de postuler l'existence potentielle de tous les schémas requis et des liens qui permettent d'y accéder. Le système des schémas, quelle que soit son utilité pratique, n'offre donc pas de prise à la réfutabilité théorique. Nous reviendrons sur cette critique (CF. CHAPITRE 7).

Les différents modèles évoqués depuis le début du chapitre organisent la construction du sens comme une procédure dirigée par la syntaxe. On peut imaginer, à l'inverse, que ce qui apparaît comme des relations syntaxiques formelles n'est que le reflet visible de processus

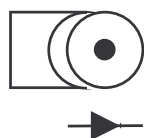
sémantiques sous-jacents. Les modèles de la linguistique cognitive vont dans ce sens, en suggérant que les relations entre les mots n'obéissent qu'à des contraintes liées au sens.

5.5. La représentation schématique du sens

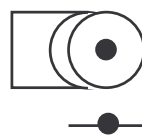
Certaines approches de la sémantique du langage, regroupées sous l'étiquette grammaires cognitives, proposent d'analyser la combinaison conceptuelle au moyen de représentations d'un type particulier, pouvant inclure des schémas graphiques. Ces schémas ont pour rôle de résumer les propriétés essentielles des représentations cognitives associées aux expressions langagières. L'idée générale est que les représentations cognitives des mots prennent leur place dans des structures schématiques associées aux constructions grammaticales pour produire la représentation cognitive correspondant à l'énoncé.

[...] we take a sentence (or a portion of discourse) to evoke in the listener a particular kind of experiential complex, here termed a **cognitive representation** or CR. The grammatical and lexical subsystems in a sentence seem generally to specify different portions of a CR. Together, the grammatical elements of a sentence determine the majority of the *structure* of the CR, while the lexical elements together contribute the majority of its *content*. The grammatical specifications in a sentence, thus, provide a conceptual framework or, imagistically, a skeletal structure of scaffolding for the conceptual material that is lexically specified. (TALMY 2000 [102] p. 21)

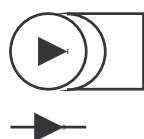
Dans cette vision de l'organisation du système conceptuel, l'élaboration des représentations cognitives allouées aux énoncés repose sur l'opposition entre structure et contenu. La structure est fournie par les éléments grammaticaux. Ces éléments incluent non seulement les entités lexicales de la classe dite fermée, comme les prépositions, les déterminants ou certains adverbess, mais également les marques morphologiques, comme l'inflexion, les cas ou les clitiques, ainsi que des entités de nature syntaxique, représentées par les catégories syntaxiques comme nom ou verbe, et les relations syntaxiques comme spécifieur ou complément. Parmi les éléments susceptibles de structurer la signification des énoncés, il faut également inclure certaines entités lexicales fondamentales qui présentent des propriétés conceptuelles particulières, comme c'est le cas du verbe *keep* dans l'exemple suivant²⁹.



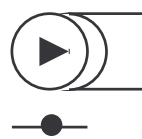
**The ball kept rolling
because of the wind blowing on it.**



**The shed kept standing
despite the gale wind blowing against it.**



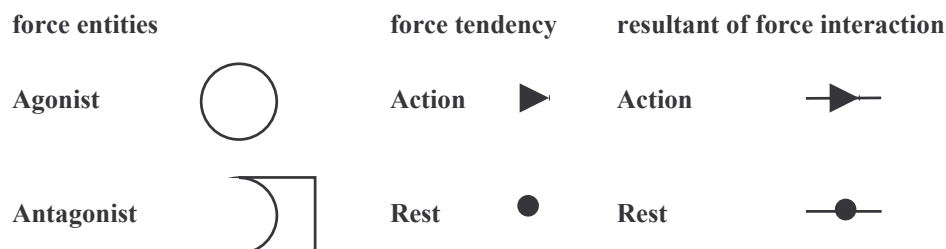
**The ball kept rolling
despite the stiff grass.**



**The log kept lying on the incline
because of the ridge there.**

²⁹ D'après (TALMY 2000 [102]).

Cet exemple montre comment le sens des énoncés peut être schématisé. L'objectif est de produire une représentation du sens de l'énoncé, en s'appuyant sur la structure introduite par des mots comme *keep*, *because* et *despite*. L'intérêt de ce type de schématisation est de se limiter à un nombre restreint d'éléments de base, de manière à présenter un caractère prédictif et productif.



La structure conceptuelle associée à la phrase *the ball kept rolling because of the wind blowing on it* prend la forme d'une action résultant de deux forces antagonistes. Dans la structure conceptuelle associée à la phrase *the log kept lying on the incline because of the ridge there*, le résultat des deux forces antagonistes est une immobilisation. Ces représentations schématiques résultent de l'interprétation d'éléments lexicaux de la classe ouverte, comme les verbes *roll* ou *lie*, et de l'interprétation d'éléments de nature grammaticale, comme la conjonction *because* ou la forme *keep + -- ing*.

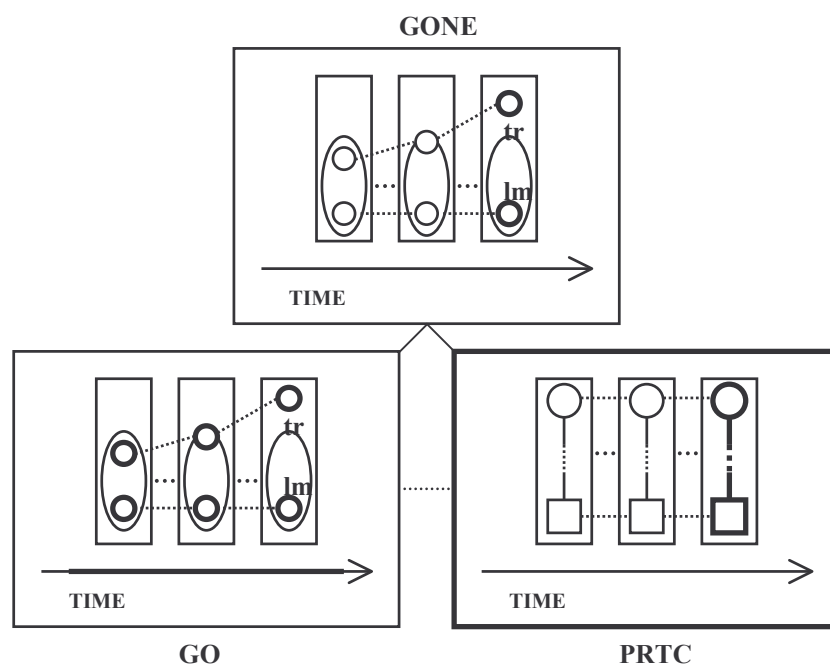
Le pari de ce type d'approche est que la structure conceptuelle contraint suffisamment l'expression langagière pour que l'une puisse être retrouvée à partir de l'autre. En particulier, les différentes langues sont supposées structurées de manière à exprimer sous forme grammaticale les composants conceptuels les plus fondamentaux. Par exemple, de nombreuses langues encodent des indications de nombre dans la morphologie. On peut penser que ces indications, qui se limitent à des distinctions singulier/duel/pluriel, jouent un rôle fondamental dans la structuration conceptuelle. Ce n'est pas le cas, en revanche, des couleurs ou des indications numériques fines comme les objets groupés "par treize" ou "par quarante-sept". Ainsi, l'étude de la grammaire des langues constitue un moyen de comprendre les composants fondamentaux du système conceptuel. Inversement, le modèle proposé pour le système conceptuel est mis à l'épreuve de sa réalisation dans l'expression de la pensée dans telle ou telle langue : la recherche de formalismes schématiques appropriés pour représenter la structure conceptuelle des énoncés a pour objectif de prédire, dans une mesure aussi large que possible, la forme langagière de ces énoncés.

Ce qui distingue les structures schématiques des structures algébriques ou logiques est qu'elles permettent d'exprimer des relations topologiques comme le voisinage ou l'inclusion, ainsi que des relations cinématiques ou dynamiques. Ce ne sont pas des représentations analogiques, puisqu'elles n'encodent ni la magnitude, ni les aspects qualitatifs comme la forme, la couleur, *et cætera*. Parmi toutes les propriétés et les relations que nous pouvons percevoir, celles qui sont retenues pour le système conceptuel sont celles qui découlent de l'observation de la grammaire.

[Cognitive grammar] claims that grammar itself serves an "imagic" function and that much of it has a figurative character. Grammar (like lexicon) embodies conventional **imagery**. By this I mean that it structures a scene in a particular way for purposes of linguistic expression, emphasizing certain facets of it at the expense of others, viewing it from a certain perspective, or construing it in terms of a certain metaphor.

(LANGACKER 1987 [64] p. 39)

La structure conceptuelle d'un énoncé apparaît comme la schématisation d'une scène. Cette schématisation constitue à la fois une simplification et un enrichissement : les propriétés métriques et qualitatives sont laissées de côté, au profit de relations explicitées comme le focus ou le rapport figure - fond, qui se traduisent directement dans la structure grammaticale. De la même manière, la structure grammaticale comporte des indications de point de vue. Grâce aux indications grammaticales liées au temps et à l'aspect, la même situation peut être considérée globalement, par son résultat ou par sa phase d'accomplissement. D'autres aspects de la richesse de la structure conceptuelle se révèlent dans le choix actif - passif, ou encore dans la forme du double accusatif : *he gave Mary a letter*, qui met en valeur le résultat de l'action alors que la forme prépositionnelle : *he gave a letter to Mary*, met le parcours accompli au premier plan. L'exemple suivant illustre le rôle de la forme du participe passé dans l'établissement d'un point de vue³⁰.



Le schéma associé au participe *GONE* diffère de celui du verbe *GO* par l'accent mis sur la partie terminale du processus. La construction grammaticale *GONE = GO + PRTC* se traduit par une combinaison au niveau des structures conceptuelles, dans laquelle la marque de participe provoque un déplacement du focus sur le résultat du processus associé à *GO*. Grâce à ce type de mécanismes de superposition, la combinaison des mots provoque la création de structures qui peuvent être davantage que la simple juxtaposition formelle de leurs composants.

Dans le type d'approche que nous considérons dans cette section, le système conceptuel est décrit à partir d'éléments qui sont supposés structurer notre fonctionnement cognitif. L'hypothèse est que cette structuration est suffisamment contraignante pour diriger l'expression langagière. Autrement dit, les combinaisons de mots reflètent des combinaisons conceptuelles, que l'on cherche à représenter par des combinaisons de schémas. Dans un tel modèle, le système conceptuel contrôle en grande partie l'expression, sans laisser de place à un niveau syntaxique défini.

³⁰ D'après (LANGACKER 1987 [64]).

Within cognitive linguistics, semantics is the primary component (which, in the form of conceptual representations, exists before, both phylogenetically and ontogenetically, syntax is fully developed). The structure of the semantic schemas puts constraints on the possible grammars that can be used to represent those schemas. (GÄRDENFORS 2000 [42] p. 165)

L'idée selon laquelle les mots et les relations entre les mots au sein d'un énoncé véhiculent des relations autres que purement algébriques nous semble fondamentale. En particulier, les mots peuvent exprimer des relations spatiales et topologiques que des formalismes purement symboliques ne peuvent pas représenter de manière naturelle. De ce point de vue, le courant de la linguistique cognitive ouvre un champ de réflexion qui nous semble incontournable. Nous devons cependant nous démarquer sur deux points.

Le premier concerne le statut de la syntaxe à laquelle ces modèles refusent toute autonomie. Les contraintes de proximité dans la liaison entre syntagmes, qui entraînent les phénomènes interprétés comme des mouvements, ou encore les contraintes positionnelles qui limitent la co-référence, semblent difficiles à expliquer par des propriétés purement sémantiques. De même, l'existence d'un jugement syntaxique, qui permet d'accepter des phrases comme bien formées ou de rejeter des phrases comme mal formées indépendamment de la signification qu'on leur accorde, serait mystérieuse si la bonne formation était subordonnée à la seule sémantique. L'ambition qui vise à éliminer du domaine de la syntaxe ce qui relève en réalité de la sémantique est louable. Nous ne souscrivons pas, cependant, à une extrapolation de ce programme qui poserait en principe l'inexistence de toute contrainte purement syntaxique. Plus généralement, l'objectif qui tend à expliquer l'ensemble des phénomènes liés à une langue par des mécanismes cognitifs de haut niveau, comme les schémas graphiques précédents, a ses limites. Il néglige le fait qu'une langue comporte de nombreux aspects non nécessaires résultant de l'habitude et de l'apprentissage par cœur. Par exemple, on dit en français *se rapprocher de la porte*, alors que dans des langues comme l'anglais ou le persan, on exprime l'équivalent de *se rapprocher à la porte*. Une interprétation cognitive consisterait à dire que le francophone mesure la distance qui reste à parcourir à l'aide d'un vecteur partant de la porte : *je suis à deux mètres de la porte*, alors que le persanophone mesurerait la même distance à l'aide d'un vecteur arrivant à la porte. Pourtant, rien n'indique, dans les paraphrases ou dans les inférences, que ces deux personnes utilisent des concepts différents. De même, lorsque le germanophone dit l'équivalent de *j'aide à mon ami* ou *je lui aide*, rien n'indique qu'il emploie un concept distinct de ce que le verbe français exprime. En revanche, si les sujets apprennent par cœur la préposition qu'il convient d'employer avec le verbe *se rapprocher* ou le cas à employer avec le verbe *aider*, l'explication cognitive peut très bien être cantonnée à une portée étymologique. Si tel est le cas, la connaissance qui permet d'employer correctement le verbe est simplement de type phonologique, un peu comme le sont les tables de multiplication. Notre souci, avec ces remarques, est donc de nous démarquer du projet qui consisterait à demander à une théorie des concepts d'expliquer la totalité des phénomènes de langue.

Le deuxième point sur lequel nous pouvons critiquer les approches courantes de la linguistique cognitive est lié à l'origine des structures mentales conférées aux énoncés. Dans le premier exemple cité, la différence des représentations schématiques associés aux phrases *the ball kept rolling* et *the log kept lying* ne doit rien à la grammaire et n'est due qu'à la différence sémantique entre les verbes *roll* et *lie*. Ainsi, comme dans les modèles de sémantique lexicale, les mots, même lorsqu'ils appartiennent à la classe ouverte, comme les verbes *roll* ou *lie*, se voient attribuer une structure conceptuelle afin que leur comportement dans les représentations schématiques puisse être distingué. Le risque est donc de devoir conférer une structure schématique précise à tous les éléments du lexique et non aux seuls éléments grammaticaux. Or une telle démarche est problématique, comme cela sera montré plus loin

(CF. CHAPITRE 7). De plus, même si l'on accepte l'idée de schématiser la totalité du lexique, le résultat n'est pas garanti. La position des forces antagonistes, dans les deux mêmes exemples, est liée à des inférences complexes sur la scène, qui veulent, dans un cas, que le vent contribue au mouvement non spontané de la balle, alors que, dans l'autre cas, le rebord s'oppose au mouvement spontané de la bûche. La méthode ne montre pas comment de telles connaissances peuvent être enfermées dans des représentations schématiques. Ce n'est pas en appliquant un schéma prédéfini qu'il aurait stocké dans l'entrée lexicale des mots *wind*, *ridge*, *roll* ou *ball* que le sujet parvient à savoir que le vent peut contribuer à un mouvement de rotation, alors qu'un rebord s'y oppose. Il semble plutôt que la direction des forces, dans ces exemples, soit le résultat d'un raisonnement, non modélisé, sur la scène, et qu'un tel raisonnement fasse appel à l'expérience de l'individu qui a vécu des situations du même type. Ainsi, l'approche par représentations schématiques ne saurait épuiser la question du sens. Elle incorpore certaines des capacités qui nous permettent de raisonner sur des scènes complexes, sans proposer de limite franche entre ce qui appartient au sens des mots et ce qui relève des capacités inférentielles générales comme le raisonnement spatial. Nous aurons l'occasion de réexaminer certaines propositions de la linguistique cognitive lorsque nous ferons nos propres suggestions sur la nature des représentations associées aux mots (CF. CHAPITRE 9).

Conclusion

Les différentes sections de ce chapitre peuvent être contrastées sur la question du lien entre la combinaison des mots et celle des concepts. Alors que la grammaire générative défend l'idée d'une syntaxe autonome, présentée comme une mécanique chargée d'apparier le sens et les séquences phonologiques, la linguistique cognitive se focalise sur la combinaison des concepts, avec idéalement un contrôle de l'ordonnancement des mots. Entre ces extrêmes, les modèles de la sémantique formelle, de la sémantique lexicale et de la sémantique procédurale proposent diverses solutions pour interfacer leur vision du système conceptuel avec le système syntaxique, allant parfois jusqu'à proposer une correspondance stricte entre les deux niveaux.

Conclusion

Cette deuxième partie de notre travail nous a permis d'étudier différentes caractérisations qui ont pu être faites concernant les concepts. Les trois chapitres de la partie correspondent aux trois interfaces que le système conceptuel entretient avec le reste de la cognition : la perception, le langage et le raisonnement. Il en est résulté des réponses contrastées, souvent incompatibles. Les propriétés qu'un concept doit posséder pour être ancré dans des relations d'évocation perceptives, pour être impliqué dans des inférences, et pour entrer dans des combinaisons dirigées par la grammaire d'une langue ne sont pas les mêmes. Il en résulte une image frustrante, selon laquelle les concepts semblent être des entités cognitives à la fois complexes et insaisissables.

La prochaine partie va se concentrer sur la caractérisation des concepts en tant que représentations internes du système conceptuel. Ce système conceptuel possède ses propres contraintes, liées notamment à l'explication des phénomènes de systématisme. Nous aurons l'occasion de constater que les solutions évoquées dans ce qui précède sont non seulement difficiles à concilier, mais présentent des défauts, parfois rédhibitoires, sur le plan de la modélisation cognitive.

Partie 3 :
Le dilemme du mentalais

Introduction

La notion de concept est parfois introduite pour expliquer le fait que les pensées sont communicables par le langage. De ce fait, les concepts acquièrent une stabilité et une universalité qui rappelle celle des mots d'une langue. Si une pensée relative au chien Fido peut être communiquée par l'emploi du mot Fido, c'est parce que le concept FIDO préexiste à la prononciation ou l'audition du mot. De plus, la possibilité de communication des pensées laisse supposer que les mêmes concepts existent dans des cerveaux différents. Cette manière d'aborder la question de la signification et de sa transmission conduit inévitablement à la notion de mentalais, une sorte de langage mental permettant la formation de la pensée.

Les pensées ne sont pas constituées, bien entendu, de suites de concepts isolés. Ces concepts sont le plus souvent assemblés. Cette observation conduit à concevoir le mentalais comme un système combinatoire, ce qui vient renforcer l'analogie avec le langage. La combinatoire des pensées serait ainsi le miroir de la combinatoire des mots, et le calcul du sens consisterait à assurer ce fonctionnement parallèle. L'acte de donner un sens à une phrase revient, dans un tel schéma, à agencer les concepts de telle manière que la construction obtenue permette d'assurer la transmission de la pensée correspondante, *id est* de produire les inférences requises.

Le langage naturel, avec sa phonologie et sa syntaxe ne serait qu'un moyen de rendre les constructions du langage mental apparentes pour les besoins de la communication. Par ailleurs, ces constructions sont supposées être le médium sur lequel opèrent les processus de raisonnement. Le mentalais constitue ainsi une idée naturelle pour assurer l'interface entre, d'une part, le système conceptuel, et d'autre part, le langage et le raisonnement. En tant que tel, il présente l'avantage considérable d'expliquer l'ensemble des phénomènes systématiques liés au langage.

Un intérêt supplémentaire de l'hypothèse de combinaison interne au mentalais est qu'elle permet d'expliquer la précision de la communication. Les êtres humains semblent capables, non seulement d'évoquer des pensées similaires aux leurs chez leurs interlocuteurs, mais aussi de les communiquer de manière arbitrairement précise quand cela est nécessaire. Une façon apparemment simple de résoudre ce problème de précision consiste à présupposer l'existence d'un dictionnaire mental. Si les concepts peuvent être paraphrasés par des descriptions exprimées à l'aide d'autres concepts, on parvient à expliquer que le sens ne dépend pas de l'individu ni de son expérience, dès lors que le dictionnaire a été correctement acquis. À l'instar de ce qui se fait en mathématique, il suffit de communiquer la description d'un concept pour assurer sa duplication dans un autre esprit, avec une précision parfaite.

Malgré cet ensemble de bonnes propriétés, la notion de mentalais pose des problèmes fondamentaux. Ces problèmes, comme nous allons le voir, ont trait d'une part à la structure interne des éléments du mentalais, et d'autre part à leurs combinaisons. Nous parvenons ainsi à un dilemme, par lequel le modèle du mentalais semble être à la fois nécessaire par les problèmes qu'il résout et, cependant, impossible à défendre en raison des difficultés insurmontables auxquelles il conduit. Dans ce qui suit, nous allons remettre en question la quasi-totalité des propriétés du mentalais que nous venons d'évoquer, sans pour autant renoncer à l'idée que les mots ont une contrepartie sémantique sur laquelle opèrent les mécanismes de construction du sens.

Chapitre 6 :
Le calcul du sens

Introduction

L'une des difficultés principales de la modélisation sémantique du langage vient du fait que les situations qu'il permet de décrire appartiennent à un monde perçu comme continu : deux pommes ne sont jamais identiques. Or, il emploie pour cela un code qui, par de nombreux aspects, est un code combinatoire digital¹. On retrouve ce caractère digital dans la phonologie, la morphologie et la syntaxe. Les représentations sémantiques sont elles de type digital, à l'image du reste du langage, ou sont elles analogiques, à l'image de notre interaction avec l'environnement sur laquelle portent les énoncés ? L'existence de phénomènes systématiques a conduit de nombreux auteurs à choisir la première option, en considérant que les représentations mentales évoquées par les mots étaient, elles aussi, organisées de manière à former un système combinatoire digital.

Parmi les aptitudes cognitives humaines, la capacité de donner un sens aux énoncés de notre langue est au centre de ce qui nous distingue des autres espèces animales. Peut-elle se modéliser comme le fonctionnement d'une machine de Turing ?

6.1. Le code du langage

Les êtres humains utilisent le langage, instancié dans les différentes langues, comme un code leur permettant de communiquer. Ce code présente la propriété remarquable d'être, par bien des aspects, un code tout-ou-rien. Dans un contexte donné, nous nommons la situation ou l'objet auquel nous sommes confrontés à l'aide d'une expression langagière. Nous pouvons ainsi désigner l'objet devant nous par le mot *table*. Ensuite, nous pouvons argumenter avec nos interlocuteurs pour démontrer que c'est bien cette expression qui est appropriée, tandis que d'autres ne le sont pas. Si l'on observe des personnes en train de discuter dans les conditions normales et spontanées, sans qu'ils soient soumis à des obligations de type psychologique ou sociologique, il arrive fréquemment qu'on les voie argumenter sur la bonne désignation d'un objet ou d'une situation. Est-ce bien une table, ou doit-on parler de bureau ?

Certes le code langagier est extrêmement flexible en fonction du contexte dans lequel il est utilisé et de l'individu qui l'utilise. Les objets qu'une étudiante en sciences cognitives nommera par le mot *table* ne sont pas exactement les mêmes que ceux qu'un antiquaire désignera par ce mot. L'antiquaire dispose de nombreux autres mots, comme *guéridon*, *console*, ou *desserte*, pour désigner les choses que l'étudiante confond sous le vocable *table*. De plus, si l'étudiante n'est pas francophone d'origine, elle pourra utiliser le mot *table* pour nommer ce qui, pour un Français, ne peut être désigné que par le mot *bureau*, voire par les mots *pupitre* ou *secrétaire*, simplement parce que cette distinction n'est pas habituelle dans sa langue maternelle. Pourtant, les mêmes personnes, lors d'un pique-nique, n'auront aucun problème pour trouver l'accord tacite qui consiste à désigner un tronc d'arbre scié par le mot *table*.

Cette variabilité dans l'emploi des mots ne signifie pas que le langage est intrinsèquement flou. Lorsque les interlocuteurs se mettent d'accord, dans chacun de ces contextes, l'objet perçu est, ou n'est pas, une table. Il ne devient pas une table à 76%. Le langage nous offre un code de communication qui peut être arbitrairement précis, dans chaque contexte, pour satisfaire aux besoins de la communication. Cette précision nous permet

¹ À titre de comparaison, le code de communication des abeilles est également combinatoire, mais analogique. L'angle que la danse produit avec la verticale dans l'obscurité de la ruche reproduit, de manière graduelle, l'angle entre la direction du soleil et la direction à prendre pour trouver la source de nourriture.

notamment d'amener notre interlocuteur à trouver la bonne référence. On l'obtient à l'aide de mots appropriés, mais aussi, lorsque ceux-ci font défaut, en construisant des syntagmes comme *table à écrire*, ou *table à dessiner des plans*. Ainsi, le langage apparaît comme un code par lequel nous pouvons communiquer à propos d'un nombre incalculable de situations, avec le plus souvent une précision qui satisfait aux besoins de l'interaction. Ce code est construit à partir d'un ensemble d'éléments de base, le lexique, qui est arbitraire, discret, dispersé, et donc digital.

Le caractère arbitraire du lexique apparaît de manière manifeste lorsque l'on compare les langues entre elles. Il est difficile de trouver de ressemblance entre les mots *pierre*, *sangue*, *stone*, *hajar*, désignant les mêmes objets perçus en français, en persan, en anglais et en arabe, et une perception quelconque que l'on peut avoir d'une pierre. Noter que le caractère arbitraire du lien symbolique n'est pas systématique dans tous les codes de communication. Dans les langues signées, certaines unités lexicales sont construites à partir de la ressemblance entre le geste et l'objet ou la situation qu'il désigne (MCNEILL 1992 [72]). Les individus entendant ne parlant pas la même langue utilisent spontanément des gestes iconiques pour se comprendre. Mais même dans les langues signées, l'iconicité se perd par le jeu de l'évolution linguistique qui restaure l'arbitraire du signe (MCNEILL 1992 [72]).

Le lexique présente en outre la caractéristique d'être discret. Cette propriété est remarquable si on la rapproche de la variation des données de la perception. Lorsque nous nommons certains objets perçus par les mots *gravillon*, *caillou*, *pierre*, *rocher*, nous imposons une échelle discrète là où notre perception nous permet de discerner une variété beaucoup plus riche.

Le lexique est non seulement arbitraire et discret, mais également dispersé. Une série de mots comme *gravillon*, *galet*, *caillou*, *pavé*, *moellon*, *pierre*, *rocher* ne réalise pas un échantillonnage uniforme de la continuité perceptive. Ces mots renvoient typiquement à des objets perçus qui ne se différencient pas seulement par leur taille, mais aussi par tout un ensemble de caractéristiques, si bien qu'ils se trouvent dispersés dans un ensemble de haute dimension. Le gravillon est typiquement anguleux, au contraire du galet qui est lisse ; le galet est associé à la rivière, le gravillon à la route ; *et cætera*. Du fait de cette dispersion, le lexique réalise un échantillonnage dépourvu de toute régularité de l'espace de la perception. Du fait de la dispersion du code, le langage ne conserve pas les rapports de voisinage fournis par la perception. Il n'existe pas de dépendance analogique, continue ou discrète, entre les éléments du lexique et les entités perçues qu'elles désignent.

Par ses aspects discrets non analogiques, le langage apparaît comme un code digital. Aucune mesure de distance entre les mots ne reproduit la distance entre les objets correspondants de la perception. Ce caractère digital s'ajoute à un aspect combinatoire pour former ce qui ressemble à une mécanique.

6.2. Le jugement sémantique

Le langage ne consiste pas en une suite de mots décorrés. À partir du lexique et de la grammaire, nous créons des expressions qui, elles-mêmes, peuvent renvoyer à une entité ou une situation perçue. Face à une collection de pierres précieuses, l'individu qui est à court de mots utilisera des syntagmes comme *l'émeraude qui ressemble à une olive*, ou *le rubis en forme de triangle*. Alors que le lexique est constitué d'une liste finie d'éléments, l'ensemble des phrases possibles est infini. Ce qui nous intéresse ici est comment déterminer l'ensemble des phrases "sensées", autrement dit les phrases auxquelles un locuteur compétent pourra accorder une signification. Cette question présuppose l'existence d'un jugement d'acceptabilité sémantique. La démarche est habituelle dans les sciences du langage.

Le même type de problème se pose aux niveaux phonologique, morphologique, syntaxique et pragmatique, car il existe un jugement d'acceptabilité pour chacun de ces niveaux : un locuteur compétent sait reconnaître une prononciation acceptable, un mot composé possible, une phrase bien formée et une intervention pertinente. Le jugement d'acceptabilité divise, idéalement, l'ensemble des énoncés en deux classes, les énoncés corrects et ceux qui ne le sont pas. Le travail de modélisation consiste alors à identifier ces classes de la manière la plus économique possible. Dans le cas du jugement sémantique, la tâche de modélisation va consister à tenter de découvrir un nombre réduit de principes permettant de prédire qu'un énoncé aura un sens ou n'en aura pas. On pourrait reprocher au jugement d'acceptabilité sémantique d'être très peu binaire, en tout cas beaucoup moins binaire que le jugement syntaxique. Même une phrase comme le jardin de la porte mange le ciel peut, à condition qu'on y mette un peu de temps, recevoir un sens. Dans ce cas, le modèle devra expliquer le caractère graduel de l'acceptabilité.

L'ensemble Σ des expressions sémantiquement acceptables possède certaines propriétés fondamentales qu'il s'agira d'expliquer. Tout d'abord, il est potentiellement infini. Il ne semble pas exister de borne supérieure, hormis celles qui sont liées aux contraintes matérielles, au nombre de significations différentes qu'un être humain peut former et exprimer par le langage. Ensuite, il possède des régularités systématiques. Une première forme de systématicité est la systématicité de substitution. Si la phrase Jean aime Marie est interprétable, alors la phrase Paul aime Marie l'est aussi. Pour les défenseurs de la systématicité sémantique, on ne trouvera pas d'être humain qui serait capable de comprendre l'une de ces deux phrases en étant strictement incapable de comprendre l'autre (FODOR & PYLYSHYN 1988 [41]). L'ensemble Σ des expressions interprétables semble être clos pour certaines opérations de substitution qui préservent la "classe sémantique" des objets substitués. Un deuxième aspect de la systématicité concerne la systématicité d'extraction. Il existe un lien obligatoire entre l'interprétation d'une expression et celle de ses composants. Si nous pouvons comprendre correctement la phrase le frère de Jacques aime Marie, alors nous pouvons comprendre le syntagme le frère de Jacques, de même que le mot frère (FODOR & PYLYSHYN 1988 [41]). Ajoutons à cela une systématicité de négation, précisant que si une phrase appartient à Σ , alors sa négation y figure également, de manière systématique.

Le caractère infini de Σ et ses aspects systématiques constituent un défi pour la modélisation. Il est exclu que Σ résulte d'un apprentissage par cœur. La production des expressions sensées ne peut donc que résulter d'un calcul. Pour déterminer la nature de ce calcul, il est utile d'examiner comment les propriétés de Σ peuvent être reproduites.

L'infinité de Σ requiert que la génération du sens soit due à un mécanisme productif. La productivité d'un mécanisme signifie qu'il peut engendrer un ensemble non borné de formes différentes. La productivité est rendue possible, notamment, grâce à des mécanismes récursifs. L'intérêt de cette notion est de souligner la possibilité que des ensembles potentiellement infinis puissent être engendrés par des moyens finis.

Pour expliquer la systématicité d'extraction, on suppose généralement que les mécanismes qui engendrent Σ sont compositionnels. Selon le principe de compositionnalité, le sens attaché à une expression peut être entièrement calculé à partir du sens des mots qu'elle contient (FODOR & PYLYSHYN 1988 [41]). On peut exprimer cette idée par l'équation suivante.

$$\mathbf{s}(m_1, m_2, \dots, m_k) = \mathbf{f}(s_1, s_2, \dots, s_k), \text{ où } s_i = \mathbf{s}(m_i)$$

La fonction $\mathbf{s}()$ est la fonction qui attache un sens une expression langagière. Notamment, elle associe un sens s_i à un mot m_i . La fonction $\mathbf{f}()$ produit le sens d'une

expression en fonction, entre autres, de sa structure grammaticale. On comprend ainsi que si une expression possède une signification, autrement dit si elle appartient à Σ , alors ses éléments appartiennent nécessairement à Σ , puisque la fonction $\mathbf{s}()$ est définie sur eux². La compositionnalité peut fournir ainsi une explication naturelle à la systématique d'extraction.

Une conception peu plausible de la construction sémantique voudrait que la fonction $\mathbf{f}()$ ne commence à être calculée qu'une fois que tous ses arguments sont disponibles. Il est manifeste que dans la plupart des situations, les sujets commencent l'élaboration d'un sens bien avant d'atteindre la fin d'une phrase, comme on peut sans doute le vérifier dans le présent paragraphe. De plus, l'agencement hiérarchique des syntagmes laisse supposer que le calcul du sens est effectué de proche en proche (ce qui n'exclut pas des retours en arrière éventuels), chaque syntagme étant évalué pour lui-même. Par exemple, pour une structure syntaxique simple comme $(m_1 m_2) m_3$, on pourrait avoir idéalement deux étapes de calcul suggérées par l'écriture suivante.

$$\begin{aligned} \mathbf{s}(m_1, m_2, m_3) &= \mathbf{f}(\mathbf{s}(m_1, m_2), \mathbf{s}(m_3)) \\ &= \mathbf{f}(\mathbf{f}(s_1, s_2), s_3) \end{aligned}$$

Ce genre d'écriture confère un certain caractère sériel à la construction du sens, hérité de la sérialité structurelle du canal linguistique³. Si l'on accepte que le sens est ainsi, dans une certaine mesure, construit par couches successives, on comprend mieux la possibilité même de la systématique d'extraction. Cette propriété peut être re-formulée d'une manière un peu différente, en écrivant l'expression $\mathbf{f}(s_1, s_2, \dots, s_k)$ comme une fonction de la variable s_k .

$$\mathbf{s}(m_1, m_2, \dots, m_k) = \mathbf{g}_{s_1, s_2, \dots, s_{k-1}}(s_k)$$

Du fait du cloisonnement interne au calcul du sens, une substitution lexicale du mot m_k par le mot n_k n'affectera qu'une partie des arguments de la fonction $\mathbf{f}()$ et ses effets resteront, dans beaucoup de cas, circonscrits. On peut s'attendre, de ce fait, à de nombreux phénomènes de systématique, car le sens construit, sa validité et les inférences qu'il déclenche ont des chances de persister en partie après la substitution. Par exemple, le sens de la phrase Jean aime une artiste est présenté comme le résultat de l'application d'une fonction sémantique au concept ARTISTE. Cette fonction est une construction associée à l'expression linguistique Jean aime ... qui prend le concept ARTISTE comme argument. L'idée, selon laquelle le sens est construit par étapes, se traduit ici par le fait que l'expression $\mathbf{g}_{s_1, s_2, \dots, s_{k-1}}()$, en tant que fonction, est un objet dont le résultat de l'application à la variable s_k n'est pas erratique. Autrement dit, une substitution affectant la variable s_k aura des conséquences limitées, ce qui engendre des phénomènes de systématique. On peut établir une analogie entre ce phénomène de régularité, due à l'imbrication du traitement sémantique, et la propriété mathématique de continuité des fonctions. Lorsque l'on réécrit une fonction $\mathbf{f}(x, y)$ sous la forme $\mathbf{g}_x(y)$, transformant l'un des arguments en paramètre, on passe de l'équation d'une surface à celle d'une famille de courbes. Ce faisant, on privilégie une direction en chaque point de la surface. Si les courbes $\mathbf{g}_x(y)$ sont dérivables, on peut localement prédire leur comportement. C'est cette idée du comportement prédictif de la fonction $\mathbf{f}(x, y)$ par rapport à l'un de ses arguments qu'on retrouve dans l'idée de systématique de substitution.

Le fait que Σ respecte la systématique de substitution peut s'expliquer par le fait que la fonction $\mathbf{f}()$ porte sur des classes de significations et non les significations elles-mêmes. Si

² L'hypothèse implicite est que $\mathbf{f}(m_i) = s_i$, puisque pour un mot isolé la fonction \mathbf{f} s'instancie par l'équation $\mathbf{f}(m_i) = \mathbf{s}(m_i)$.

³ Cet aspect sériel est certainement loin d'être strict. L'idée est plutôt de critiquer la possibilité d'un parallélisme radical. La sérialité provient non seulement du traitement successif des syntagmes imbriqués, mais également et plus simplement de l'ordre des mots dans l'énoncé.

l'on substitue un élément par un élément de la même classe, le mécanisme continue de s'appliquer. Lorsque les chimistes ont découvert des régularités comme celle représentée par l'équation *acide + base* → *sel + eau*, ils étaient en droit de l'expliquer par le fait que le mécanisme sous-jacent n'était pas sensible aux réactifs eux-mêmes, mais seulement à leur caractère acide ou basique. De même, les régularités de substitution laissent supposer que les mécanismes compositionnels ne sont sensibles qu'à la classe sémantique des mots. Cette idée peut être illustrée par la réécriture suivante de l'expression $\mathbf{f}(s_1, s_2, \dots, s_k)$.

$$\mathbf{s}(m_1, m_2, \dots, m_k) = \mathbf{g}_{s_1, s_2, \dots, s_{k-1}}(\mathbf{t}[u]), \text{ où } \mathbf{s}(m_k) = \mathbf{t}[u] \text{ pour } u \in D_{\mathbf{t}[\]}$$

Selon cette équation, la fonction $\mathbf{g}_{s_1, s_2, \dots, s_{k-1}}()$ accepte comme argument des éléments qui partagent une certaine propriété. Cela revient à postuler l'existence d'une forme de typage⁴, que nous représentons à l'aide de la fonction $\mathbf{t}[\]$. Le typage permet de prédire l'interprétation systématique de l'adjectif foncé lorsqu'il s'applique à des noms de couleur comme *vert*, *jaune*, *émeraude*, *et cætera*. Dans ce cas, la fonction $\mathbf{t}[\]$ filtre les seuls concepts de couleur, en laissant sa variable libre de parcourir tout l'ensemble des couleurs. De même, on comprend pourquoi un lien systématique existe entre le fait que la phrase *Jean aime Marie* soit interprétable et le fait que la phrase *Paul aime Marie* le soit également. Les entités désignées par le prénom *Jean* et par le prénom *Paul* partagent un ensemble de propriétés, notamment celle d'être humain, qui sont censées garantir l'existence d'une interprétation. Tout se passe comme s'il existait un patron de pensée, du genre <humain> aime Marie, permettant à n'importe quelle instanciation du type <humain> de conduire à une pensée possible. Pour illustrer ce phénomène lié à la compositionnalité, il est utile de considérer un contre-exemple. Le sens d'une expression idiomatique comme *retourner sa veste*, dans l'esprit de la plupart des locuteurs, n'est pas lié au sens du mot *veste*. Ce défaut de compositionnalité se traduit par une lacune dans la systématisme : le fait de pouvoir comprendre l'expression *retourner son manteau* (dans son sens matériel, par exemple pour un manteau réversible) n'est pas systématiquement lié au fait de comprendre l'expression *retourner sa veste* (dans son sens figé de brusque changement d'opinion ou de camp). La nature de ce genre d'exception permet de comprendre le cas général, qui est que la compositionnalité est un excellent moyen d'expliquer les phénomènes de systématisme. Pour prendre un autre contre-exemple, l'ensemble des recettes gastronomiques est engendré par un mécanisme productif, la combinaison d'ingrédients. Cependant, ce mécanisme n'est pas connu pour être compositionnel : la production des recettes gastronomiques, à l'exclusion des autres, procède par essais - erreurs, car on ne connaît pas de fonction $\mathbf{f}()$ permettant de prédire le goût du résultat à partir de celui des ingrédients.

Le principe de compositionnalité est parfois jugé insuffisant pour expliquer les phénomènes de systématisme. Pour assurer systématiquement le lien entre la signification s_i d'un mot m_i et une représentation calculée comme $\mathbf{f}(s_1, \dots, s_i, \dots, s_k)$, on peut considérer que cette représentation doit conserver une mémoire de l'élément s_i . Le moyen le plus simple est d'exiger que cette représentation contienne, dans un sens matériel, ses composantes. Ainsi, selon le principe de constituance, l'accès systématique à la représentation sémantique associée au prénom *Jean* à partir de la représentation sémantique associée à la phrase *Jean aime Marie* est assurée par le fait que la première fait partie intégrante de la seconde (FODOR & PYLYSHYN 1988 [41]). La présence effective des constituants au sein des représentations garantit leur accessibilité systématique. On peut traduire ce principe par le fait qu'il s'agit de représentations récursives : des représentations qui contiennent d'autres

⁴ De manière étonnante, comme nous le verrons, certains auteurs qui mettent en avant ce type de systématisme refusent un tel typage de concepts, sans proposer d'explication alternative.

représentations de même nature. Le principe de constituance est une exigence forte, qui pose le problème du rapport entre le composé et ses constituants. Dans le cas de représentations résultant de l'interprétation des expressions du langage, les principes de compositionnalité et de constituance supposent que la structure de l'expression langagière se transporte dans la structure de la représentation mentale qui lui correspond. C'est ainsi que ces deux principes, constituance et compositionnalité, parviennent à rendre compte des phénomènes de systématique sémantique.

6.3. La combinatoire du langage

La question devient de savoir quel type de calcul est capable, pour engendrer un ensemble Σ doté des propriétés d'infinité et de systématique, d'être productif et compositionnel en produisant des représentations respectant le principe de constituance. Dans une certaine acception du mot "calcul", il s'agit de modéliser le comportement d'un système, le cerveau, à l'aide des mathématiques, en conservant une distinction radicale entre le modèle et le système. Le physicien ne prétend pas qu'une planète résout une équation différentielle pour déterminer sa prochaine position. De même, le chercheur en sciences cognitives qui s'inscrit dans la tradition cybernétique suppose que le système qui engendre les combinaisons langagières est de nature dynamique, régi par des équations d'évolution et des principes de minimisation d'énergie, sans croire que les calculs de son modèle sont réellement effectués par le système lui-même (VAN GELDER 1998 [105]). Les mécanismes qui sous-tendent le langage, comme tous les processus naturels, peuvent en principe être décrits comme des processus dynamiques, ne serait-ce qu'au niveau de leur implémentation neuronale. Cependant, la nature du langage permet d'espérer mieux. Le langage est caractérisé par des changements d'états discontinus et des unités hiérarchisées : phonèmes, mots, syntagmes. Il passe par un canal sériel, et son traitement présente de ce fait des aspects séquentiels. La syllabe et les syntagmes se prêtent à des analyses qui produisent des structures arborescentes par enchâssement. Tout ceci suggère fortement la possibilité d'un niveau de description qui se situe au-dessus du niveau dynamique. En ce qui concerne les aspects sémantiques, le fait que les phénomènes systématiques concernent des substitutions ou des extractions d'éléments plaide également en faveur d'un niveau de description que l'on peut qualifier de symbolique ou de computationnel, consistant en un calcul de type combinatoire.

Les techniques combinatoires, en mathématiques, regroupent les méthodes d'énumération et de génération d'objets finis discrets. Contrairement aux modèles dynamiques, les mécanismes symboliques qui expliquent un phénomène cognitif sont supposés réalisés dans le système qu'ils modélisent (VAN GELDER 1998 [105]). Le calcul de gradient ne prétend pas constituer une image réaliste du comportement de la bille qui roule sur une pente. Son mouvement peut être aussi modélisé par un calculateur analogique utilisant des courants électriques pour représenter les paramètres physiques. En revanche, tous les systèmes connus capables d'effectuer des calculs combinatoires sur des structures discrètes sont des variantes ou des spécialisations de la machine de Turing. En conséquence, si un aspect de la cognition se laisse modéliser par un mécanisme computationnel combinatoire, il est naturel de faire l'hypothèse que le cerveau effectue une combinatoire du même type.

Cette question de la nature combinatoire du calcul sémantique s'est trouvée au centre d'une polémique (FODOR & PYLYSHYN 1988 [41] ; FODOR & McLAUGHLIN 1990 [40] ; SMOLENSKY 1988 [97] ; SMOLENSKY 1990 [98]). Pour prétendre représenter le sens des expressions du langage, des mécanismes non combinatoires comme le connexionnisme devraient eux aussi pouvoir engendrer la systématique, ce que cette polémique mettait en discussion. Le point crucial est que la systématique ne peut être due au hasard des connexions

neuronales, si ces connexions sont supposées être directement responsables des liens qui relient le support matériel des différentes représentations mentales (FODOR & PYLYSHYN 1988 [41] ; FODOR & MCLAUGHLIN 1990 [40]). La propriété de systématisme exclut toute idée de lacune ou de panne partielle. Si la pensée correspondant à la phrase Jean aime Marie peut être produite, alors il y a 100% de chances que la pensée correspondant à la phrase Jacques aime Marie et la pensée correspondant au prénom Marie puissent être produites, non 99% ou 99.9%. Or, un dispositif qui représente les liens entre les constituants de ces représentations par des connexions matérielles ne peut assurer ces 100%. La raison fondamentale qui empêcherait les dispositifs connexionnistes d'assurer la systématisme serait leur manque de constituance. Sans constituance, les représentations ne peuvent que se déclencher les unes les autres, de manière causale, par les liens physiques qui les relient, alors que dans un dispositif doté de la propriété de constituance, une représentation a accès, de manière certaine, à ses propres composants. Le débat peut alors porter sur la question de savoir si le connexionnisme est en mesure d'assurer la propriété de constituance (CHALMERS 1993 [12]).

Dans cette partie, nous voulons explorer les conséquences de l'option combinatoire. Si le calcul du sens est un calcul combinatoire et que le fonctionnement du cerveau est, à cet égard, le même que celui d'une machine de Turing particulière, alors les propriétés de productivité, de systématisme, de compositionnalité et de constituance peuvent être garanties sans problème. Ce calcul combinatoire et les représentations symboliques sur lesquelles il porte constituent un authentique langage mental.

6.4. Le langage mental

Le projet d'un modèle combinatoire de la sémantique est de reproduire le jugement d'acceptabilité sémantique d'un énoncé quelconque d'après sa structure et la nature de ses éléments. Or, nous avons eu l'occasion de montrer qu'un projet analogue, concernant le jugement épistémique, pouvait être critiqué. La critique portait sur l'existence d'énoncés tautologiques non triviaux, dont la validité puisse être à coup sûr déterminée d'après les éléments de l'énoncé (CF. CHAPITRE 4). Si le jugement épistémique à propos de tout énoncé nécessite un accès à des connaissances externes, ne risque-t-on pas d'avoir une situation similaire dans le cas du jugement sémantique ? Par exemple, pourquoi, dans la plupart des contextes, la phrase Jean mange la pomme est sensée, alors que la phrase Jean mange la porte ne l'est pas ? Est-ce la signification du verbe manger qui constitue ici l'élément crucial, ou est-ce le fait contingent que les individus se nourrissent plus souvent de pommes que de portes ? Certaines approches logiques tentent de ramener le jugement sémantique à une question épistémique en notant que le premier énoncé, contrairement au deuxième, peut être vrai dans certains domaines d'interprétation (CF. CHAPITRE 5). Ce faisant, elles n'offrent en guise de signification qu'un critère relationnel contingent. Il est plus plausible de dire que, pour juger de l'acceptabilité sémantique des énoncés, nous effectuons un calcul combinatoire sur la signification des mots. Autrement dit, le jugement sémantique ne se réduit pas à une simple vérification dans une base de données. Lorsque nous jugeons les phrases Jean entre dans la pièce ou le calme entre dans la pièce comme sémantiquement acceptables, alors que, dans la plupart des contextes, la phrase le salon entre dans la pièce ne l'est pas, nous ne vérifions pas dans une base de données les occurrences des entités qui peuvent "entrer" dans d'autres entités ; nous effectuons un calcul sur la signification des mots.

L'approche combinatoire du calcul du sens ne se limite pas à reproduire la performance sémantique à l'aide d'un calcul formel. Elle prend ce calcul "au sérieux" et s'interroge sur sa nature. Si l'on adopte cette approche, il devient essentiel de préciser les propriétés de ce calcul

et les représentations sur lesquelles il porte. Si la production et la compréhension des énoncés se réalisent par une combinatoire de symboles, quelle est la nature de ces symboles, correspondent-ils à des pensées élémentaires, et ces pensées sont-elles en bijection avec les mots de la langue ?

Si la construction du sens s'effectue par un calcul combinatoire, nous avons affaire à un langage mental, parfois dénommé mentalais (FODOR 1975 [34]). L'existence du mentalais est une hypothèse forte sur le plan cognitif. Il ne s'agit pas simplement de dire que l'état du monde décrit par le langage naturel peut être aussi décrit par un langage formel plus précis, mais bien que l'esprit humain manipule des représentations structurées en utilisant des mécanismes combinatoires systématiques. L'hypothèse suppose que les opérations qui agissent sur ces représentations, par exemple pour produire des inférences, opèrent uniquement d'après les caractéristiques structurelles des représentations. On retrouve donc, au niveau du mentalais, ce qui, au niveau syntaxique, fait d'une langue comme le français ou le persan un mécanisme combinatoire productif et systématique. Contrairement aux langues, toutefois, le mentalais est supposé universel, ce qui permet de traduire un énoncé d'une langue à l'autre en préservant son sens. L'hypothèse de l'existence du mentalais se situe dans le cadre de l'approche fonctionnaliste en sciences cognitives. Si le calcul du sens est réellement effectué sur des bases combinatoires, alors peu importe le fait que ce soit un cerveau qui serve de substrat à ces computations (FODOR 1975 [34]). L'important est d'étudier la computation pour elle-même. Selon cette approche, l'essentiel pour comprendre l'élaboration du sens des énoncés est d'étudier le mentalais. Une bonne compréhension des lois de ce langage de la pensée, pour les partisans de l'hypothèse du mentalais, devrait permettre d'envisager sa reproduction matérielle à l'aide d'ordinateurs, qui seraient ainsi capables de produire exactement les mêmes interprétations que les êtres humains.

Le langage mental a pour fonction, entre autres, d'offrir une interprétation sémantique des expressions langagières, que nous avons appelée, dans la deuxième partie, une représentation conceptuelle. L'hypothèse du langage mental décrit la sémantique comme une double interprétation. Les expressions langagières sont interprétées dans le langage mental, sous forme de représentations conceptuelles. Mais le langage mental est lui-même un système symbolique, dont les constructions doivent être à leur tour interprétées. Cette interprétation se trouve dans le monde perçu. Ainsi, le contenu d'une expression langagière est une représentation conceptuelle ; le contenu d'une représentation conceptuelle, quant à lui, demeure dans notre interaction avec le monde.

Si le mentalais existe, ce que nous aurons l'occasion de discuter plus loin, alors une question essentielle est de déterminer la nature des représentations sur lesquelles il opère. Le modèle combinatoire suppose que ces représentations sont construites par la combinaison d'unités élémentaires, appelés concepts. Les concepts peuvent être définis par leurs propriétés représentationnelles et causales (CF. CHAPITRE 3). Les propriétés représentationnelles d'un concept constituent son contenu. Les propriétés causales d'un concept, quant à elles, sont liées aux mécanismes du calcul dans lesquels elle est impliquée, et lui confèrent une forme.

The representational properties of a creature's mental state are supposed to be sensitive to, and hence carry information about, the character of its environment. The causal properties of a creature's mental state are supposed to determine the course of its mental processes, and, eventually, the character of its behavior. [...] Concepts are the least complex mental entities that exhibit both representational and causal properties; all the others (including, particularly, beliefs, desires and the rest of the "propositional attitudes") are assumed to be complexes whose constituents are concepts, and whose representational and causal properties are determined, wholly or in part, by those of the concepts they're constructed from. (FODOR 1994 [36] p. 96)

Si les représentations conceptuelles sont constituées par les concepts qui les composent, comment définir la nature des concepts eux-mêmes ? Comme le système conceptuel ne peut, par définition, être observé directement, le problème peut être considéré à partir de ses trois interfaces : raisonnement, perception et langage.

On peut partir de l'idée selon laquelle les concepts sont les représentations conceptuelles les plus simples, sachant que les représentations conceptuelles sont définies comme les objets du langage mental qui ne peuvent pas recevoir de valeur de vérité. Il est habituel d'opérer une distinction entre les représentations propositionnelles qui peuvent recevoir une valeur de vérité, comme l'interprétation de la phrase le chat marche sur le mur, et les représentations conceptuelles proprement dites, qui ne le peuvent pas, comme l'interprétation du syntagme le chat noir du voisin. Cette distinction tire sa pertinence de l'hypothèse, admise dans certaines théories, selon laquelle seules les représentations propositionnelles peuvent être objet des attitudes propositionnelles, comme le désir ou la croyance, qui sont à l'œuvre dans les mécanismes de raisonnement (FODOR 1994 [36]). Notons que la distinction entre représentations propositionnelles et représentations conceptuelles est prédite par un critère linguistique : les syntagmes nominaux renvoient à des représentations conceptuelles, alors que les phrases sont interprétées comme des représentations propositionnelles. Cette distinction, cependant, ne nous aide pas à déterminer la nature des concepts en tant qu'éléments constitutifs des représentations conceptuelles, car elle n'a pas de conséquence sur la complexité structurelle des représentations correspondantes. Les deux formes de représentation, propositionnelle et conceptuelle, peuvent être incluses l'une dans l'autre. Par exemple, l'interprétation du syntagme l'homme qui croit que le chat marche sur le mur sera construite par la composition récursive des deux types de représentation. L'interprétation d'un syntagme nominal comme le chat noir du voisin, au même titre que celle de la phrase le chat marche sur le mur, doit nécessairement être une représentation conceptuelle composée, en vertu du principe de constituance, qui veut que le sens d'un syntagme contienne le sens des mots qui y apparaissent.

Si l'on cherche à caractériser les concepts à partir de leur contenu, on peut postuler que les concepts correspondent à des catégories perceptives. Dans le cadre d'un langage mental combinatoire, le problème peut être simplifié en partant de l'opposition entre concepts primitifs et concepts complexes. Dans certains modèles, les concepts primitifs sont ceux qui sont directement liés aux données de la perception, comme les couleurs, ou aux catégories de base de la cognition, comme l'espace ou le temps. Du fait de leur lien privilégié avec l'expérience, l'acquisition de ces concepts primitifs est souvent considérée comme un problème à part. Tous les autres concepts seraient des concepts complexes, ce qui signifie qu'ils sont acquis par des mécanismes inductifs généraux. Notons que le caractère primitif d'un concept peut être considéré comme différent de sa structure simple ou complexe. Rien n'interdit qu'un concept réputé primitif comme ROUGE corresponde à une représentation conceptuelle composée, alors qu'un concept considéré comme appris, par exemple CARBURATEUR, ne se révèle doté d'aucune structure.

Pour définir la nature des concepts élémentaires, il est tout aussi légitime de partir de la troisième interface, en considérant leur correspondance avec le langage. Nous sommes confrontés à deux systèmes symboliques, le langage proprement dit et le langage mental. Le langage mental, en tant que système symbolique, ressemble sur certains points au langage proprement dit. Il possède sa grammaire : les règles de combinaison conceptuelles. Il possède ses phrases, les constructions conceptuelles, et ses mots, les concepts. Les mots du langage naturel sont les plus petites unités susceptibles d'être à la fois interprétées et manipulées par la combinatoire grammaticale. Par analogie, les concepts apparaissent comme les plus petites unités du langage mental ayant un contenu et pouvant entrer dans des combinaisons conceptuelles. Il s'agit alors d'établir la correspondance entre ces deux systèmes. Selon cette

idée de correspondance, chaque fois que nous faisons un jugement sémantique, c'est-à-dire que nous donnons un sens à une expression langagière, nous lui associons une représentation conceptuelle⁵. La question de la nature des concepts élémentaires peut alors se poser ainsi : les mots du langage naturel, dans cette correspondance, sont-ils associés aux "mots" du langage mental ? Si les concepts lexicaux, associés aux mots du langage, sont définis de manière indépendante par leurs propriétés causales ou représentationnelles, cela n'a rien d'évident.

Le fait que les expressions langagières composés soient interprétés comme des représentations conceptuelles composées ne signifie pas que les mots qui les constituent correspondent à des concepts simples. Pour de nombreux auteurs, les concepts lexicaux correspondent, à leur tour, à des représentations conceptuelles composées possédant une structure récursive, puisque consistant en la combinaison d'autres représentations conceptuelles, un peu comme les molécules du vivant qui sont décrites comme des assemblages de molécules plus simples. Le dilemme est donc de savoir s'il existe une correspondance entre les mots et les concepts simples, ou si les concepts lexicaux sont pour la plupart des représentations conceptuelles composées. Dans le premier cas, les concepts verbalisés par les mots seront les éléments les plus simples du langage mental ; dans le deuxième cas, il faut admettre l'existence de concepts ineffables, indicibles.

Dans les deux chapitres qui suivent, nous allons tour à tour traiter les deux options. Nous considérerons tout d'abord l'approche moléculaire, qui voit les concepts lexicaux non comme des entités minimales, mais comme des représentations conceptuelles composées, récursives. Ensuite, nous aborderons l'option selon laquelle les concepts lexicaux sont d'authentiques représentations minimales du langage mental, des concepts atomiques.

Conclusion

L'approche symbolique computationnelle semble présenter un cadre approprié et prometteur pour expliquer les phénomènes systématiques de la sémantique du langage naturel. Le calcul du sens, effectué au sein d'un système conceptuel de ce type, consiste à former des expressions dans un langage mental. Dans le reste de cette partie, nous explorons les conséquences de cette hypothèse de manière à en cerner les atouts et les limites. Notre analyse se focalise sur une question précise. Si le système conceptuel possède la forme d'un langage mental, quelle est la nature des représentations de ce langage qui sont associées aux entités lexicales, dans une langue humaine quelconque ? Ces représentations, les concepts lexicaux, peuvent être soit des représentations moléculaires, possédant une structure récursive, soit des représentations atomiques, dépourvues de toute structure interne. Après avoir détaillé les différents moyens qui ont pu être proposés pour représenter les concepts lexicaux, que ce soit sous forme moléculaire ou sous forme atomique, nous montrerons qu'un certain nombre de paradoxes se présentent dans chaque cas. La liste de ces paradoxes nous conduira à penser que l'approche symbolique computationnelle, introduite dans le présent chapitre, conduit à une impasse.

⁵ Le fait que ces deux systèmes symboliques sont supposés coexister s'accorde bien avec l'hypothèse de modularité.

Chapitre 7 :
Concepts moléculaires

Introduction

Dans la deuxième partie nous avons eu l'occasion de mentionner des modèles du système conceptuel qui postulent des structures pour les concepts associés aux entités lexicales du langage naturel. Ces modèles sont issus de cadres théoriques différents, relevant de disciplines aussi variées que la philosophie, la psychologie, la linguistique ou la logique. Certains d'entre eux ont conduit à des techniques de représentation des connaissances utilisées en intelligence artificielle. Pour le sujet qui nous intéresse dans cette partie, nous cherchons à savoir dans quelle mesure ces techniques à base de représentations structurées peuvent être compatibles avec les contraintes posées par les principes de compositionnalité et de constituance. Ayant réalisé ce premier tri, nous nous demanderons si des structures ainsi représentées peuvent recevoir une plausibilité cognitive, en tenant compte des contraintes sur les mécanismes du raisonnement, du langage, et de l'apprentissage. En d'autres termes, la question sera de savoir si des représentations conceptuelles structurées comme celles qui sont couramment proposées peuvent avoir une contrepartie cognitive hébergée par un cerveau humain.

7.1. Motifs pour structurer les concepts lexicaux

Comme nous avons pu le constater dans la deuxième partie de ce mémoire, bon nombre d'auteurs font l'hypothèse que les représentations sémantiques associées aux mots, les concepts lexicaux, sont des entités structurées. Ils sont poussés en cela par deux motivations essentielles.

La première motivation est liée au rôle inférentiel des concepts (CF. CHAPITRE 4). Les concepts peuvent déclencher des inférences qui activent d'autres représentations mentales. Or, certaines de ces inférences apparaissent comme des inférences fortuites, liées au contexte et aux associations que le sujet établit sur le moment. Par exemple, le mot chimpanzé peut suggérer à certains, dans le contexte d'un zoo, le fait que l'animal désigné aimera les cacahuètes. Certaines inférences semblent cependant être d'une autre nature, apparaissant comme des inférences obligatoires. Un bon exemple est celui des taxonomies. Si une entité est classée comme CHIMPANZÉ, dans la plupart des contextes, elle sera également classée comme ANIMAL, ou encore comme MAMMIFÈRE, voire comme PRIMATE ou PONGIDÉ, si l'individu dispose de ces concepts. Il semble donc exister un lien obligatoire entre le concept CHIMPANZÉ et ceux d'ANIMAL, de MAMMIFÈRE, *et cætera*. Un autre exemple d'inférence obligatoire est celle qui mène du concept à ses attributs. Ainsi, on sait qu'une fleur est une entité fragile, qu'elle a une couleur et un parfum. En combinant taxonomie et attributs, on produit de nouvelles inférences. Par exemple, en apprenant qu'une pensée est une fleur, on apprend également qu'elle est fragile, qu'elle a une couleur, *et cætera*. Un concept peut aussi déclencher des inférences qui dépassent sa propre spécification. Dans des circonstances standard, nous savons, en entendant la phrase il a tué le chat, que le chat est mort et nous pouvons formuler la phrase le chat est mort. De plus, il semble que cette inférence ne soit pas différente de celle qui nous amène de la phrase il a tué le chien à la phrase le chien est mort. Autrement dit, l'inférence semble déclenchée par le mot tuer pour aboutir au mot mourir. Où résident les connaissances qui rendent ces inférences possibles ?

L'hypothèse moléculariste consiste à supposer que l'information qui nous guide pour faire les inférences obligatoires est constitutive du concept. Le concept CHIMPANZÉ doit contenir une information permettant de retrouver sa position dans l'arbre taxonomique correspondant

aux connaissances dont le sujet dispose. Il doit exister une information intégrée au concept FLEUR qui détermine son caractère fragile, et qui lui fournit des traits de couleur et de parfum dont les valeurs restent à instancier. Celui qui ne comprend pas que le chat est mort, après avoir entendu la phrase il a tué le chat, n'attache sans doute pas le concept correct au mot tuer. Pour que le concept TUER puisse, à lui seul, produire l'inférence obligatoire concernant la mort du patient de l'action décrite, il faut qu'il contienne, au moins en puissance, l'information permettant cette inférence. Autrement dit, qu'il ait une structure interne. Les inférences fortuites, qui peuvent être liées à l'histoire personnelle du sujet, par exemple la pensée faiblement suggérée selon laquelle le chat occis était malade, peuvent s'expliquer par bien d'autres mécanismes externes, par exemple une matrice associative. En revanche, les inférences obligatoires, qui sont systématiquement valides lors de l'usage d'un mot, ne peuvent pas être soumises aux aléas d'un mécanisme externe. Leur existence suggère fortement la présence d'une structure interne pour les concepts.

La deuxième motivation pour conférer des structures internes aux concepts lexicaux est liée au caractère compositionnel des concepts (CF. CHAPITRE 5). Les concepts peuvent se combiner pour engendrer des représentations composées. Or, toutes les combinaisons ne sont pas sémantiquement admissibles. L'information qui contraint cette combinatoire serait constitutive des concepts. Prenons un premier exemple. Dans la plupart des contextes, des syntagmes comme pendant qu'il sautillait ou pendant qu'il toussait sont interprétées comme la répétition de la situation exprimée par les verbes sautiller ou tousser. Ceci suggère que ces verbes partagent la propriété de désigner des événements intrinsèquement itératifs, alors qu'un verbe comme manger, dont la conjugaison à l'imparfait n'engendre pas systématiquement de répétition, sera plutôt caractérisé comme un verbe d'activité. Il est donc tentant d'attacher l'étiquette ÉVÉNEMENT - ITÉRATIF à un concept comme SAUTILLER, de manière à en faire un élément de sa structure. Cette solution présente l'intérêt d'expliquer la systématisme de certaines possibilités de combinaisons : la phrase il a sautillé pendant une heure est sémantiquement bien formée, alors que la phrase il a sautillé en une heure semble incorrecte, tandis que pour le verbe manger, les deux tournures sont possibles.

L'intérêt de l'hypothèse moléculaire pour la composition des concepts se fait particulièrement sentir dans l'interface avec la syntaxe. La présence d'une structure conceptuelle peut par exemple expliquer le comportement différencié de certains verbes. Dans la phrase le sous-marin a coulé le bateau, le mot bateau reçoit le même rôle sémantique que dans la phrase le bateau a coulé, où le même verbe est employé sous sa forme intransitive. Or, le couple de phrases le loup a mangé la brebis et la brebis a mangé fonctionne différemment : le mot brebis n'a pas, en tant que complément direct dans la première phrase, le même rôle sémantique qu'il a dans la deuxième phrase, où il remplit la place du sujet. Cette particularité du verbe couler pourrait être expliquée par le fait que le concept COULER possède une information, au sein de sa structure, indiquant que le complément, lorsqu'il est présent, hérite du rôle du sujet de la version intransitive. En revanche, le concept MANGER ne posséderait pas cette information structurelle. Un autre exemple, où la structure conceptuelle peut expliquer la combinaison des significations, nous est fourni par la notion de causalité. Si le sens de la phrase le chat est mort est systématiquement lié à celui de la phrase il a tué le chat, c'est que le sens de la première apparaît comme le résultat de la situation décrite par la seconde. Ce lien entre significations composées peut s'expliquer, dans la structure du concept TUER, par la présence d'une information exprimant qu'une action particulière provoquant un changement d'état s'exerce sur le complément. La nature conceptuelle du lien causal, que l'on traduit par la présence du concept CAUSE dans la structure du concept TUER, est suggérée par le fait que le lien semble exister indépendamment des situations qu'il relie. Le même lien causal pourra ainsi expliquer la connexion systématique entre les concepts ABATTRE et TOMBER.

L'idée d'une structure semble évidente lorsque l'on s'intéresse aux représentations conceptuelles associées aux syntagmes ou aux mots possédant une structure morphologique transparente. Ainsi, en vertu du principe de constituance, la représentation associée au syntagme chat noir doit contenir le concept associé au mot chat ainsi que celui associé au mot noir. En vertu du principe de compositionnalité, la structure qui lie ces concepts au sein de la représentation associée au syntagme chat noir est tirée de la syntaxe. L'hypothèse moléculariste, toutefois, est plus forte. Lorsque l'on suppose que le concept TUER contient celui de CAUSE, ou un lien fixe vers celui de MOURIR, on se situe en dehors de l'hypothèse de compositionnalité. Il n'y a rien, dans la forme phonologique du mot tuer qui indique la relation entre ces trois concepts. La structure est donc attachée au concept, non au mot.

L'hypothèse moléculariste est également plus forte que le simple fait de doter les concepts lexicaux d'une structure interne. On pourrait imaginer que cette structure soit exogène, c'est-à-dire que ses composants soient non conceptuels. Or, les structures postulées par l'hypothèse moléculariste pour les concepts lexicaux sont de nature endogène, c'est-à-dire que leurs composants appartiennent au niveau conceptuel. C'est pour cette raison qu'il s'agit de représentations récursives. Dans sa version stricte, l'hypothèse moléculariste suppose que les composants des concepts lexicaux sont également des concepts lexicaux. Ainsi, le composant ÉVÉNEMENT dont nous avons doté le concept SAUTILLER ne serait autre que le concept associé au mot événement. De même, l'idée de cause présente dans le concept TUER serait rendue par le concept CAUSE associé au mot cause. Dans cette version, la structure postulée par l'hypothèse moléculariste s'apparente à celle d'une paraphrase. Ainsi, le mot tuer aurait la même structure conceptuelle que le syntagme causer la mort de.... Si l'on adopte une version moins stricte du molécularisme, on autorise la présence, dans les structures conceptuelles, de concepts abstraits, c'est-à-dire de concepts qui ne peuvent être nommés. Ces structures, bien que restant endogènes, comportent des éléments conceptuels ineffables. Par exemple, on peut prétendre que les composants conceptuels CAUSE ou ÉVÉNEMENT dans les exemples précédents n'ont qu'un vague rapport avec les concepts associés aux mots cause et événement. Dans ce cas, la question se pose de proposer une description scientifiquement acceptable de ces concepts ineffables.

L'analogie avec la chimie, qui nous a suggéré la qualification de "moléculaire" pour les structures conceptuelles, est intéressante. En chimie classique, les composés sont décrits par une formule censée représenter leur structure. Ces formules sont récursives. Ainsi, l'acide sulfurique peut s'écrire H_2SO_4 , formule dans laquelle apparaît une autre formule, celle de l'ion SO_4^{2-} . Les structures les plus simples sont les atomes ou les ions atomiques, dont la structure interne sort du cadre de la chimie. Ainsi, on peut considérer que les écritures de la chimie classique reposent sur des structures endogènes dont les composants sont "chimiquement interprétables", c'est-à-dire qu'ils correspondent à des substances¹. Le problème fondamental, en ce qui concerne la sémantique du langage, est de savoir si la composition du sens des mots dans une phrase peut être décrite d'une manière qui ressemble à la composition des substances dans une réaction chimique. Selon l'hypothèse moléculariste, les mots, comme les phrases, renvoient à des structures conceptuelles. La question qui se pose est de savoir comment représenter concrètement cette chimie du sens.

¹ Le statut de l'électron, qui intervient par exemple dans les équations d'électrolyse, pourrait être présenté comme celui d'une entité ineffable au niveau chimique. De même, une description des ions à partir des orbitales atomiques pourrait conférer une structure exogène aux éléments de base de la chimie.

7.2. Méthodes pour structurer les concepts lexicaux

Une méthode assez naturelle pour doter les concepts lexicaux d'une structure constituée d'éléments eux-mêmes conceptuels consiste à s'inspirer des définitions mathématiques. En mathématique, chaque nouvelle notion est définie en fonction des notions déjà définies ou de notions primitives, comme la notion d'ensemble en théorie des ensembles. Le souci fondamental est d'éviter la circularité des définitions. Si l'on transpose cette méthode définitionnelle dans le cas des concepts, on cherchera à offrir une définition pour chacun d'entre eux, hormis pour un petit nombre de concepts primitifs. Ainsi, dans l'exemple du verbe sautiller, c'est en tant que concept primitif que le concept ÉVÉNEMENT intervient dans la structure du concept SAUTILLER. Certains concepts composés seront naturellement antérieurs à d'autres dans cette démarche définitionnelle, simplement parce qu'ils sont plus généraux. Ainsi, on peut considérer que le concept MAMMIFÈRE est plus général que le concept CHIMPANZÉ et peut entrer dans sa définition. Dans d'autres cas, il peut s'agir d'un choix de commodité, par exemple si l'on définit le concept MONTRER à partir du concept DÉSIGNER plutôt que l'inverse.

Il est remarquable que les dictionnaires, dont on attend des définitions précises et non ambiguës, ne s'inscrivent pas dans la démarche définitionnelle. Aucune garantie n'est offerte contre la circularité². Au contraire, chaque entrée utilise de nombreux autres concepts, sans se limiter à des concepts plus généraux ou définis plus simplement. Il est donc cohérent d'assumer pleinement cette interrelation complexe entre concepts lexicaux et d'abandonner l'idée que l'on puisse construire un catalogue non circulaire de définitions. Selon cette méthode relationnelle, le sens d'un concept est l'ensemble des relations qu'il entretient avec les autres concepts. Il semble plausible, par exemple, de supposer que les concepts sont caractérisés par leur rôle dans une théorie locale impliquant plusieurs autres concepts. Une théorie recouvre l'ensemble des relations conceptuelles que l'on utilise pour gérer les connaissances dans un contexte particulier. Par exemple, la notion de RACCORD prend son sens dans le réseau de connaissances que le plombier mobilise lorsqu'il réalise une installation de chauffage central. Les différents concepts de cette théorie locale : TUBE, RACCORD - DROIT, RACCORD - T, BRASURE, FILASSE, SECTION, FILETAGE, *et cætera*, n'ont de sens, selon cette démarche relationnelle, que les uns par rapport aux autres.

Pour étayer la plausibilité théorique de l'hypothèse moléculariste, nous allons montrer comment elle a pu être déclinée en Intelligence Artificielle en suscitant le développement de plusieurs formalismes de représentation des connaissances et de traitement automatique du langage naturel³.

Formalismes pour la méthode définitionnelle

Une première idée, assez intuitive, pour représenter un concept par une définition, consiste à traduire le concept dans un formalisme logique. Par exemple, on cherchera à donner au concept TUER une structure logique qui reflète l'idée du syntagme causer la mort de. Sans précaution, on risque de se trouver contraint d'adopter un formalisme logique d'ordre supérieur à un, permettant au concept CAUSE d'agir sur d'autres représentations conceptuelles.

² Par exemple une référence croisée entre les mots crime et délit n'est pas exclue.

³ Les exemples de formalismes que nous avons retenus pour cet exposé, tant pour la méthode définitionnelle que pour la méthode relationnelle, sont des versions simplifiées et généralisées de formalismes existants. Nous juxtaposons ces formalismes, en les instanciant sur un même exemple, afin d'illustrer notre propos concernant les structures conceptuelles récursives. Bien évidemment, ces formalismes ont été conçus pour aborder des problèmes différents. Chacun présente ses avantages et désavantages propres, que nous ne chercherons pas à opposer ici.

On risque également de devoir utiliser une modalité, car le lien entre le concept TUER et le concept MOURIR est de type temporel. L'inconvénient présenté par une logique modale d'ordre deux réside d'une part dans la complexité des procédures déductives, et d'autre part dans le manque de lisibilité des structures conceptuelles obtenues qui comportent une imbrication entre opérateurs modaux et quantifications sur les prédicats. Il est cependant possible d'éviter toute imbrication formelle des concepts en les décomposant jusqu'à obtenir une structure "plate", que l'on peut exprimer à l'aide de la logique de premier ordre.

$$\text{tuer}(x, y) \equiv_{\text{déf}} \text{cause}(x, e) \ \& \ \text{mourir}(e, y) .$$

Cette mise à plat utilise des variables de liaison qui permettent d'éviter la complication formelle. Ainsi, la variable e apparaît dans le prédicat cause , ce qui la contraint à désigner un événement causé par une action de l'individu x . Elle réapparaît également dans le prédicat mourir pour désigner cette fois l'événement qui consiste en la mort de l'individu y . Ce partage de variable permet de réaliser la liaison sémantique entre prédicats qui, autrement, nécessitait une imbrication.

La solution de la mise à plat présente plusieurs avantages. La décomposition peut être poussée aussi loin que nécessaire, tout en évitant les imbrications. L'écriture suivante fait apparaître des notions cinématique (passage), dynamique (cause), et des concepts généraux comme la vie ou la mort, que l'on pourrait accepter comme primitifs.

$$\text{tuer}(x, y) \equiv_{\text{déf}} \text{cause}(x, e) \ \& \ \text{passage}(e, y, d, f) \ \& \ \text{vie}(d) \ \& \ \text{mort}(f) .$$

Un autre avantage de la mise à plat est lié au fait que le formalisme obtenu est d'ordre un. Il peut donc être directement manipulé par un moteur de déduction pour produire les inférences que l'on peut attendre d'un système sémantique. Le langage informatique Prolog, par exemple, peut opérer sur des expressions déclaratives comme celles de notre exemple. Par son mécanisme dit de résolution, ce langage peut remplacer toute occurrence du prédicat $\text{tuer}(x, y)$ par la définition qui lui est fournie. Ainsi, un programme conçu en Prolog peut décider de la satisfaction d'un terme qui lui est soumis, en remplaçant le terme par sa définition, puis en répétant récursivement l'opération pour les nouveaux termes ainsi introduits, jusqu'à parvenir éventuellement à des faits connus comme vrais⁴.

Enfin, la technique de mise à plat rend visibles les parentés entre concepts, puisque l'on retrouve les mêmes prédicats de base dans différentes définitions. Ainsi, le concept MOURIR peut être défini à l'aide des concepts primitifs introduits précédemment.

$$\text{mourir}(x) \equiv_{\text{déf}} \text{passage}(x, d, f) \ \& \ \text{vie}(d) \ \& \ \text{mort}(f) .$$

Cette parenté entre les définitions nous rapproche de la propriété de constituance lexicale, puisque la définition du concept TUER contient, aux variables de liaison près, celle du concept MOURIR.

Toutefois, cette recherche de définitions sans imbrications présente les défauts de ses qualités. D'une part, ces définitions font appel à des variables de liaison dont la pertinence cognitive peut apparaître comme discutable. D'autre part, étant plates, elles font disparaître la hiérarchie inhérente à la structure syntaxique, ce qui rend l'interface avec la syntaxe délicate à concevoir. On peut donc légitimement préférer l'option opposée, qui consiste à assumer pleinement l'imbrication des concepts en exprimant les définitions à l'aide d'un formalisme fonctionnel.

⁴ Prolog utilise généralement l'hypothèse du monde clos, ce qui signifie que les termes qu'il ne peut pas démontrer sont considérés comme faux. On peut cependant se passer d'une telle hypothèse en ce qui concerne la production d'inférences. Par ailleurs, nous ne mentionnons pas ici les risques de bouclage liés aux définitions auto-référentes.

$tuer(x, y) =_{\text{déf}} cause(x, passage(y, vie, mort))$.

Cette définition fait l'économie des variables de liaison. Les variables x et y sont des variables conceptuelles qui prennent comme valeur les représentations conceptuelles associées aux arguments syntaxiques du verbe *tuer*. Elles sont donc de même nature que les concepts *VIE* et *MORT*.

Les définitions fonctionnelles présentent l'avantage de parfaitement respecter le principe de constituance lexicale. Ainsi, la définition du concept *TUER* contient le concept *MOURIR* qui peut à son tour être défini.

$tuer(x, y) =_{\text{déf}} cause(x, mourir(y))$.

$mourir(x) =_{\text{déf}} passage(x, vie, mort)$.

Sous cette forme, la définition pourrait sembler peu exploitable, que ce soit pour son utilisation dans le raisonnement ou pour l'interface avec la syntaxe. Il faut en particulier indiquer le moyen de récupérer la valeur des variables x et y à partir de la structure de l'énoncé linguistique où le verbe *tuer* apparaît. Il existe cependant des formalismes à base de représentations fonctionnelles qui répondent parfaitement à cette exigence. La composition prend alors la forme d'une dérivation. Dans l'illustration que nous donnons ci-dessous, la dérivation sémantique qui permet l'intégration des arguments dans le prédicat est assurée par un formalisme de calcul de Lambek. Cette dérivation est parallèle au décodage syntaxique qui, lui, est assuré par une grammaire catégorielle. Elle opère sur des entrées lexicales fortement structurées, comme celles que nous donnons pour les verbes *causer*, *mourir* et *tuer*.

$causer =_{\text{déf}} \lambda P. \lambda x. \underline{causer}(P)(x) :: np \backslash s / s$.

$mourir =_{\text{déf}} \lambda x. \underline{mourir}(x) :: np \backslash s$.

$tuer =_{\text{déf}} \lambda y. \lambda x. \underline{causer}(\underline{mourir}(y))(x) :: np \backslash s / np$.

À chaque entrée lexicale, on associe une définition sémantique, ainsi qu'une catégorie syntaxique. Par exemple, le verbe *causer*, reçoit une définition sémantique, le terme *causer*, construite autour d'une fonction supposée prédéfini, *causer*. Celle-ci admet deux arguments, P et x , qui correspondent respectivement à l'événement causé et l'individu acteur. Le terme *causer* reçoit également une catégorie syntaxique $np \backslash s / s$. Cette expression exprime qu'en tant que verbe, le mot *causer* prend un groupe propositionnel à sa droite ($/s$) et un groupe nominal à sa gauche ($np \backslash$) pour devenir à son tour un groupe propositionnel (s). Le verbe *mourir* reçoit une définition similaire, si ce n'est que la fonction *mourir* ne prend qu'un seul argument x et que le terme *mourir* n'accepte, syntaxiquement, qu'un groupe nominal à sa gauche ($np \backslash$). La définition sémantique de l'entrée lexicale *tuer* est construite autour des fonctions *causer* et *mourir*. Le développement qui suit montre comment le sens de la phrase *A tue B* se construit au fur et à mesure que les syntagmes sont agrégés.

$$A :: np \quad \quad \quad tuer :: np \backslash s / np \quad \quad \quad B :: np \quad \quad \quad (1)$$

$$A :: np \quad \quad \quad \underline{(tuer :: np \backslash s / np)(B :: np)} \quad \quad \quad (2)$$

$$A :: np \quad \quad \quad \underline{(\lambda y. \lambda x. \underline{causer}(\underline{mourir}(y))(x) :: np \backslash s / np)(B :: np)} \quad \quad \quad (3)$$

$$A :: np \quad \quad \quad \lambda x. \underline{causer}(\underline{mourir}(B))(x) :: np \backslash s \quad \quad \quad (4)$$

$$\underline{(\lambda x. \underline{causer}(\underline{mourir}(B))(x) :: np \backslash s)(A :: np)} \quad \quad \quad (5)$$

$$\underline{\underline{causer}(\underline{mourir}(B))(A) :: s} \quad \quad \quad (6)$$

Les syntagmes nominaux A et B sont définis par les termes $A::np$ et $B::np$. Le terme *tuer* s'applique successivement au terme B puis au terme A. L'expression (S) (T) pour deux termes S et T quelconques signifie que le terme T est offert comme argument au terme S. L'expression $(\lambda x.U(x)) (T)$ se réduit à celle de $U(T)$. Cette réduction signifie que le terme U, considérée comme fonction de la variable x en raison de la présence de l'expression λx , est appliqué au terme T. L'expression λx peut se comprendre comme une indication sur la manière de considérer le terme T en tant qu'argument du terme U. Ainsi, par la présence de l'expression λy dans la définition du terme *tuer* en (3), le terme B peut venir en (4) remplacer la variable y de la fonction *mourir*. L'étape suivante, en (5) et (6), voit l'association du terme A à la variable x de la fonction *causer*. Cet exemple illustre la manière dont les arguments sont intégrés à la définition sémantique au fur et à mesure qu'ils sont acceptés au niveau syntaxique.

Dans ce genre de représentation fonctionnelle, le comportement syntaxique et la structure argumentale sémantique des mots sont codés dans l'entrée lexicale correspondant au mot. L'interface avec la syntaxe est ainsi assurée, en quelque sorte, par le concept lui-même, dans la mesure où l'expression fonctionnelle qui le définit peut être considérée comme une procédure interprétable par une machine générale. Cette richesse permet au concept de contenir des instructions pour la production d'inférences. Lors de l'instanciation des variables par le décodage syntaxique, ces instructions déclenchent la dérivation de nouvelles structures, jusqu'à l'obtention d'un résultat, par exemple le fait que le complément du verbe *tuer* se retrouve à la place de l'argument de la fonction associée au verbe *mourir*.

Cette solution, qui consiste à enfermer dans chaque concept les instructions qui lui permettent de récupérer ses arguments en tirant partie de l'agencement syntaxique, n'est praticable que si cette information procédurale dérive de schémas généraux. Le comportement sémantique semblable de concepts comme *TUER* et *LICENCIER* ne saurait être dû à des similitudes structurelles fortuites. On en vient naturellement à décrire ces deux concepts verbaux comme héritant d'un schéma général où apparaissent les concepts *CAUSE* et *PASSAGE*. Cette logique d'héritage conduit à concevoir une hiérarchie conceptuelle. On dispose alors de la liberté de produire des concepts par spécification de patrons. Le risque est cependant d'engendrer ainsi des concepts inutilisables, par exemple exprimant le fait qu'un objet cause un autre objet. C'est pourquoi l'héritage s'accompagne généralement d'un typage.

$$\text{tuer}(x, y) \rightarrow_{\text{déf}} [\text{événement cause}([\text{entité } x], \\ \text{événement passage}([\text{entité } y], [\text{lieu vie}], [\text{lieu mort}])]] .$$

Une telle représentation typée reste pleinement adaptée au principe de constituance lexicale.

$$\text{tuer}(x, y) \rightarrow_{\text{déf}} [\text{événement cause}([\text{entité } x], [\text{événement mourir}([\text{entité } y])]] .$$

$$\text{mourir}(x) \rightarrow_{\text{déf}} [\text{événement passage}([\text{entité } x], [\text{lieu vie}], [\text{lieu mort}])]$$

Les concepts *MOURIR* et *TUER* dérivent des patrons plus généraux comme les suivant.

$$[\text{événement cause}([\text{entité } x], [\text{événement } i])]$$

$$[\text{événement passage}([\text{entité } x], [\text{lieu } u], [\text{lieu } v])]$$

Ces même patrons peuvent être instanciés de diverses manières pour donner d'autres concepts, par exemple le concept *LICENCIER* ou le concept *DÉPLACER*.

$$\text{licencier}(x, y) \rightarrow_{\text{déf}} [\text{événement cause}([\text{entité } x], \\ \text{événement partir}([\text{entité } y])]] .$$

$\text{partir}(x) \rightarrow_{\text{déf}} [\text{événement passage}([\text{entité } x], [\text{lieu emploi}], [\text{lieu chômage}])]$.

$\text{déplacer}(x, y) \rightarrow_{\text{déf}} [\text{événement cause}([\text{entité } x], [\text{événement bouger}([\text{entité } y])])]$.

$\text{bouger}(x) \rightarrow_{\text{déf}} [\text{événement passage}([\text{entité } x], [\text{lieu } _], [\text{lieu } _])]$.

Notons que les structures des concepts DÉPLACER et BOUGER correspondent aux patrons généraux mentionnés précédemment. Cette structuration en patrons typés suggère que les concepts EMPLOI et CHÔMAGE dans le concept LICENCIER, de même que les concepts VIE et de MORT dans le concept TUER, se comportent comme des lieux dans un certain espace. Le choix d'un typage a donc des conséquences sur la structuration de nombreux concepts, et revêt une importance cognitive considérable.

L'effet du typage est de poser des contraintes sur le type des arguments qui peuvent être offerts à la structure. Si l'on opte pour cette logique d'héritage, on peut abandonner l'idée que les concepts contiennent les instructions permettant les interfaces avec la syntaxe et le raisonnement. L'interface syntaxique peut être assurée par des règles opérant sur les patrons conceptuels généraux, par exemple une règle qui transforme le sujet du syntagme verbal, qui est externe dans la structure syntaxique, en argument interne du concept. Une telle règle peut s'écrire, dans le cas d'un verbe transitif, comme suit.



L'expression XP est un syntagme dont la tête est la catégorie syntaxique X0. L'expression S désigne la position de spécifieur du syntagme XP et l'expression CP celle de son complément. Elles correspondent respectivement aux variables conceptuelles E₁, E₂. Ces concepts sont arguments de la variable conceptuelle F de type catégorie. Celle-ci est instanciée avec le concept associé au mot qui occupe la position X0. De même, la production d'inférences peut être réalisée en attachant des règles aux prédicats les plus généraux, exprimant par exemple que d'une structure de type cause on peut logiquement déduire la sous-structure de type événement correspondant à son effet.

Le contrôle de la construction du sens, dans cette optique, n'est pas effectué par des instructions procédurales contenues dans les concepts, mais par des déclarations de contraintes sur les types. Si ces dernières sont suffisamment détaillées, par exemple sous la forme de structures de traits typés, le contrôle de la construction du sens peut être assuré par un mécanisme unique et général, l'unification par appariement sous contraintes. L'illustration suivante présente une hiérarchie de types sémantiques dans une représentation inspirée du formalisme ALE (*Attribute Logic Engine*) qui est un moteur d'unification de structures de traits typés.

prédicat	sous-types	[unaire, binaire]
	traits	[séquence : séquence].
unaire	sous-types	[]
	traits	[argument : entité].
binaire	sous-types	[]
	traits	[argument1 : entité, argument2 : entité].
séquence	sous-types	[]
	traits	[étape1 : situation, étape2 : situation, étape3 : situation].
situation	sous-types	[état, événement, processus, nul]
	traits	[thème : entité].
état	sous-types	[].
événement	sous-types	[].
processus	sous-types	[]
	traits	[cause : entité].
entité	sous-types	[abstrait, concret].
concret	sous-types	[objet, lieu].
objet	sous-types	[animé, inanimé].

Ce fragment de hiérarchie conceptuelle indique comment les types, dans ce formalisme, peuvent se particulariser en sous-types. Par exemple le type entité, par le biais de ses sous-types concret et objet, se particularise en animé et inanimé. La structure interne d'un type est donnée par ses traits. Chaque trait a un nom et un type⁵. Ainsi, le type binaire, censé représenter les prédicats qui admettent deux arguments, est caractérisé par deux traits, *argument1* et *argument2*, qui sont tous deux de type entité. La structure d'un type est héritée par ses sous-types. Ainsi, binaire possède non seulement sa propre structure de traits indiquant la présence de deux arguments, mais également le trait *séquence*, qui représente sa structure temporelle, héritée du type prédicat. Le fait de typer les traits permet d'imposer des contraintes sur les types. Par exemple, le type séquence comporte trois traits *étape1*, *étape2* et *étape3* dont les valeurs sont de type situation. Ce choix contraint la structure temporelle des prédicats en maximum trois situations successives. Pour un verbe comme dormir, les traits *étape2* et *étape3* prendront la valeur nul pour traduire le fait que le concept DORMIR ne comporte qu'une seule étape. Pour réveiller, on pourra utiliser deux des trois étapes, par exemple avec les valeurs respectives des types processus et événement, pour traduire que l'événement du réveil est l'aboutissement d'un processus causal.

Les contraintes posées sur les structures de traits sont cumulatives. Ainsi, les traits du type séquence ne sont soumis qu'à la contrainte d'être du type situation, mais lorsqu'on définit la structure dormir en tant que prédicat, on peut contraindre davantage son trait *étape1* en lui imposant d'être du type état. Illustrons cette possibilité sur le concept TUER, qui peut être défini par la structure suivante.

⁵ Il arrive que l'on donne à un trait le nom de son type, mais il ne s'agit que d'une homonymie de commodité.

structure *mourir* est du même type que la valeur du trait *étape2* dans la structure *tuer*. Il suffit d'étiqueter cette sous-structure pour souligner la parenté entre les deux concepts de départ.

$mort \Rightarrow_{\text{déf}}$
événement, *thème*: animé.

La structure *mort* est ainsi définie comme une spécification du type événement. Le lien entre les concepts *TUER* et *MOURIR* s'effectue alors simplement en remplaçant le type événement par la structure *mort* dans les deux expressions, respectivement dans les valeurs des traits *étape2* et *étape1*. Ce nouveau lien entre les structures temporelles des concepts *TUER* et *MOURIR* permet alors de produire des inférences. Une règle d'inférence générale peut lier un processus causal comme la structure *tuer* à un événement comme la structure *mourir* en unifiant le trait *étape2* de la première structure avec le trait *étape1* de la deuxième. Grâce à l'unification provoquée par cette règle d'inférence, l'entité instanciée dans le deuxième argument du verbe *tuer* se retrouve argument du verbe *mourir*. Le système, grâce à ses structures riches, produit ainsi une inférence correcte portant sur les concepts et leurs arguments.

L'un des avantages des structures de traits est qu'il est facile de les développer ou d'augmenter leur richesse. Nous pouvons ainsi ajouter le trait *manière* au type prédicat. Dans la structure *tuer*, ce trait pourra être instancié par la valeur portée par un adverbe éventuel, par exemple *froidement*, *sauvagement*, *et cætera*. La présence de ce trait, même s'il n'est pas instancié par unification avec un adverbe, impose tout de même une contrainte, celle du type associé au trait. Cette contrainte peut ainsi permettre de proscrire que certains adverbes, comme *gentiment*, puissent être attachés au verbe *tuer*.

Grâce à ce système où les concepts prennent la forme de structures de traits typés, on parvient à exprimer tous les concepts qui correspondent à des assemblages de ces types prédéfinis. Ces types forment une hiérarchie d'héritage dans laquelle sont prosrites circularité et redondance. L'hypothèse de non circularité, qui est à la base de la méthode définitionnelle, revêt ici un caractère technique. Il serait difficile d'identifier les types d'après les traits spécifiés si le système de types présentait des cycles. L'évitement de la redondance, quant à lui, est motivé à la fois sur le plan technique et sur le plan cognitif. Deux structures identiques doivent nécessairement dériver d'un même type. Cependant, une telle exigence conduit à définir de nombreux types intermédiaires sans réalité cognitive.

Pour d'autres formalismes, l'essentiel est de caractériser formellement les concepts, sans poser d'a priori sur d'éventuelles circularités ou redondances. Ces formalismes s'inscrivent dans la méthode que nous avons appelée relationnelle.

Formalismes pour la méthode relationnelle

Les premiers formalismes utilisés en intelligence artificielle pour stocker la connaissance d'un domaine de manière explicite consistaient à mémoriser une liste de règles logiques portant sur des propositions ou des prédicats. Cette technique est adaptée à la production du raisonnement, mais se révèle inadéquate lorsqu'il s'agit de représenter le sens des énoncés du langage. Par exemple, on peut stocker, indépendamment du concept *TUER*, une règle du type suivant.

$tuer(x, y) \supset mourir(y)$.

Toutefois on peut regretter, d'une part, que cette règle ne soit pas une partie intégrante du concept *TUER*, et d'autre part, qu'elle rende imparfaitement l'idée de cause par une implication logique. Pour améliorer la représentation sémantique, il est nécessaire de disposer

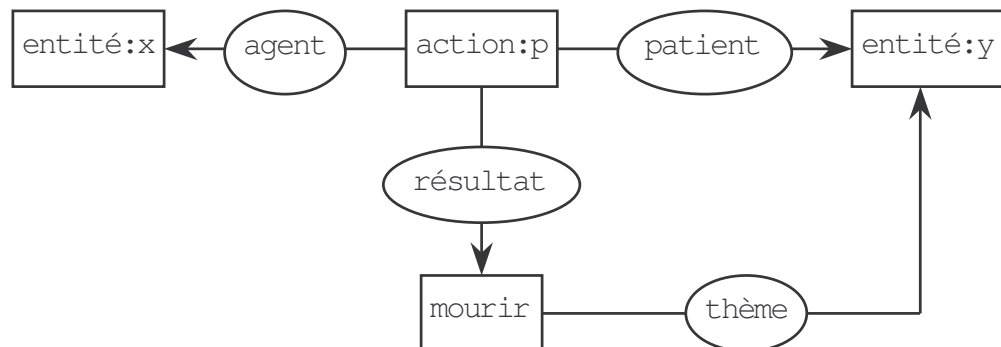
de tout un répertoire de relations entre concepts que la logique n'offre pas⁶. L'idée à la base des réseaux sémantiques consiste à représenter explicitement la mise en relation des concepts. Dans sa forme la plus immédiate, cette approche place les concepts aux nœuds d'un graphe dont les arêtes sont les relations. Un réseau sémantique est un graphe orienté et étiqueté dans lequel les concepts, représentés par les nœuds (les rectangles dans le schéma) sont reliés par des relations qui sont représentées par les arêtes (étiquetées par les ovales dans le schéma).



Un réseau sémantique en tant que tel ne pose aucune autre contrainte sur les concepts que les relations qu'il exprime. L'intérêt de la représentation sous forme de réseau est l'inclusion de chaque concept avec ses liens dans un grand ensemble d'associations conceptuelles. Par exemple, le concept d'ARME est lié au concept de TUER par une relation de type MOYEN, le concept de MOURIR est associé à celui de DEUIL, *et cætera*. Cependant, il s'agit bien d'une représentation moléculaire, dans la mesure où l'on appelle concept un nœud du réseau muni de ses arêtes. Ainsi, tout concept contient une référence explicite à certains autres concepts.

De manière équivalente, on peut appeler concept une portion du réseau incluant tous les nœuds liés à un nœud donné, ainsi que les arêtes internes à cet ensemble de nœuds. Les graphes conceptuels, dans une version élaborée des réseaux conceptuels, peuvent inclure un système de typage et des variables permettant de lier les concepts⁷.

tuer:p



Le typage offre un système d'héritage entre concepts. Ainsi, le type action qui porte sur la variable p permet d'exprimer le fait que le concept tuer est une instance du concept action. Le typage reste dans l'esprit des réseaux conceptuels. Il remplace une arête étiquetée instance qui pointerait vers le nœud parent, ce qui permet d'alléger la représentation schématique.

Les aspects statiques d'un graphe conceptuel peuvent être rendus dans une écriture logique dans un système de premier ordre typé.

```

    [tuer:p]: [action:p] & [entité:x] & [entité:y] & [mourir:q];
    agent(x,p) & patient(y,p) & résultat(q,p) & thème(y,q).
  
```

⁶ Les réseaux sémantiques qui sont décrits ici peuvent être ramenés à un ensemble de règles logiques. Cependant, ils offrent une description des relations conceptuelles plus proche de l'intuition. De plus, ils peuvent être utilisés comme un graphe pour propager des associations.

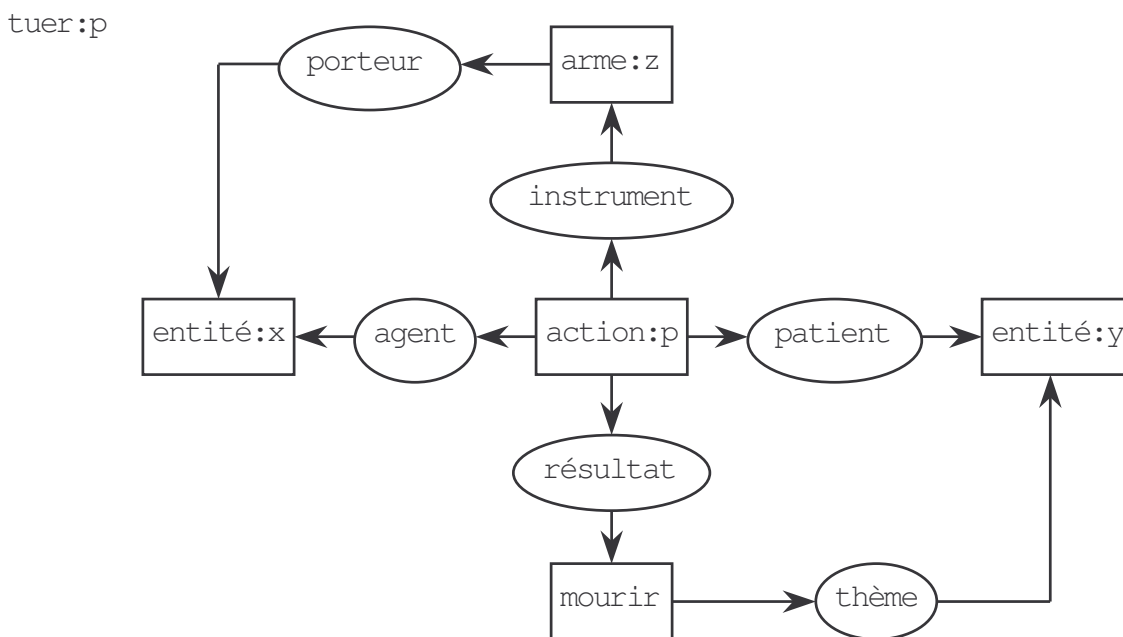
⁷ Le mot graphe conceptuel est parfois utilisé dans un sens différent, pour parler de la représentation d'une phrase.

Sous cette forme, caractériser le concept TUER apparaît comme définir un nouveau type qu'une variable peut prendre. La liaison avec d'autres concepts, par exemple le concept MOURIR, est exprimée par l'intermédiaire des variables p et q qui interviennent séparément dans la partie typage de la clause, sous les types `tuer` et `mourir`, et conjointement dans la partie condition de la clause, dans la relation `résultat`. Le concept MOURIR peut à son tour recevoir une mise en relation du même type.

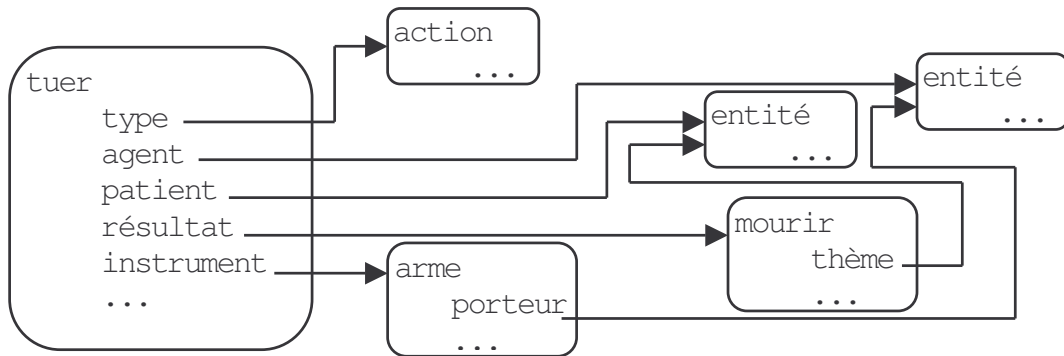
[mourir:p]: [événement:p] & [entité:x] & [vie:e₁] & [mort:e₂];
thème(x,p) & antérieur(e₁,p) & ultérieur(e₂,p).

Le typage permet ainsi de respecter le principe de constituance lexicale, puisque ici la définition du type `mourir` fait partie intégrante de la définition du type `tuer`. Notons que ce n'était pas le cas avec les définitions déclaratives plates qui n'offraient que des liaisons par partage de variables.

Le schéma précédent offre ce que l'on peut appeler le graphe canonique du concept TUER. Si l'on réduit le réseau conceptuel à un ensemble de définitions canoniques, il ne manque que la contrainte de non circularité pour retomber dans un système définitionnel. En réalité, un graphe conceptuel peut offrir bien plus. Par exemple, il peut contenir des relations dont la validité est celle de la typicité.



Ce schéma du verbe `tuer` présente des aspects non nécessaires. Il décrit les situations typiques qui tombent sous le concept TUER. Pour cette raison, la transcription statique de ces schémas fait souvent appel, non à des relations logiques d'arité fixe, mais à des structures pouvant admettre un nombre variable d'éléments comme celle qu'offrent les langages de trames. La trame liée au concept TUER peut être décrite comme suit.



La différence fondamentale entre ce type de présentation et l'écriture logique est qu'ici, ce sont les relations qui sont attachées aux concepts plutôt que l'inverse⁸. Ainsi le concept TUER est représenté par une trame au sein de laquelle des liens pointent vers d'autres trames. Ces liens expriment des relations avec d'autres concepts. La relation la plus simple est la relation d'instanciation, indiquée par le lien *type*. La relation entre les concepts TUER et MOURIR est décrite par le lien *résultat* dans la trame *tuer* qui pointe vers la trame *mourir*. De plus le lien *patient* de la trame *tuer* pointe vers la trame *entité*, qui est par ailleurs pointé par le lien *thème* de la trame *mourir*. Cette dernière peut être décrite à son tour comme suit.

```

mourir:
  type → événement;
  thème → entité (1);
  antérieur → [vie:
                type → état;
                thème → [animé:
                          type → entité (1)]];
  ultérieur → [mort:
                type → état;
                thème → [inanimé:
                          type → entité (1)]];
  circonstance → vieillesse;
  etc.
    
```

La notion de lien, contrairement à celle de trait, n'évoque pas l'idée de nécessité. Les liens typiques sont toujours présents. D'autres, moins typiques, peuvent être déclenchés par le contexte. Dans la plupart des contextes, le concept TUER est en relation avec le concept MOURIR, si bien que l'évocation du premier déclenche le deuxième par le lien *résultat*. Le lien *instrument*, moins typique, est importé par la présence d'une préposition comme *par*. De la même façon, la présence d'un adverbe, dans le contexte, amènera un lien *manière*. Ce mécanisme qui permet l'adjonction de nouveaux liens par les mots du contexte est à contraster avec le mécanisme d'unification qui opérait sur les structures de traits. Les traits agissent comme des contraintes, ils doivent tous être présents dès le départ. Dans un langage de trames, les liens ne posent pas de contraintes, mais offrent des possibilités. Non seulement ces liens sont internes aux concepts, mais c'est aussi le cas des procédures qui règlent la

⁸ En Prolog par exemple, un réseau sémantique sera stocké sous formes de règles du type `sorte_de(albatros,oiseau)`, alors qu'un système de trames comportera des déclarations du type `albatros(sorte_de,oiseau)`. Dans le premier cas, c'est la relation qui est mise en avant, alors que dans le second cas c'est le concept.

compositionnalité. Ainsi, la présence d'un adjectif déclenche une procédure de cet adjectif qui adjoint un lien manière à la trame tuer. Cette possibilité des concepts de se modifier mutuellement dans la composition est due à la possibilité offerte, comme dans les langages informatiques à objets, de doter les objets de procédures capables d'agir sur eux-mêmes ou de déclencher des procédures dans un autre objet.

La méthode relationnelle s'oppose à la méthode définitionnelle surtout par sa motivation qui privilégie l'expressivité au détriment du souci d'évitement de la circularité et de la redondance. Il est intéressant, toutefois, de détailler leurs ressemblances et les aspects qui les opposent.

Convergences et divergences au sein de l'approche moléculaire

La méthode définitionnelle et la méthode relationnelle se rejoignent sur bien des points. Ce sont deux manières de conférer une structure symbolique aux concepts. Toutes deux présentent la flexibilité qui permet de remplacer les valeurs binaires par des pondérations représentant, par exemple, des coefficients de probabilité ou de plausibilité. De telles techniques numériques sont motivées par le souci d'introduire, au sein de formalismes conçus au départ pour gérer des valeurs binaires, le flou que l'on constate dans l'ambiguïté des langues naturelles.

Les deux approches utilisées pour représenter la connaissance conceptuelle sont souvent interchangeables. Nous avons vu comment la représentation des relations nécessaires d'un graphe conceptuel peut conduire à un catalogue de définitions. Plus généralement, on peut toujours réduire différentes portions d'un graphe de relations pour en faire des arbres qui seront lus comme des structures de traits. Ce faisant, en privilégiant un parcours du graphe pour convertir celui-ci localement en arbre, on perd bien entendu une partie de l'information, notamment la distance entre les nœuds.

Une autre manière de concevoir la conversion d'une représentation relationnelle en une représentation définitionnelle consiste à se placer dans le cadre des théories. Une théorie est censée résumer la connaissance propre à un domaine limité. Elle regroupe un certain nombre de dépendances au sein d'un ensemble limité de concepts. On produit une définition pour certains termes d'une théorie donnée, considérés comme nouveaux, en rassemblant les dépendances qu'ils entretiennent avec les autres termes de la théorie. Ce processus, connu sous le nom de ramsification, consiste à remplacer les occurrences d'un nouveau terme par une variable dans l'ensemble des dépendances qui constituent la théorie, et à ajouter une quantification existentielle sur cette variable. On introduit ainsi une précédence implicite de tous les autres termes de la théorie par rapport au nouveau terme défini. Il est possible d'itérer ce processus pour plusieurs termes de la théorie, ce qui conduit à une expression logique à plusieurs variables. Notons toutefois que si l'on effectue ce traitement pour tous les concepts, il ne reste plus que la structure logique de la théorie qui fournit une seule définition, trop générale pour présenter un intérêt.

Nous pouvons cependant opposer les formalismes des deux méthodes sur la question cruciale de la constituance lexicale, selon laquelle les structures conceptuelles liées aux concepts lexicaux contiennent, littéralement, d'autres structures conceptuelles. Une technique définitionnelle comme les descriptions fonctionnelles offre la constituance de manière naturelle. La définition fonctionnelle d'un concept contient d'autres concepts, si bien que l'on peut remplacer le concept MOURIR par sa définition dans la définition du concept TUER sans changer le comportement de ce dernier. D'un autre côté, une technique comme les langages de trames est naturellement sans constituance : un concept comporte la référence à d'autres concepts, mais n'inclut pas ces concepts. Toutefois, certains formalismes définitionnels, par exemple les descriptions logiques plates, peuvent être dépourvus de constituance.

Inversement, on peut imaginer des formalismes relationnels avec constituance, si les concepts sont identifiés non à des nœuds, mais à des sous-graphes.

La différence essentielle entre la méthode définitionnelle et la méthode relationnelle concerne le caractère potentiellement holistique de la seconde. Par construction, grâce à l'hypothèse de non circularité, un système définitionnel fournit des définitions locales pour les concepts, puisque chaque concept est défini au maximum en fonction des seuls concepts définis avant lui. Dans la méthode relationnelle, il n'existe aucune garantie que l'on puisse isoler une portion du réseau des relations pour produire l'équivalent d'une définition. La notion de théorie n'offre pas le critère de délimitation voulu. Au contraire, l'exemple des théories scientifiques montre que les théories ont tendance à s'interconnecter. S'il n'existe pas de critère de délimitation, les relations qui partent d'un concept donné percolent dans l'ensemble des connaissances conceptuelles disponibles. Ainsi, la méthode relationnelle, en s'exposant au holisme, s'interdit de produire des descriptions utilisables pour les concepts, ce qui est précisément ce que la méthode définitionnelle s'attache à fournir.

Dans ce qui suit, nous allons tour à tour considérer les avantages et les inconvénients qui peuvent être tirés d'une représentation moléculaire des concepts lexicaux, en observant que ces deux aspects sont à la mesure l'un de l'autre.

7.3. Les atouts de l'approche moléculariste

L'un des principaux problèmes qu'une théorie computationnelle de la sémantique se doit de résoudre est celui de la systématité. Certaines combinaisons de mots produisent des résultats sémantiques prévisibles. Les phrases *il a tué le chat*, *il a tué le chien*, *il a tué le cheval*, *et cætera* produisent des pensées qui ne sont pas sans rapport entre elles. Le rapport entre ces phrases se constate par exemple dans le fait que cet ensemble de phrases peut être mis en correspondance avec l'ensemble des phrases *le chien est mort*, *le chat est mort*, *le cheval est mort*, *et cætera*. Attribuer la propriété de systématité à la sémantique revient à dire que la ressemblance sémantique entre les phrases de notre exemple n'est pas due à une analogie approximative, mais est le résultat d'une composition par un mécanisme qui présente la bonne propriété d'être clos pour certaines opérations de substitution qui préservent la "classe sémantique" des objets substitués. Comme nous allons le vérifier, l'hypothèse moléculariste est un moyen élégant et radical pour rendre compte de la systématité.

L'explication de la systématité des constructions sémantiques, dans le cadre de l'hypothèse du mentalais, passe par une certaine conception de la compositionnalité. Le principe de compositionnalité stipule que le sens d'une phrase peut être entièrement calculé à partir du sens de ses mots. Nous avons exprimé cette idée à partir de la fonction $\mathbf{f}()$ qui calcule le sens d'une expression langagière composée en attachant un sens s_i à chaque élément m_i de cette expression. Ainsi le sens d'une suite de mots (m_1, m_2, \dots, m_k) est le résultat de la construction $\mathbf{f}(s_1, s_2, \dots, s_k)$. La fonction $\mathbf{f}()$ tient compte de sa structure grammaticale pour produire le sens d'un énoncé. L'observation de phénomènes sémantiques systématiques suggère la possibilité de la réécriture de cette fonction à l'aide de fonctions régulières du type $\mathbf{g}_{s_1, s_2, \dots, s_{k-1}}(s_k)$. Pour donner une forme à ces fonctions, l'hypothèse de constituance paraît particulièrement naturelle. Selon cette hypothèse, l'élément s_k occupe une place dans une structure associée à $\mathbf{g}_{s_1, s_2, \dots, s_{k-1}}$, ce qui permet de rendre compte facilement des effets limités des substitutions. Ainsi, les sens associés aux phrases *il a tué le chien* et *il a tué le chat* correspondent à une même structure, simplement l'une comporte le concept CHIEN à l'endroit où l'autre fait figurer le concept CHAT. Si l'on répète le même raisonnement pour les différentes représentations sémantiques s_i qui offrent des phénomènes systématiques, on

arrive inévitablement à concevoir le sens associé à un énoncé comme une structure moléculaire, c'est-à-dire comme un assemblage récursif dans lequel chacun des éléments s_i trouve sa place. Ainsi, dans une analyse fondée sur le principe de constituance, la ressemblance sémantique entre les deux phrases de notre exemple résulte de la présence, dans chacune des représentations conceptuelles associées, du même concept TUER. Dans cette perspective, il est possible de re-formuler le calcul du sens par l'équation suivante.

$$\mathbf{s}(m_1, m_2, \dots, m_k) = \mathbf{g}_{S_1}(s_2, s_3, \dots, s_k)$$

Si l'on reproduit le raisonnement concernant le comportement systématique de certains énoncés pour l'appliquer aux mots eux-mêmes, par exemple l'aspect causal systématique évoqué par les verbes tuer, licencier, bousculer, *et cætera*, on arrive à une conclusion similaire, selon laquelle le sens des concepts lexicaux est lui aussi moléculaire. Le comportement sémantique des mots de notre exemple serait ainsi dû à la présence, comme constituant des concepts correspondants, du même concept CAUSE. Il est donc naturel de prolonger le principe de constituance sans tenir compte de la frontière des mots, en adoptant un principe de constituance lexicale qui étend l'hypothèse de constituance concernant les représentations sémantiques associées aux énoncés. Ainsi, dans cette optique, les effets systématiques attachés à la construction il a tué ... proviennent non seulement de la présence du concept TUER dans la structure sémantique associée, mais s'expliquent par la structure interne (causale, événementielle, *et cætera*) de ce concept.

Le corollaire de l'hypothèse de constituance lexicale est que toute représentation sémantique est décomposable, à l'instar des molécules chimiques que l'on peut décomposer en d'autres molécules plus simples, jusqu'au niveau de base où l'on s'attend à trouver les atomes de sens.

Le fait de postuler une structure interne pour les concepts lexicaux est une solution efficace pour résoudre le problème des combinaisons systématiques. Chaque concept est une structure pouvant posséder des "places" vides. Ces places sont comblées, étape par étape, par la fonction $\mathbf{f}()$ lors du décodage syntaxique. En fondant le mécanisme de construction du sens sur la combinaison de structures, on explique du même coup la compositionnalité et la systématisme. Cette dernière, lorsqu'elle peut être observée, s'explique très simplement par le fait que la substitution d'un mot par un autre laisse l'essentiel de la structure invariante, puisque cette structure est "apportée" par les autres mots.

L'approche moléculariste - compositionnelle (MC) de la construction du sens que nous venons d'esquisser présente d'autres caractéristiques notables. La première est une caractéristique que l'on peut qualifier d'esthétique. En postulant l'existence de structures récursives pour les concepts lexicaux, l'approche MC ne pose pas de différence de nature entre la structure du sens des énoncés et celle des concepts eux-mêmes. Autrement dit, il n'y a pas de frontière sémantique liée au caractère lexical ou non lexical d'une expression linguistique. La structure interne du concept lexical TUER réapparaît dans la représentation conceptuelle associée à la phrase il a tué le chat. Elle est simplement instanciée, complétée par exemple par le concept CHAT. L'approche MC ne postule donc qu'un seul type de représentation structurée qui sert à la fois à stocker la signification des concepts lexicaux et à représenter le sens des énoncés.

Une deuxième caractéristique importante de l'approche MC est qu'elle fait supporter l'essentiel de la complexité sémantique aux structures, pour ne postuler que des mécanismes d'une grande simplicité. Par exemple, si l'on utilise une représentation à base de structures de traits typés, la composition s'effectue par le seul mécanisme d'unification. L'information qui permet le calcul du sens est entièrement contenue dans la structure des concepts lexicaux. Il n'y a idéalement aucune règle sémantique, aucun opérateur global affectant le sens. C'est l'ensemble des contraintes inscrites dans les structures qui dirige la combinatoire du sens, et

la dynamique se résume à la seule procédure d'unification entre structures. La même idée de simplification des mécanismes sémantiques généraux du fait de l'enrichissement des structures existe dans les modèles à base de trames. Les structures conceptuelles contiennent en elles-mêmes des procédures qui sont déclenchées au moment opportun, lors de l'adjonction des trames due au décodage syntaxique.

L'approche MC présente aussi l'intérêt d'offrir un traitement des appariements sémantiques inappropriés. Les combinaisons inacceptables sont repérées soit dynamiquement, par l'échec de la procédure générale d'appariement, soit statiquement par leur caractère mal structuré. Dans les deux cas, il est facile d'envisager un traitement graduel de l'erreur. Cette gradualité reproduit celle du jugement sémantique spontané. Elle permet aussi de traiter les métaphores. Une métaphore est analysée comme une erreur d'appariement dont les conséquences structurelles restent de portée limitée.

Ajoutons à l'attrait de l'approche moléculaireiste - compositionnelle que le calcul sémantique qu'elle offre s'interface parfaitement avec celui des relations syntaxiques, au point que dans certaines approches, rien ne permet de les distinguer⁹. En particulier, il est facile d'utiliser les indications grammaticales au sein de la phrase pour guider la composition sémantique. De plus, comme ces calculs d'appariement sont essentiellement des calculs locaux, la construction sémantique peut s'effectuer en parallèle avec la construction syntaxique, ce qui représente un avantage considérable.

La même remarque peut être faite en ce qui concerne l'interface avec le système inférentiel. Par exemple, il est possible de déduire de la structure du concept TUER le fait que l'entité désignée par le complément du verbe tuer cesse d'être en vie. Ces inférences sont produites directement à partir des structures, sans qu'il soit nécessaire de construire une base de règles indépendante.

Mentionnons enfin que le fait d'associer une structure déterminée à chaque concept lexical permet d'envisager une variété de mécanismes d'apprentissage. Les méthodes de l'apprentissage symbolique automatique, qu'elles soient inductives ou déductives, sont toutes définies sur des structures (CF. ANNEXE). Par exemple, l'appariement entre deux structures permet de produire, par induction, une nouvelle structure qui les subsume et qui constitue la moins générale des généralisations possibles. Ces méthodes fournissent autant de modèles pour expliquer l'acquisition des concepts au cours de l'ontogenèse.

Malgré ces atouts apparemment décisifs, l'approche moléculaireiste présente des défauts qui vont se révéler rédhibitoires, comme nous allons le voir, dans le cadre d'une modélisation cognitive de la compétence langagière.

7.4. Les limites de l'approche moléculaireiste

L'intérêt, tant pratique que théorique, de l'approche moléculaireiste de la représentation des concepts semble indiscutable, voire décisif. Pourtant, nous sommes amenée à remettre en question la plausibilité cognitive des structures postulées par cette approche, et même leur validité en tant que moyen de représenter le sens.

Problèmes de la méthode définitionnelle

La méthode définitionnelle aborde l'entreprise de la sémantique lexicale d'une manière directe et ambitieuse, en tentant d'élaborer un catalogue de définitions. Cette démarche

⁹ Une pratique courante consiste à regrouper sous le même attribut *syn-sem* l'ensemble des informations relatives à la syntaxe (comme la catégorie syntaxique, le genre ou l'aspect) et ceux qui sont relatifs à la sémantique (comme la structure temporelle, la sous-catégorisation des compléments ou la quantification).

suppose que les concepts soient ordonnés dans un treillis, chaque concept étant défini, directement ou indirectement, en fonction de tous les concepts qui lui sont antérieurs, si bien que la complexité des définitions croît le long de chaque branche du treillis. Le concept MOURIR est ainsi antérieur au concept TUER, car il entre dans la définition de ce dernier. L'aboutissement de l'entreprise nécessite un choix plausible d'un nombre limité de concepts simples et primitifs. Le concept ÉVÉNEMENT serait un concept simple, du fait que sa structure est supposée atomique, et serait un concept primitif, du fait que son ancrage est considéré comme immédiat. Or cette hypothèse de treillis définitionnel a des conséquences qui se révèlent problématiques, tant en ce qui concerne ses éléments de base que ses constructions complexes.

Le premier problème concerne l'absence de critère permettant de décider de la précédence définitionnelle. La précédence est essentielle si l'on veut éviter les circularités qui seraient fatales à l'entreprise. En mathématique, le choix de précédence n'est dicté que par des considérations techniques et l'existence de plusieurs ordres possibles est parfaitement acceptable. Lorsque l'on fait de la modélisation cognitive, en revanche, on ne peut se satisfaire d'un choix arbitraire. On pourrait penser que si la précédence définitionnelle a une existence cognitive, on doit en voir la manifestation éclatante dans les dictionnaires, qui résultent de l'activité cumulée de lexicographes humains soucieux d'organiser de manière systématique leurs intuitions concernant la définition des concepts. Or, le respect de la précédence n'est qu'un souci marginal dans le savoir-faire des lexicographes. Dans Le Robert le mot liberté est défini par référence au mot contrainte : "possibilité d'agir sans contrainte", alors que la définition du mot contrainte inclut le mot liberté : "l'entrave à la liberté d'action". Peut-on décider de la précédence définitionnelle entre les deux mots liberté et contrainte ? Bien entendu, les structures conceptuelles, dans la mesure où elles ont une réalité cognitive, n'ont peut-être que peu de rapport avec les descriptions d'un dictionnaire d'usage courant. Il n'en reste pas moins que la plausibilité d'un système de définitions conceptuelles repose de manière cruciale sur l'existence d'un critère de précédence. Or, un tel critère est généralement absent des théories concernées. La plupart des théories définitionnelles se contentent de déclarer un certain nombre de concepts comme simples et développent, pour définir les autres concepts, un ensemble de structures construites à l'aide de ces concepts simples. Un tel système ne supprime les ambiguïtés de précédence que par des choix arbitraires. Aucun critère de simplicité ou de généralité n'est proposé pour décider, par exemple, que la liberté est une absence de contrainte plutôt que l'inverse. On souhaiterait disposer d'un critère empirique, fourni par la psychologie, permettant de conclure qu'un concept B est plus complexe qu'un concept A, si bien que le concept B ne peut pas entrer dans la définition du concept A. Inversement, si le concept B est défini à partir du concept A, on s'attend à des manifestations comportementales attestant du fait que le concept B est une représentation mentale plus complexe à manipuler que le concept A, ce qui permettrait de conforter le choix de précédence. Malheureusement, la psychologie semble ne rien fournir de tel.

La méthode définitionnelle prévoit que les concepts sont d'inégale complexité, puisque certains d'entre eux en incluent d'autres en tant que composants définitionnels. Or, cette complexité définitionnelle ne se traduit par aucun effet décelable au niveau comportemental. Il serait naturel de s'attendre à ce que l'acquisition des concepts s'effectue selon des degrés croissants de complexité au cours de l'ontogenèse. Pourtant, bien que dans la plupart des théories, les concepts supposés les plus simples soient aussi les plus généraux et les plus abstraits, ce sont ceux qui sont maîtrisés le plus tard. De plus, les jeunes enfants produisent facilement des concepts leur permettant de discriminer les objets, par exemple un train électrique d'un train automoteur, alors qu'ils acquièrent plus difficilement les concepts leur permettant de catégoriser les objets par généralisation. Or, dans les modèles courants de l'apprentissage symbolique, la discrimination passe par une complexification, alors que la

généralisation est obtenue par une simplification (CF. ANNEXE). En outre, la discrimination suppose un critère permettant de sélectionner les caractéristiques pertinentes qui doivent être retenues pour différencier les objets, alors que la généralisation peut être réalisée sur les définitions d'après des critères formels généraux. Si les définitions, et les modèles symboliques qui expliquent leur acquisition, constituent une image plausible de l'établissement du système conceptuel humain, la priorité ontogénétique de la discrimination reste inexpliquée.

Le problème de la complexité définitionnelle concerne non seulement l'acquisition des concepts, mais également leur manipulation lors de l'interprétation des énoncés linguistiques et lors de la production des inférences. Les concepts les plus dérivés dans le treillis des définitions devraient normalement nécessiter un temps de traitement plus important. Ces différences de temps de traitement devraient même pouvoir servir de base à une vérification empirique de la précedence définitionnelle. Or, aucune corrélation n'a pu être mise en évidence entre la complexité supposée de telle ou telle définition et le temps de réaction dans des tâches où les concepts correspondants doivent être manipulés (FODOR & GARRETT & WALKER & PARKES 1980 [38]).

La difficulté inhérente à la méthode définitionnelle réside dans la priorité donnée au mécanisme combinatoire. Ce mécanisme engendre des combinaisons conceptuelles, de complexité croissante, qui sont attribuées à des mots. Non seulement ce modèle explique mal les modes d'acquisition et de manipulation des concepts, comme nous venons de le voir, mais c'est la conception même du système combinatoire qui est problématiques.

La combinatoire définitionnelle, interne au système conceptuel, peut produire une infinité potentielle de concepts dont très peu auront une utilité. Au nom de quel principe doit-on décider si une construction conceptuelle possède une réalité cognitive ? Il est nécessaire de disposer d'un critère de vraisemblance cognitive qui filtre les constructions conceptuelles pour ne retenir que celles qui peuvent jouer un rôle représentationnel, langagier ou inférentiel¹⁰. Une première idée consiste à considérer l'interface du système conceptuel avec la perception. Dans une vision réaliste, ce critère est externe et correspond à l'existence d'une référence dans le monde. Ce genre de critère étant à la fois indisponible et trop limitatif, il est légitime de se tourner vers la deuxième interface du système conceptuel, le langage. Cette solution a aussi ses limites. Ce n'est pas parce qu'une expression "se dit" ou "ne se dit pas" dans une langue donnée que la structure conceptuelle correspondant à cette expression peut, ou ne peut pas, être pensée. Ainsi, un persanophone dira jouer avec le Brésil là où un francophone dira jouer contre le Brésil ou même, en français contemporain, jouer le Brésil. Si ces expressions évoquent des représentations sémantiques similaires chez les locuteurs concernés, il faut expliquer comment, dans une conception compositionnelle, les prépositions avec, contre et l'absence de préposition peuvent avoir le même effet sémantique. On peut bien entendu postuler qu'il existe un concept AVEC-1, différent du concept AVEC standard, qui aurait un sens proche du concept CONTRE. Ou encore que la forme transitive du verbe jouer renvoie à un concept JOUER-1 différent du concept JOUER standard. On voit mal, dans une perspective définitionnelle produisant des structures qui reflètent le sens des mots effectivement présents dans la phrase, comment éviter ce genre d'hypothèse invérifiable et peu parcimonieuse. Une autre solution consisterait à dire que les locuteurs qui emploient les prépositions avec, contre ou la forme transitive du verbe jouer ne forment pas la même conceptualisation. Cette position est difficile à tenir, dans la mesure où ces différents locuteurs, toutes choses égales par ailleurs, vont produire exactement les mêmes

¹⁰ Les expressions d'une langue humaine obtiennent leur bonne formation du consensus de la communauté linguistique correspondante. Lorsqu'il s'agit du mentalais, le langage mental universel, le postulat de la bonne formation d'une construction plutôt qu'une autre doit être justifié par un critère cognitif.

inférences¹¹. Une troisième solution consiste précisément à considérer la troisième interface du système conceptuel, celle qui permet de produire des inférences. Ainsi, il peut être envisagé d'appliquer la théorie de la pertinence pour filtrer les représentations sémantiques dignes d'intérêt : ce serait celles qui occasionnent la production d'inférences. Utiliser le critère de pertinence de cette manière présente l'inconvénient d'être *post hoc*, car la production d'inférences à partir d'une représentation sémantique présuppose l'existence cognitive de cette représentation. Ce n'est donc pas du côté des interfaces qu'il faut chercher le critère de plausibilité cognitive des constructions conceptuelles. Le problème de la méthode définitionnelle est qu'elle n'offre pas de moyen formel pour isoler les constructions conceptuelles possibles parmi toutes celles que le modèle peut proposer. Rien n'empêche le système de tourner à vide et de produire des constructions aberrantes.

L'absence de critère de vraisemblance cognitive concerne également les éléments conceptuels introduits pour expliquer la systématique de certains phénomènes. La systématique des combinaisons et des inférences est expliquée par la ressemblance structurelle des concepts lexicaux mis en jeu. Inversement, l'existence de la systématique exige que cette ressemblance entre structures ne soit pas fortuite, qu'elle résulte d'un mécanisme d'héritage. Il faut donc partir à la recherche de ce qui, dans une famille de concepts, provoque leur comportement systématique. Or, là encore, le risque existe que soient postulées des entités conceptuelles abstraites dont le statut cognitif est peu vraisemblable. Considérons les paires suivantes.

Il croit les enfants. / Il croit les enfants malades.
Il voit les enfants. / Il voit les enfants malades.

Dans chaque paire, le syntagme *les enfants* peut avoir le statut d'argument du verbe. Cependant, les phrases de droites sont ambiguës. Considérons la lecture dans laquelle le syntagme *les enfants* est sujet de l'adjectif *malade*. Dans ce cas, le syntagme *les enfants* cesse d'être argument du verbe. Il ne s'agit plus de croire la parole des enfants, mais de croire qu'ils sont malades. Certains auteurs expliquent ce type de comportement des verbes en ajoutant un attribut à leur définition. Ainsi, les verbes *croire* ou *voir* auront un attribut "montée d'argument" qui leur permet d'accepter comme complément syntaxique un syntagme qui n'est pas un argument sémantique. Ce comportement est manifeste quand le verbe n'a pas de rôle disponible pour un complément d'objet.

*Il suppose les enfants. / Il suppose les enfants malades. / Il les suppose malades.

D'autres verbes ne posséderont pas cet attribut, ce qui explique les exemples suivants, où dans les phrases de droites la lecture selon laquelle le syntagme *les enfants* est sujet de l'adjectif *malade* est incorrecte.

Il nourrit les enfants. / Il nourrit les enfants malades.
Il rappelle les enfants. / Il rappelle les enfants malades.

Dans le cas de certains verbes, pour lesquels cette lecture est parfaitement acceptable, le choix est moins net, comme si une double lecture était possible.

Il imagine les enfants. / Il imagine les enfants malades.
Il dessine les enfants. / Il dessine les enfants malades.

¹¹ Ce point serait systématiquement mis en doute, quel que soit l'exemple choisi, si l'on se plaçait dans une version extrême du relativisme culturel. Une telle position, que l'on pourrait baptiser solipsisme culturel, décréterait que la communication n'est possible qu'au sein d'une même culture. Nous n'abordons pas cette discussion ici.

Dans ces phrases, le syntagme *les enfants* semble pouvoir être à la fois argument du verbe et argument de l'adjectif : *il dessine les enfants* et *il dessine la situation dans laquelle les enfants sont malades*. Il est donc nécessaire d'introduire un nouvel attribut sémantique pour distinguer ces cas des précédents, par exemple "montée partielle d'argument", qui serait possédé par des concepts comme IMAGINER ou DESSINER. Il semble difficile d'expliquer le comportement différent de ces verbes dans la méthode définitionnelle sans introduire quelque attribut sémantique abstrait. Or, un tel attribut risque d'être inintelligible. La plausibilité d'un attribut comme "comestible", introduit pour expliquer l'opposition entre les phrases *il mange la pomme* et **il mange la porte*, peut être envisagée sans problème, car un tel attribut concret s'interprète également par rapport à la perception et à la production d'inférences. Une catégorie sémantique abstraite comme l'attribut "montée d'argument" ne correspond à rien de perceptif qui puisse la rendre intelligible. La même situation existe en syntaxe, par exemple avec l'introduction d'entités invisibles comme les traces. Pour justifier l'existence des traces laissées par le mouvement des syntagmes, les syntacticiens montrent qu'elles expliquent des phénomènes visibles comme l'accord du participe passé. L'existence d'éléments sémantiques abstraits demande le même effort de justification, avec la différence qu'il s'agit d'hypothèses particulières concernant des classes restreintes de mots, contrairement aux traces en syntaxe qui constituent une hypothèse générale. Le problème, encore une fois, est l'absence de critère de vraisemblance cognitive. La méthode définitionnelle peut aisément expliquer tous les phénomènes langagiers, mais le risque est de postuler une quantité d'éléments sémantiques dépourvus de vraisemblance et d'intelligibilité. De plus, l'explication ainsi obtenue est de portée limitée, car elle n'enlève rien au mystère du comportement sémantique irréductible des éléments conceptuels de base qui sous-tend celui des structures.

Le comportement systématique de certaines classes de concepts dans des contextes différents est expliqué par la présence, dans les structures de ces concepts, d'un élément commun. Il s'agit sans doute là de la principale force des structures moléculaires en sémantique. Pourtant, cette solution laisse de côté un aspect fondamental du problème de l'invariance inter-contextuelle. Un élément comme "montée d'argument" explique le comportement systématique de tous les concepts qui le possèdent, mais comment expliquer son propre comportement systématique par rapport à ses arguments ? Nos exemples auraient fonctionné tout aussi bien si les expressions *enfants* et *malades* étaient remplacés respectivement par celles de *lacs* et *poissonneux*, ou encore par celles de *cyclamens* et *fanés*. Si l'élément "montée d'argument" est atomique, son comportement systématique reste inexplicable. Dans ce cas, c'est l'atout principal des structures qui s'effondre. Si le trait n'est pas atomique, on tombe sur un problème de régression, car il faudra bien *in fine* expliquer l'origine du comportement systématique par des propriétés non structurelles (FODOR 1998 [37]). On constate que le programme de la méthode définitionnelle, qui est d'expliquer l'ensemble des phénomènes sémantiques systématiques en considérant que les concepts sont des représentations fortement structurées, ne tient pas ses promesses. Une bonne partie du comportement sémantique des concepts ne provient pas des structures elles-mêmes, mais des éléments irréductibles qui les constituent.

La méthode définitionnelle repose sur la possibilité d'élaborer des structures complexes à partir d'un ensemble limité de concepts, inanalysables en termes d'autres concepts, les concepts primitifs. L'un des problèmes de cette méthode, qui n'est pas des moindres, concerne le statut de ces concepts primitifs. L'absence de consensus sur leur nature démontre que nous ne disposons pas de critère d'élimination clair permettant d'isoler ceux qui font partie de la liste. Contrairement à la chimie ou à la physique corpusculaire, dans lesquelles il existe des phénomènes permettant de conclure à la non atomicité de certains objets, on ne sait pas comment tester l'atomicité des concepts candidats pour être primitifs. Est-ce que le concept COULEUR est un concept primitif qui entre dans la structure du concept ROUGE, ou est-ce

l'inverse ? On pourrait penser que les concepts primitifs sont ceux pour lesquels les individus sont incapables de fournir une définition. Or, il semble que ce cas ne se présente jamais ! Comme le montre l'exemple des dictionnaires, le fait de tolérer la circularité permet de trouver une description pour n'importe quel mot. Dans des domaines circonscrits, certains concepts peuvent ne nécessiter aucune définition. Ainsi, en ébénisterie, il n'est pas besoin de définir le concept d'arbre. Cependant, ce caractère pseudo-primitif de certains concepts varie considérablement selon le contexte. Il semble que la capacité spontanée qui permet à chacun de former une définition dans un contexte donné corresponde à un processus complexe de nature inférentiel. Il ne consiste pas, dans la plupart des cas, à expliciter une définition stockée en mémoire. Même dans ce dernier cas, l'incohérence dans la précédence laisse peu d'espoir de découvrir la liste des concepts primitifs par introspection.

L'absence d'intuition fiable concernant les concepts simples se révèle aussi à propos de leur statut linguistique. Nous ne savons pas si ces concepts simples, dépourvus de définition, sont des concepts lexicaux, ou si à l'inverse il s'agit de concepts ineffables. Dans le premier cas, il existe potentiellement un mot m_i qui renvoie de manière non ambiguë à chacun de ces concepts primitifs p_i . L'emploi du mot m_i dans des constructions linguistiques provoque alors systématiquement l'inclusion du concept p_i dans le sens construit. Ainsi, si le concept CAUSE est considéré comme primitif, on doit pouvoir trouver un mot, le mot *cause* ou un autre, qui permette d'énoncer le concept CAUSE dans tous les contextes. Un concept structuré est supposé s'adapter, par une variété de traits, à tous ses emplois, même métaphoriques. Un concept primitif n'a pas cette ressource, si bien que l'emploi du mot associé doit être totalement rigide. Comprendre des expressions comme *cause finale* pourrait être, de ce fait, exclu. Les auteurs préfèrent généralement contourner cette difficulté en acceptant le fait que les concepts primitifs puissent être ineffables, autrement dit qu'ils ne puissent pas être nommés. Ainsi, si le concept CAUSE est primitif, il ne correspond à aucun mot, et certainement pas au mot *cause*. Mais dans ce cas, on peut s'interroger sur le statut scientifique de ce genre d'hypothèse. Il est habituel, par exemple en physique corpusculaire, de postuler l'existence d'entités auxquelles on n'a provisoirement pas accès. Cependant, la légitimité de telles hypothèses repose sur leur intérêt pour diminuer la complexité des données. Or, le fait de postuler l'existence de concepts ineffables n'élimine aucun des problèmes qui motivent la méthode définitionnelle. Ces entités ineffables conservent de manière inexplicée un comportement systématique dans des champs sémantiques différents (FODOR 1998 [37]). De plus, à partir du moment où leur existence est admise, il n'existe aucune raison pour les considérer comme atomiques, puisqu'elles peuvent contenir d'autres entités ineffables. L'hypothèse des concepts ineffables apparaît donc motivée par des considérations liées à la cohérence interne de la méthode définitionnelle, plutôt qu'à un souci d'explication. Ce défaut se révèle également lorsqu'il s'agit du statut représentationnel des concepts primitifs.

Les concepts primitifs, qu'il s'agisse de concepts lexicaux ou d'entités ineffables, sont supposés être des représentations. Leur lien avec le langage étant fortuit ou inexistant, il faut expliquer comment ils sont reliés au monde perçu. L'un des atouts de la méthode définitionnelle est qu'elle permet de déduire de la structure des concepts leurs propriétés intentionnelles, c'est-à-dire, selon le point de vue adopté dans ce travail, leur interface avec la perception. Encore faut-il que soient connues les propriétés intentionnelles des concepts qui sont à la base de toute structure, autrement dit des concepts primitifs. Or, aucun mécanisme n'est proposé dans la méthode définitionnelle pour établir l'ancrage des concepts primitifs dans la perception. La réponse classique à ce problème consiste à considérer que les concepts primitifs, dotés de leurs propriétés intentionnelles, sont innés. Il ne suffit pas de considérer que la capacité d'éprouver la sensation de rouge est une propriété innée de l'esprit humain. Il faut en outre supposer que le concept ROUGE, avec sa capacité à être déclenché dans les bonnes conditions et celles-là seulement, est inné. De même, si le concept CAUSE est considéré comme

faisant partie du répertoire des concepts primitifs, il faut supposer que les conditions d'évocation de ce concept, dans des situations variées observées dans le monde, n'attendent aucun apprentissage. Ce type de prédisposition est plus volontiers admis dans le cas des émotions. Ainsi, l'enfant n'aurait pas besoin d'apprendre les conditions dans lesquelles il doit être triste, gai ou avoir mal. En ce qui concerne les concepts, en revanche, l'idée d'une connexion fixe avec le monde perçu ne va pas de soi. Certaines données psychologiques vont dans le sens d'une prédisposition des très jeunes enfants à supposer que les noms désignent des objets simples plutôt que des parties d'objets composés ou des réunions d'objets (MARKMAN 1990 [70] ; SPELKE 1990 [100]) La démonstration est cependant purement comportementale. Le savoir-faire du bébé qui identifie correctement la présence de quatre objets n'est pas de même nature que le concept associé au mot quatre. Une distinction analogue oppose les couleurs en tant que représentations perceptives et les concepts de couleurs. Les premières seraient innées, alors que les seconds, étant soumis à une variabilité culturelle importante, ne pourraient prétendre au statut de primitives conceptuelles (SAUNDERS & VAN BRAKEL 1997 [94]). Les affirmations concernant le caractère inné de certains concepts se heurtent à l'impossibilité de les tester directement, dès lors que ces concepts sont considérés comme ineffables. Lorsque ces concepts sont nommables, comme dans le cas des couleurs, l'hypothèse semble réfutée. Nous reviendrons sur ce point (CF. CHAPITRE 8).

La méthode définitionnelle s'inscrit dans le projet d'un modèle compositionnel de la sémantique, qui suppose que le sens des composés puisse être déduit de celui des composants. Cette exigence rappelle la notion d'analyticité pour le jugement épistémique (CF. CHAPITRE 4). Un énoncé est dit analytique si la détermination de sa valeur de vérité peut se faire sur la base de sa structure sémantique interne. Ainsi, si le sens du mot *célibataire* est défini comme équivalent au syntagme *personne adulte non marié*, l'énoncé *un célibataire n'est pas marié* sera analytiquement vrai. Par analogie, nous dirons que le pari de l'approche compositionnelle est de rendre tous les énoncés sémantiquement analytiques en conférant une structure aux énoncés et en offrant les moyens de déduire leur sens à partir de leur structure. Dans une optique extensionnelle, le sens d'un énoncé est constitué par l'ensemble des situations possibles auxquelles l'énoncé s'applique en acquérant la valeur vraie. Dans une perspective cognitive, l'objectif est de prédire les conditions d'évocation des énoncés et les inférences que ces énoncés peuvent occasionner. Le problème central est de trouver les bonnes structures qui permettront d'atteindre cet objectif. L'intérêt de la méthode définitionnelle est que ce problème peut être ramené à la détermination de la bonne structure pour les concepts lexicaux. Or, comme nous allons le constater, un tel projet ne va pas de soi. L'analyticité sémantique doit s'appliquer en premier lieu aux concepts lexicaux eux-mêmes. En particulier, la définition d'un concept X doit permettre de prédire les conditions dans lesquelles le mot x est évoqué, et celles-là seulement. En d'autres termes, la définition doit, idéalement, être complète et correcte. Or, la découverte de telles définitions se heurte aux obstacles de la faisabilité pratique. Les définitions qui semblent valides sont rares. Il est même permis de questionner la possibilité de proposer une définition, ne serait-ce que pour un seul concept, en dehors des mathématiques ou des sciences formalisées. Les humains s'accordent généralement sur l'emploi des mots en contexte, ce qui rend la communication possible. Or cet accord semble échapper à toutes les tentatives de formalisation définitionnelle.

Toute tentative pour définir un concept lexical se heurte à la variété des situations dans lesquelles le mot est utilisé. La proposition consistant à assimiler le concept *TUER* à la représentation conceptuelle associée à la définition "causer la mort d'un être vivant" est vite mise en échec par l'observation de multiples cas non prévus d'application du mot *tuer*. Il est par exemple possible d'utiliser ce verbe dans l'expression *tuer l'ennui*, ce qui élargit le champ des possibilités de la classe sémantique de son complément d'objet. Il est certes toujours

possible d'enrichir une définition pour lui permettre de couvrir davantage de situations. Le problème qui se pose alors est de savoir si l'on peut borner la définition. Quelle garantie peut-on avoir qu'aucun autre cas imprévu ne se présentera ? La richesse sémantique d'un concept, quel qu'il soit, est telle qu'il semble impossible de l'enfermer dans une définition figée. On pourrait penser que pour certains concepts au moins, on dispose de définitions indiscutables. L'exemple classiquement cité est celui du concept CÉLIBATAIRE, qui est défini par la représentation conceptuelle associée au syntagme *personne adulte non mariée*. Mais les cas limites ne tardent pas à se présenter : une personne qui a fait un mariage blanc peut être considérée comme célibataire. Il faut donc inclure ce cas dans la définition, en trouvant un élément qui subsume les éléments "non marié" et "marié fictif". Que dire des individus qui sont séparés depuis dix ans ou des enfants mineurs dont le mariage est célébré avant l'âge ? Dans de nombreux contextes, ces individus seront considérés comme célibataires. Il faut donc corriger la définition. Il faudra encore la modifier pour tenir compte des amnésiques rétrogrades qui ignorent avoir été mariés, des individus dont le mariage, bien que célébré, est légalement nul, et ainsi de suite. La prise en compte des nouvelles situations conduit à une inflation de conditions disjonctives dont on ne peut pas prédire la limite.

Toute définition s'expose au risque d'être inexacte, c'est-à-dire de couvrir des situations ne correspondant pas au concept. La définition "causer la mort d'un être vivant" permet au mot tuer de s'appliquer, de manière erronée, à l'acte d'un juge qui, en condamnant l'accusé, cause sa mort. Il faut donc cette fois restreindre la définition pour l'empêcher de s'appliquer à tort, par exemple en insérant une précision supplémentaire sur le type de lien causal et sur la situation résultante. De même, la représentation conceptuelle associée au syntagme *personne adulte non mariée* inclut à tort les personnes pacsées, celles qui vivent en concubinage, *et cætera*¹². Il faut donc exclure ce cas de la définition en remplaçant l'élément "non marié" par un élément plus restreint qui soit subsumé à la fois par l'élément "non marié" et par l'élément "non pacsé". Ensuite, il faudra gérer le cas des veufs, le cas des personnes qui vivent en concubinage notoire, *et cætera*. Pourtant, le concept CÉLIBATAIRE est l'archétype du concept qui suscite l'envie de définir.

Il est légitime de se demander, si les définitions ont une quelconque réalité cognitive, pourquoi elles sont si difficiles à obtenir qu'il faille sans cesse les complexifier, soit inclusivement, soit exclusivement. Le problème vient de ce que l'on exige des définitions qu'elles soient à la fois correctes et complètes au regard des conditions d'utilisation des concepts qu'elles sont censées prédire. Or ces deux exigences, non seulement semblent impossibles à satisfaire, mais se révèlent contradictoires en pratique. Toute précision concernant un élément exclut de fait des exemplaires acceptables, alors qu'inversement le relâchement d'un élément inclut des exemplaires indésirables. L'entreprise est encore compliquée par le phénomène de gradualité.

La gradualité de la plupart des paramètres perceptifs pose un problème insurmontable aux systèmes définitionnels classiques. Personne ne confond un gland et un chêne¹³. Un gland qui vient de germer n'est pas non plus un chêne. Peut-on décider de manière non arbitraire à partir de quel moment le plant provenant du gland devient un (petit) chêne ? Bien évidemment non. Or, toute décision arbitraire expose la définition à être inadaptée pour ce qu'on attend d'elle, à commencer par la prédiction des conditions d'emploi du mot chêne. De la même façon, peut-on définir correctement un fleuve ? Un petit ruisseau côtier ou un grand affluent ne sont pas des fleuves. Le cours d'eau doit être grand et se jeter dans la mer. Mais la taille requise pour être un fleuve est une matière d'appréciation subjective. Poser un seuil

¹² Notons que ces emplois dépendent du contexte. Une personne en concubinage est considérée célibataire dans certains contextes et pas dans autres.

¹³ L'exemple est emprunté à (CHOMSKY 1975 [15]). Une version plus connue est celle du tas de sable. L'ajout d'un seul grain de sable à quelque chose qui n'est pas un tas ne donne pas un tas. Et pourtant...

arbitraire, comme “un cours d'eau se jetant dans la mer est un fleuve si et seulement si son débit dépasse 300 m³/s”, serait grossièrement insatisfaisant. Ainsi, il semble impossible de donner une définition nécessaire et suffisante des concepts qui, comme pour le concept FLEUVE, reposent sur une gradualité qui semble être décidée, à chaque emploi, par un accord subjectif entre les communicants. Une définition doit être indépendante de ce type de subjectivité. Il est toujours possible de confier la gestion contextuelle de la gradualité à des procédures qui utilisent une information autre que celle qui se trouve dans la définition. Il s'agirait là d'un aveu d'échec pour l'approche définitionnelle, car si de telles procédures existent, on peut imaginer qu'elles se chargent de l'intégralité de la construction du sens, en faisant de la recherche de définitions un jeu inutile et vain.

L'ambition qui suscite la recherche de définitions complètes et correctes dépasse largement le fait d'assurer une sorte de correspondance entre le lexique, le monde perçu et les capacités inférentielles. Il s'agit de reproduire la compétence humaine qui nous permet d'accepter certaines phrases comme sensées. Or l'outil des définitions paraît dérisoire pour cerner l'acceptabilité sémantique des énoncés, tant l'ensemble des entités couvertes par un concept semble varier à l'infini selon le contexte. Quelles sont les entités possibles qui tombent sous le concept CHAT¹⁴ ? Une peluche en forme de chat ? Une image de chat ? Un type de félin d'une taille et d'une couleur précises ? Que dire d'un chat à rayures noires et violettes ? D'un chat de quatre mètres de long ? D'un chat de quatre mille mètres de long ou de quatre mille années-lumière de long ? Certes, notre connaissance sur le monde interdit de telles entités, mais nous n'avons aucun problème pour donner un sens à ces syntagmes. Est-ce à dire que la dimension d'un chat ne fait strictement pas partie de sa définition ? Il s'agit pourtant d'une connaissance dont on dispose lorsque l'on connaît le concept CHAT. Nous utilisons ces connaissances pour produire des inférences. Par exemple, nous utilisons la connaissance concernant les dimensions normales d'un chat pour comprendre qu'il peut échapper à un chien en fuyant par un trou. Il n'en reste pas moins que tout ce que nous savons à propos d'un chat, qu'il s'agit d'un mammifère, qu'il possède deux poumons, qu'il ne fait pas un kilomètre de long, *et cætera*, peut être mis en défaut dans certains contextes, et pourtant nous continuons à reconnaître le concept CHAT.

Pour échapper à la critique de l'imprévisibilité du jugement sémantique, il est possible de séparer la connaissance sémantique, inscrite dans les concepts, d'une autre forme de connaissance, stockée sous forme de règles dans le système inférentiel. Ainsi, nous saurions qu'un kangourou possède deux poumons, non parce que cette connaissance fait partie du concept KANGOUROU, mais parce qu'elle fait partie d'un corps de connaissances zoologiques de base. Ces connaissances propres au système inférentiel ne produiraient, le plus souvent, que des inférences non systématiques. Par exemple, un sujet pourrait décider qu'un serpent est dépourvu de poumons sur la base de son apparence (absence de cage thoracique visible), ou décider qu'il en possède parce que tous les reptiles en possèdent. Ce genre d'inférences, liées au raisonnement, utiliserait des connaissances variées sur les concepts, sans dépendre exclusivement des concepts eux-mêmes. Cette solution, cependant, crée autant de problèmes qu'elle en résout. Certes, il devient difficile de réfuter le système définitionnel en lui opposant la variabilité de l'acceptabilité sémantique. Mais ce résultat se paie de la nécessité de concevoir une deuxième forme de connaissance, et d'organiser la compétition entre les deux formes de connaissances. Comment tracer la limite entre ce qui relèverait de la connaissance purement sémantique, inscrite “en dur” dans les concepts, d'une connaissance plus labile, sensible au contexte, qui serait stockée en dehors des concepts, dans une base de connaissances (FODOR 1994 [36]) ? Là encore, aucun critère n'a été proposé, tant chacune des connaissances propres à un concept semble pouvoir être réfutée par le contexte sans que le

¹⁴ Cet exemple est adapté de (FODOR 1994 [36]).

concept le soit. Le projet d'une double connaissance, dans l'état, reste caduc sur le plan théorique. Le rêve d'analyticit , avec ses d finitions compl tes et correctes d pourvues de toute ambigu t , semble donc illusoire. Il existe plusieurs man res pour tenter de sauver une partie du r ve. L'une d'elle est de se contenter de d finitions partielles.

Les auteurs ont, depuis longtemps, per u   quel point il  tait difficile de trouver des d finitions qui enferment tout ce qu'un concept contient de potentialit s cognitives (JACKENDOFF 1983 [49] ; JACKENDOFF 1990 [50]). Par exemple, les verbes fr ler et effleurer ne sont pas d'exactes synonymes (les voitures se sont fr l es / il a effleur  sa chevelure). On peut accepter que les nuances de sens entre les verbes fr ler, effleurer, caresser, toucher,  rafler, heurter, cogner, frapper, emboutir, tamponner soient en partie li es   la perception des situations correspondantes et ne puissent pas  tre rendues de mani re suffisamment pr cise au sein d'une d finition. De fait, il n'est pas facile de distinguer entre les verbes fr ler et effleurer dans une d finition conceptuelle. Toutefois, les comportements syntaxiques et inf rentiels de ces deux mots, dans ce qu'ils ont de syst matique, sont   peu pr s les m mes. On peut imaginer de ramener le programme d finitionnel   la d couverte, pour chaque concept lexical, d'un ensemble minimal de caract ristiques n cessaires, que l'on pourrait appeler le noyau d finitionnel, permettant d'expliquer la syst maticit  de l'emploi du mot et la production des seuls ph nom nes que l'on peut observer syst matiquement pour ce concept. En d'autres termes, l'ambition ne serait plus de reproduire tous les ph nom nes cognitifs li s au langage et au raisonnement, mais simplement les ph nom nes syst matiques. Ainsi, la d finition du concept TUER donn e par une repr sentation comme cause ( v nement (passage (vie, mort))) est cens e fournir une structure suffisante pour assurer l'interface avec la syntaxe et le calcul des inf rences syst matiques. Ce noyau d finitionnel est cens  expliquer, en outre, le fait que le concept conserve certaines caract ristiques quand il est utilis  dans des champs s mantiques et des contextes diff rents. En revanche, le noyau d finitionnel n'est suffisant ni pour caract riser les occurrences du concept (comme l'expression tuer l'ennui), ni pour produire l'ensemble des compositions ou des inf rences non syst matiques (comme l'utilisation d'une arme dans l'acte de tuer). C'est donc une notion restreinte de compl tude qui est adopt e ici, limit e   la production des ph nom nes syst matiques. Gr ce   cette restriction, on semble  viter certaines des difficult s signal es plus haut, notamment les soucis li s   la gradualit  et au caract re ouvert de l'acceptabilit  s mantique. Par exemple, le caract re graduel que l'on peut d couvrir dans l'acte de tuer pourra  tre report  sur la gradualit  de la fronti re entre vie et mort. De m me, la variabilit  de l'acceptabilit  s mantique du mot tuer pourra  tre imput e   celle de cause. En dehors de ces ph nom nes, le noyau d finitionnel semble pouvoir jouer son r le par rapport   la syntaxe et   la production des inf rences syst matiques. En r alit , une bonne partie des probl mes subsistent. Les difficult s concernant la correction des d finitions demeurent. De plus, la notion restreinte de compl tude, li e   la production de toutes les inf rences syst matiques, est probl matique, car on ne sait pas tracer la ligne de d marcation entre les  l ments qui expliquent la syst maticit  et ceux qui sont, de ce point de vue, superflus. Cette s paration est fortement d pendante du contexte, comme on peut s'en convaincre avec l'exemple du concept C LIBATAIRE. Selon que le contexte a trait   la moralit ,   la fiscalit ,   l'h ritage ou   l'adoption, les  l ments qui permettent la production des inf rences que les individus font syst matiquement connaissent une variation significative. Rajoutons que la nature des  l ments caract ristiques des concepts qui permettent d'effectuer les op rations non syst matiques reste myst rieuse. Pour certains auteurs, il s'agirait d' l ments sub-conceptuels, faisant partie du syst me perceptif (JACKENDOFF 1983 [49] ; JACKENDOFF 1990 [50]). Ainsi, la distinction entre fr ler et effleurer ne serait pas conceptuelle, mais perceptuelle. Pour que cette id e, qui revient    vacuer une bonne partie de la connaissance lexicale hors du champ de la s mantique, soit acceptable, il faut pr ciser

comment les deux types de connaissance s'articulent. Le lien est essentiel pour expliquer l'apprentissage des concepts ou tout simplement la verbalisation d'une scène perçue. Enfin, il est regrettable qu'une bonne partie du comportement des concepts soit ainsi renvoyée hors du modèle, que ce soit la gradualité, l'invariance inter-contextuelle ou la flexibilité de l'acceptation en contexte (FODOR 1998 [37]). Il est donc légitime de douter non seulement des chances de succès de la méthode définitionnelle, même si ses ambitions sont revues à la baisse, mais également de son intérêt théorique.

L'essentiel des difficultés rencontrées lors de la détermination de définitions acceptables et efficaces invite naturellement à explorer une forme atténuée de définition, qui ne porterait que sur des caractéristiques régulières. Les caractéristiques cherchées ne sont plus ni nécessaires, ni suffisantes. Elles permettent de reconstituer, pour un contexte standard ou moyen, les inférences ou le jugement d'acceptabilité habituellement liés au concept considéré (SOWA 1984 [99]). Par exemple, la notion de scénario permet de décrire avec un certain détail les conditions habituelles d'emploi d'un concept donné, ce qui permet de produire une variété de compositions et d'inférences. D'un autre côté, un scénario ne peut pas être contredit, puisqu'il ne s'applique que par défaut. Quel que soit l'intérêt de cette technique, elle ne répond pas à ce qui fait la raison d'être des définitions, l'explication de la systématisme, car c'est précisément cette idée qui est remise en question avec la notion de scénario.

Une autre technique consiste à introduire, dans les définitions, des plages d'admissibilité censées représenter la typicité (SMITH & MEDIN 1981 [96]). Lorsque l'objectif consiste simplement à remplacer les contraintes binaires par des seuils ou du flou, nous avons toujours affaire à des définitions. Ces définitions quantitatives présentent les mêmes écueils que les définitions binaires : il faut déterminer les différents éléments qui doivent figurer dans la définition, et pour chaque élément, sa plage d'admissibilité. De plus, l'application du principe de compositionnalité à ce type de structure devient problématique (CF. CHAPITRE 4).

En revanche, lorsque des valeurs numériques ou des fonctions continues sont introduites pour pondérer la présence des éléments de la définition, on sort du paradigme définitionnel classique. La notion de structure formelle, sujette à des manipulations elles-mêmes formelles, se retrouve remise en question. En particulier, l'introduction de coefficients de pondération remplace la constituance des structures définitionnelles par une forme de pseudo-constituance mal définie. On ne sait pas si le concept de TITRE ou celui d'INDEX font partie intégrante du concept de LIVRE s'ils sont affectés de pondérations valant respectivement 0,99 et 0,05. Il ne s'agit pas d'affirmer ici que des structures contenant des éléments numériques sont sans intérêt pour modéliser les concepts. Notre critique porte ici sur le fait que les coefficients numériques affectent la définition récursive des concepts moléculaires. C'est donc le mélange ontologique entre le graduel et le définitionnel qui est ici problématique. Le résultat en est un outil, plutôt qu'un modèle cognitif plausible de la structure des concepts.

Enfin, signalons que les modèles prônant les techniques définitionnelles ne proposent généralement aucun dispositif par lequel les individus testeraient la cohérence et la non circularité de leurs définitions.

Ces critiques de la méthode définitionnelle nous semblent réhivitoires. Un premier groupe de problèmes est lié à l'existence même des structures définitionnelles : le problème de la précédence, l'absence d'effet de complexité, le manque de vraisemblance cognitive. Un autre groupe de problèmes concerne les fondements de la méthode, les concepts primitifs : l'absence d'explication pour leur invariance inter-contextuelle, pour leur statut linguistique, pour leurs propriétés intentionnelles, et l'absence de critère d'élimination permettant d'en isoler la liste.

Le troisième groupe de problème concerne la difficulté de rendre les énoncés sémantiquement analytiques, liée à l'impossibilité d'obtenir des définitions complètes et

correctes. Enfin, nous avons donné des arguments montrant que des versions affaiblies de l'approche, que ce soit la recherche de noyaux définitionnels ou le repli sur des définitions partielles, ne pouvaient pas prétendre résoudre la question de la systématique en sémantique. Cette accumulation de réfutations nous oblige à voir dans la méthode définitionnelle rien de plus qu'un outil technique commode pour des réalisations pratiques de traitement du langage et du raisonnement, sans plausibilité cognitive. Nous nous tournons maintenant vers la seconde méthode permettant de structurer les concepts.

Problèmes de la méthode relationnelle

La méthode relationnelle peut sembler plus prometteuse, puisqu'elle n'exige pas de listes d'éléments définitoires. Surtout, elle ne pose pas en principe l'exigence de non circularité qui se révèle si problématique avec l'approche définitionnelle. Elle échappe également à la difficile question des concepts primitifs. Il faut cependant admettre que la méthode relationnelle est loin de représenter la solution du problème de la représentation sémantique.

L'objectif premier de la méthode relationnelle est de représenter la signification des entités lexicales par la place que chaque concept lexical occupe dans un réseau conceptuel. Lorsqu'on s'intéresse à la combinaison des concepts, pour représenter le sens des syntagmes et des phrases, la manipulation de ce réseau est moins immédiate. À quelle partie du réseau correspond le sens du syntagme le chat de Jean dans la phrase le chat de Jean court ? Le nœud associé au concept COURIR demande la présence d'un autre nœud auquel il sera lié par une arête représentant la relation thème. Il pourrait s'agir d'un nouveau nœud, un nœud recruté, pointant vers le sens du syntagme le chat de Jean. Or, les nœuds étant sémantiquement atomiques, le système obtenu se retrouve dépourvu de la propriété de constituance, sans offrir de moyen alternatif d'obtenir la systématique. Dans la méthode relationnelle, un concept est caractérisé par un nœud et un ensemble d'arêtes le reliant à d'autres nœuds. Comment garantir la systématique de certaines inférences lorsque les informations qui permettent ces inférences dépendent de la présence d'un ensemble de liens dans le réseau ? Par exemple, les deux phrases le chat miaule et le chat de Jean miaule déclencheront toutes deux l'inférence conduisant au fait que le chat ouvre la bouche. Cette inférence ne peut être possible que par la présence d'une série de relations entre le concept CHAT et d'autres concepts. Or, si la représentation conceptuelle associée au syntagme le chat de Jean ne contient pas, dans le sens du principe de constituance, le concept CHAT avec l'ensemble de ses relations, comment cette inférence peut-elle se produire ? L'absence fortuite d'une seule de ces relations peut anéantir la systématique. Or, rien ne permet d'assurer la présence effective de toutes les relations requises. La perte de constituance, dans ce cas, semble donc rédhibitoire pour la systématique.

Le sens du syntagme le chat de Jean correspond plus naturellement à un sous-graphe du graphe conceptuel global. Un sous-graphe possède une structure interne, ce qui permet d'assurer la constituance si le sous-graphe contient les représentations, elles-mêmes sous forme de sous-graphes, de ses constituants. En revanche, le mécanisme compositionnel qui permet de combiner les sous-graphes n'est pas immédiat¹⁵. Ainsi, pour assurer que le sens de la phrase le chat de Jean court contient celui du verbe courir et celui du syntagme le chat de Jean, il faut disposer d'un moyen de connecter dynamiquement ces deux sous-graphes, par exemple à l'aide de variables. La difficulté vient du fait qu'un tel système n'est pas récursif. Ainsi, dans la phrase le chat de Jean court plus vite que le chat de Jacques, le sous-graphe correspondant au syntagme le chat est supposé intervenir dans deux syntagmes différents, ce que le système

¹⁵ Il est possible, dans un graphe, de substituer un sous-graphe à un nœud, à condition que le sous-graphe ait le même nombre d'arêtes adjacentes que le nœud. Pour ce faire, il faut déterminer d'avance les nœuds dans le sous-graphe auxquels les arêtes du nœud remplacé seront attribuées.

de variables ne permet pas. Une telle difficulté peut être contournée en considérant que le concept CHAT peut être instancié plusieurs fois dans le graphe. Par exemple, dans un langage de trames, la solution consiste à juxtaposer deux exemplaires de la trame associée au concept CHAT dont les liens *propriétaire* sont spécifiés différemment. Cependant, cette solution conduit à modifier dynamiquement la structure du réseau. Or, une telle opération ne va pas de soi. Faut-il dupliquer les arêtes qui pointent vers la classe *chat* pour que les trames *chat-1* et *chat-2* puissent jouer le même rôle sémantique que leur modèle ? Rien n'indique que ces manipulations, jointes à celles qui sont nécessaires pour représenter le sens de la phrase, n'auront pas de conséquences sur la cohérence de l'ensemble du réseau. Une telle solution est donc peu envisageable. L'intérêt des réseaux conceptuels est avant tout de stocker la connaissance dans une structure de graphe statique qui offre des possibilités de parcours dynamiques. Dans cette acception, l'impossibilité de suivre la récursivité de la syntaxe entraîne une perte de compositionnalité qui limite fortement les ambitions du modèle.

Une autre critique fondamentale que l'on peut adresser à la méthode relationnelle, lorsqu'il s'agit de la considérer comme un modèle cognitif plausible, est qu'elle utilise une ontologie non bornée. Rien, dans la théorie, ne vient limiter le nombre des types différents de nœuds ou d'arêtes qui sont nécessaires pour représenter le sens, dès que le domaine conceptuel à représenter augmente. Il est possible d'éviter tout typage pour les nœuds du graphe relationnel, autrement dit de n'utiliser le typage que par commodité. Dans ce cas, le rôle sémantique d'un concept provient entièrement de l'ensemble de relations qui le relie aux autres concepts du réseau conceptuel. Il n'en va pas de même pour les arêtes. La seule manière de produire des inférences à partir du réseau conceptuel est que les relations soient connues ou que leurs propriétés puissent être déduites. Ainsi, le système inférentiel doit connaître la relation *résultat* pour inférer que son élément cible est réalisé dans le futur lorsque son premier argument est réalisé dans le présent. En conséquence, soit la liste des relations possibles est fixe et donnée à l'avance, soit elle dérive d'un typage productif. La première solution impose une forte contrainte à la méthode relationnelle, en bornant *a priori* l'ensemble des relations disponibles aux rôles sémantiques fondamentaux. La deuxième solution réintroduit une combinatoire définitionnelle pour les relations, avec les inconvénients que nous avons considérés dans la section précédente. De plus, les deux solutions réintroduisent des éléments primitifs dans le système.

L'existence d'un réseau conceptuel, quelle que soit l'ontologie sur laquelle il est bâti, pose le problème de savoir comment ce réseau se met en place au cours de l'apprentissage. Contrairement à la méthode définitionnelle qui est supposée élaborer des structures en combinant les concepts primitifs, on voit mal, si l'on adopte la méthode relationnelle, par où aborder la mise en place du réseau. Il est possible d'imaginer un mécanisme permettant de créer de nouveaux nœuds et d'établir des arêtes entre eux et les nœuds existants. Cependant, dans la mesure où ces nœuds sont atomiques et sont liés à la perception par des liens d'évocation, l'apprentissage de leur intentionnalité se pose avec la même intensité que lorsqu'il s'agissait de comprendre l'origine de l'intentionnalité des concepts primitifs dans l'approche définitionnelle. La difficulté est même accrue, puisqu'il s'agit de l'intentionnalité de chaque nœud du graphe. La solution qui consiste à considérer cette intentionnalité comme innée est donc difficilement envisageable. Nous reviendrons sur ce problème (CF. CHAPITRE 8). L'apprentissage des relations se révèle tout aussi problématique. Aucun mécanisme évident ne permet de concevoir un système pour créer de nouveaux types de relations. Les types de relations utilisées dans les graphes conceptuels, par exemple, n'existent qu'au sein du graphe. Elles ne peuvent pas être inférées de la seule perception. La relation *instrument* est une relation conceptuelle. Le sujet qui possède cette relation peut rechercher dans le monde perçu des indices lui permettant de l'activer. En revanche, aucun mécanisme d'apprentissage, statistique ou non, n'a été proposé qui permette à un sujet de

créer la relation *instrument de novo*. Les mécanismes d'apprentissage symbolique qui ont été proposés reposent tous sur l'instanciation de patrons relationnels (CF. ANNEXE). Dans l'état actuel de la théorie, il faut donc supposer qu'un certain nombre de schémas généraux de relations sont disponibles dès le départ, de manière innée. Ainsi, la relation *instrument* peut être vue comme dérivant d'un schéma général d'action causale, par exemple en tant que cause intermédiaire. Toutefois, si les relations sont dérivées de patrons relationnels, elles possèdent nécessairement une structure interne, ce qui nous fait retomber dans le schéma définitionnel. On perd alors tout espoir de se démarquer de l'approche qui a été critiquée dans la section précédente.

Si la mise en place du réseau conceptuel est délicate à modéliser, il en est de même de ses transformations. Certaines théories psychologiques font état de changements radicaux et coordonnés de la connaissance au cours du développement (PIAGET 1932 [82]). En particulier, l'accès à de nouvelles opérations peut entraîner une réorganisation des connaissances relatives à un domaine jusqu'à ce que ces connaissances forment une nouvelle structure. Ainsi, le jeune enfant se forme une théorie de la justice en opérant plusieurs transitions. Il passe en particulier d'un stade égocentrique, où est injuste tout ce qui l'affecte négativement, à un stade symétrique où sa propre action peut lui apparaître comme injuste envers les autres. Or, ce type de transition semble soudain. À partir du moment où l'enfant est capable de se mettre mentalement à la place des autres, il opère rapidement tous les changements qui lui permettent d'accéder au stade conceptuel ultérieur. La méthode relationnelle ne contient pas, dans sa panoplie théorique, de moyens convaincants pour expliquer de telles transitions épistémiques. Il lui manque en particulier l'idée de clôture opérationnelle qui contraint, dans le modèle du développement que nous venons d'évoquer, les formes que peut prendre la connaissance d'un domaine. C'est parce que ces formes sont fortement contraintes que l'on observe de soudaines transitions entre stades, ce que la méthode relationnelle ne peut expliquer.

Le problème de l'apprentissage se retrouve pour les modifications et la mise à jour du réseau. Lorsqu'un enfant apprend à distinguer le sens des adjectifs maigre et étroit¹⁶, comment est-il supposé modifier son réseau conceptuel ? Le mécanisme d'apprentissage doit être capable de "lire" le graphe pour opérer les modifications qui s'imposent (HOFSTADTER 1995 [47]). Le problème ne se limite pas à concevoir un mécanisme de mise à jour du graphe. Il faut également être en mesure d'en limiter l'application. Or, aucune garantie n'est offerte pour borner l'étendue des mises à jour à opérer dans le réseau, suite à une modification ponctuelle. Ce problème, connu sous le nom de problème du cadre, provient du fait que toute modification risque de provoquer le besoin d'une nouvelle mise à jour, si bien que rien n'exclut en théorie que l'ensemble du graphe s'en retrouve affecté (MCCARTHY 1963 [71]). La difficulté disparaît si l'on se contente d'une approximation, par exemple en bornant *a priori* le nombre de mises à jours que l'on effectue en cascade. On peut imaginer que les individus se contentent et s'accommodent de ce genre de limitation. Le problème est que les graphes ainsi produits sont les fruits de mises à jour partielles dont la cohérence n'est pas garantie. Rien n'est proposé pour gérer ce genre de graphe, ni pour en vérifier la cohérence à un moment donné en un temps raisonnable.

Un problème similaire se pose lorsque l'on cherche à délimiter un concept lexical dans le réseau conceptuel. Si la mise à jour de la signification d'un mot peut faire sentir ses effets dans n'importe quel autre endroit du réseau, il n'est plus possible d'isoler le sens d'un mot. Ce sens dépendra, de proche en proche, du sens de tous les concepts représentés dans le réseau. Une attitude, face à ce problème, consiste à assumer l'impossibilité d'isoler le sens en adoptant la position du holisme conceptuel. Cette attitude de principe, quoique séduisante, est

¹⁶ Exemple réel.

malheureusement impraticable. S'il n'existe aucun moyen d'isoler les sens des différents concepts les uns des autres, c'est la notion même de concept lexical qui disparaît (FODOR & LEPORE 1992 [39]). Comment, dans ces conditions, définir théoriquement ce qui est appris lorsque l'usage d'un nouveau mot est maîtrisé ? Quel sera le mécanisme d'apprentissage correspondant ? À supposer que tout cela soit disponible, il reste que le sens d'un mot pour un individu A ne peut pas être comparé au sens du même mot pour l'individu B, si bien qu'on perd tout espoir de modéliser la communication entre A et B. On perd même la possibilité de modéliser l'acquisition des concepts chez les sujets, puisque le sens d'un mot à un moment donné et son sens à une époque ultérieure sont incommensurables.

L'alternative au holisme conceptuel consiste à introduire un principe de localité. Chaque concept est identifié à un sous-réseau impliquant un ensemble déterminé d'autres concepts. Un tel principe permet d'envisager des mécanismes d'apprentissage, de maintien de cohérence et de communication qui opèrent localement, même si d'autres mécanismes comme l'association conservent une portée holistique. Le problème est de fournir un critère opérationnel de localité. Même à supposer que l'on dispose d'un tel critère, on se retrouve dans la situation de l'approche définitionnelle : on ne sait pas déterminer les relations qui composent le sous-réseau local d'un concept donné (FODOR 1994 [36]). Il est d'ailleurs possible, pour chaque concept, de convertir le sous-réseau qui le caractérise en une définition, si bien que le système se retrouve équivalent à un système définitionnel, avec l'exigence de non circularité en moins.

Conclusion

Essayons de récapituler les différentes critiques que l'on peut adresser aux modèles moléculaires du système conceptuel.

Liste des critiques adressées à la méthode définitionnelle

Dans la dernière section de ce chapitre, nous avons formulé un certain nombre de problèmes concernant la méthode définitionnelle.

- Absence de critère de précedence pour ordonner le treillis des définitions.
- Absence d'effet psychologique de la complexité définitionnelle. Les concepts généraux ne sont pas acquis en premier et ne sont pas manipulés plus rapidement.
- Absence de filtre pour éviter les constructions invraisemblables de la combinatoire. Les candidats (conserver ce qui existe, ce qui se dit ou ce qui produit des inférences) sont inadéquats.
- Absence de méthode pour justifier la vraisemblance des éléments sémantiques abstraits introduits de manière *ad hoc*.
- Absence d'explication de l'invariance inter-contextuelle. Le comportement systématique des composants fondamentaux reste inexpliqué.
- Absence de critère d'élimination pour décider quels sont les concepts primitifs.
- Absence de statut linguistique pour les concepts primitifs. Tant la correspondance lexicale stricte que l'ineffabilité sont problématiques.

- Absence d'explication de l'intentionnalité des concepts primitifs.
- Impossibilité de trouver des définitions correctes et complètes.
- Absence d'explication de la gradualité sémantique.
- Absence d'explication de la souplesse de l'acceptabilité sémantique en contexte. Tous les attributs définitionnels peuvent être mis en défaut.
- Solution : isoler une partie de la connaissance relative au concept dans un noyau conceptuel pour laisser le reste à une base de connaissances contingentes.
 - > Absence de critère de délimitation.
 - > Absence d'articulation entre les deux ingrédients conceptuels.
 - > Persistance du problème de correction.
- Solution : isoler une partie de la connaissance relative au concept portant sur des caractéristiques régulières.
 - > Absence de critère de délimitation.
 - > Perte de la systémativité.
- Solution : remplacer les contraintes binaires des définitions par des plages d'admissibilité.
 - > Persistance des problèmes concernant le choix de la structure définitionnelle.
 - > Perte de la compositionnalité.
- Solution : affecter les traits définitionnels d'un coefficient pondérant leur probabilité de prise en compte.
 - > Perte de la constituance lexicale ; on sort du paradigme définitionnel.
- Absence d'explication des mécanismes permettant le contrôle de la cohérence et de la non circularité des définitions.

À ces différentes critiques, nous ajouterons deux problèmes que nous évoquerons dans les chapitres suivants, et qui ne sont certainement pas les moindres.

- Absence de mécanisme pour contrer la monotonie de la composition (CF. CHAPITRE 8). L'assemblage des définitions pour former le sens des énoncés conduit à des représentations de taille non plausible.
- Absence d'explication de l'omnipotence de la conceptualisation (CF. CHAPITRE 9). Nous pouvons conceptualiser tout ce que nous percevons. Or, nous ne pouvons avoir de duplication conceptuelle pré-stockée pour toute représentation perceptuelle.

Liste des critiques adressées à la méthode relationnelle

Nous nous sommes ensuite attachée à répertorier un certain nombre de problèmes concernant la méthode relationnelle.

- Problème du calcul du sens des syntagmes : la représentation du sens d'un syntagme par un nœud recruté fait perdre la constituance et donc la systémativité.

- Problème du calcul du sens des syntagmes : la représentation du sens d'un syntagme par un sous-graphe fait perdre la compositionnalité ; la duplication dynamique des nœuds ne duplique pas les sous-graphes.
- Problème de l'ontologie non bornée : même s'il est possible d'éviter de typer les nœuds, ce n'est pas possible en ce qui concerne les arrêtes.
- Problème de la mise en place du réseau : les mécanismes d'acquisition des concepts, tant pour les nœuds que pour les arrêtes, restent inexpliqués ; le postulat de patrons relationnels ramène la solution au paradigme définitionnel.
- Problème des transitions épistémiques : les mécanismes de réorganisation majeure du réseau restent inexpliqués.
- Problème de la mise à jour du réseau : les mécanismes de modification des concepts, tant pour les nœuds que pour les arrêtes, restent inexpliqués ; le problème du cadre met en doute la possibilité des modifications locales et du maintien de la cohérence.
- Transposition au problème du holisme : l'acceptation du holisme fait perdre tout pouvoir prédictif au modèle.
- Insuffisance d'un principe de localité : la recherche d'un critère de délimitation ramène la solution au paradigme définitionnel.

Il faut, là encore, ajouter les problèmes liés à la monotonie et à la duplication que nous considérerons dans les chapitres suivants.

Le problème de l'acquisition des concepts simples, dépourvus de structure, a été simplement mentionné dans ce chapitre. Il fera l'objet d'un examen plus approfondi dans le chapitre suivant.

On constate ainsi que les modèles moléculaires, qu'ils soient définitionnels ou relationnels, présentent des inconvénients que l'on peut considérer comme rédhibitoires. Au premier abord, ces modèles semblent résoudre les problèmes de systématisme par l'emploi de structures riches, ce qui représente un avantage considérable. *A posteriori*, on s'aperçoit que la méthode consiste à vouloir enfermer dans des structures statiques ce qui n'est que la trace de mécanismes compositionnels dynamiques. Par leur aspect statique, ces représentations répondent mal à la sensibilité au contexte et se retrouvent très souvent inadéquates. La solution qui consiste à augmenter encore davantage la richesse des structures pour tenter d'anticiper toutes les situations d'emploi des concepts est vaine, car on n'en perçoit pas la limite. Si l'on abandonne le projet moléculaire, il devient légitime d'envisager des concepts dépourvus de toute structure, sur lesquels opéreraient des mécanismes compositionnels puissants.

Chapitre 8 :
Atomisme conceptuel

Introduction

Les théories qui décrivent la sémantique du langage naturel en tant qu'un calcul symbolique opérant sur des entités mentales présupposent la notion d'atomes de sens, définis comme les entités mentales les moins complexes dotées de propriétés représentationnelles et causales. Pour la plupart des adeptes de l'approche moléculariste, ces atomes de sens sont les quelques concepts simples à partir desquels tous les autres sont déterminés. Les autres représentations conceptuelles, qu'il s'agisse des concepts lexicaux ou des représentations associées aux syntagmes ou aux phrases, sont des représentations composées, que l'on peut hiérarchiser selon leur degré de complexité. Dans cette hiérarchie, les concepts lexicaux n'occupent pas de place privilégiée. Seules des considérations externes, liées aux mécanismes grammaticaux ou au système inférentiel, permettent de distinguer les représentations conceptuelles qui s'expriment par un seul mot, par un syntagme ou par une proposition. Par exemple, des représentations conceptuelles peuvent se distinguer parce que l'une s'exprime verbalement par un verbe, et constitue ainsi un concept lexical, une autre par un syntagme verbal doté d'une structure grammaticale arborescente, une autre encore exprimera une pensée propositionnelle susceptible de provoquer un jugement de vérité. Pourtant, dans l'hypothèse moléculariste, il s'agit dans chaque cas d'agencements conceptuels qui pourront avoir des structures proches.

L'originalité de l'atomisme conceptuel n'est donc pas de postuler l'existence d'atomes de sens. L'hypothèse est plus forte. Elle consiste à imposer une condition formelle sur le niveau de la complexité des concepts lexicaux. Selon l'approche atomiste, la plupart des mots du lexique évoquent des représentations mentales dépourvues de structure interne de type conceptuel¹.

I'm interested in such questions as [...] : 'What is the structure of the mental representation DOG?' And my answer will be that, on the evidence available, it's reasonable to suppose that such mental representations *have no structure*; it's reasonable to suppose that they are atoms.
(FODOR 1998 [37] p. 22)

L'atomisme conceptuel se traduit, de manière immédiate, par un changement de granularité par rapport à l'approche moléculariste. Le calcul du sens reste une sorte de chimie, mais au lieu de ne considérer que les seules primitives conceptuelles comme atomiques, c'est l'ensemble des concepts lexicaux qui deviennent inanalysables conceptuellement. Cependant, l'atomisme conceptuel est motivé par des considérations plus fondamentales, liées au statut représentationnel des concepts.

8.1. Le refus des descriptions

La motivation principale pour considérer les concepts lexicaux comme atomiques est essentiellement négative. Elle est liée à l'impossibilité d'établir des structures adéquates pour ces concepts. Elle repose aussi sur le refus de voir le lien qui unit les concepts lexicaux aux représentations perceptuelles dépendre du résultat d'un calcul sur une structure.

¹ L'hypothèse n'interdit pas le fait que des mots dont la structure morphologique est transparente, comme antiparlementaire, puissent être associés à une représentation composée.

Les approches molécularistes ont un point commun : les concepts y sont caractérisés par les conditions de leur usage. La nature intrinsèque des concepts apparaît donc comme seconde, déduite du rôle qu'ils jouent par rapport au langage et au système inférentiel.

The substance of current theories lies in what they say about *the conditions for having the concept X*. It's the story about *being* the concept X – the story about concept *individuation* – that they treat as parasitic: the concept X is just *whatever it is that a creature has* when it has the concept.
(FODOR 1994 [36] p. 98)

Les approches molécularistes s'attachent à caractériser les représentations conceptuelles liées aux entités lexicales de manière à expliquer leur capacité à engendrer le comportement verbal et le raisonnement. La tendance naturelle est de privilégier cette capacité, autrement dit de considérer que les concepts se résument à la structure qui permet d'expliquer les performances verbale et inférentielle. De la même façon que les pièces d'un jeu d'échecs prennent leur sens par le rôle qu'elles jouent dans la production du jeu, on peut ne voir dans les concepts que ce qui explique les interactions qu'ils ont avec les autres concepts, dans le but d'expliquer la mécanique conceptuelle qui sous-tend le langage et le raisonnement. Cette conception rappelle la métaphore du dictionnaire. Dans un dictionnaire, chaque mot est paraphrasé par des synonymes ou des périphrases dans des contextes différents, il est caractérisé par son champ et ses traits sémantiques, par des indications sur son usage, *et cætera*. Autrement dit, un mot est décrit par le système dans lequel il existe, à savoir le langage. Dans l'approche moléculariste, ce principe est reproduit tel quel dans le système conceptuel : les représentations conceptuelles associées aux mots sont caractérisées par une description en termes d'autres représentations conceptuelles.

La version radicale de cette idée est que le contenu des mots est entièrement constitué par les relations inter-conceptuelles. Dans cette version le caractère représentationnel des concepts, leur lien avec le monde perçu, est oublié. Le système symbolique, à la manière d'un jeu, acquiert une indépendance totale, l'interprétation n'existe plus. Avoir des concepts revient à pouvoir jouer le jeu linguistique ou le jeu inférentiel. Le risque d'une telle conception est d'obtenir une mécanique qui tourne à vide, sans ancrage dans les situations que le langage est censé représenter et communiquer. À cela s'ajoute la difficulté de concevoir le jeu conceptuel, en tant que contrepartie sémantique du comportement linguistique, comme préexistant à la caractérisation des concepts eux-mêmes, un peu comme si le jeu d'échecs pouvait être défini indépendamment des pièces du jeu.

La version modérée est que les concepts, certes, servent à représenter le monde perçu, mais ce n'est pas ce qui les caractérise. L'essentiel réside dans leurs propriétés causales, qui déterminent leur participation dans les mécanismes de calcul. Un bon modèle se doit alors de définir une structure formelle pour ces concepts qui explique leur comportement dans les aspects computationnels de la cognition. La force de cette lignée moléculariste est de fournir, comme un sous-produit, une explication de l'ancrage des concepts lexicaux, grâce aux principes de compositionnalité et de constituance au niveau lexical. Les structures moléculaires sont interprétables par un calcul effectué sur l'interprétation de leurs constituants. Ainsi c'est par le fait que ces structures peuvent être interprétées, que les concepts lexicaux représentent quelque chose dans le monde perçu.

En imposant le primat du lien représentationnel, l'atomisme conceptuel s'attaque à l'incapacité de la version radicale du molécularisme à représenter des entités et des situations du monde perçu. C'est le caractère représentationnel qui devient déterminant.

Concepts are the constituents of thoughts; as such, they're the most elementary mental objects that have both causal and representational properties. Since, however, concepts are individuated by their representational *and not* by their causal properties, all that has to be specified in order to identify a concept is what it is the concept of.

(FODOR 1994 [36] p. 112)

Il semble cependant difficile de justifier l'atomisme conceptuel face à la version modérée du molécularisme compositionnel. Cette approche semble, en effet, résoudre l'essentiel des problèmes en permettant le calcul du sens des énoncés tout en assurant le lien représentationnel avec le monde perçu. Paradoxalement, c'est le rejet de cette version du molécularisme qui est la principale justification de l'atomisme conceptuel. L'impossibilité de trouver les bonnes définitions, la difficulté d'imaginer quelles sont les primitives conceptuelles et l'inefficacité explicative du holisme font du molécularisme compositionnel une voie séduisante, mais impraticable (CF. CHAPITRE 7). Si les concepts étaient identiques à une structure, celle-ci devrait être facile à déterminer. Or ce n'est le cas pour aucun concept, en dehors des mathématiques et des sciences formalisées. Les structures qui sont couramment proposées, comme celle de "causer la mort de" pour TUER, apparaissent comme des paraphrases de portée locale, plutôt que comme des descriptions fiables.

Les arguments développés contre le molécularisme sont des arguments qui nous semblent décisifs. Les concepts lexicaux ne sont pas équivalents à des définitions. La richesse de leurs variations en fonction du contexte échappe aux prédictions qu'une définition figée peut produire. Les concepts lexicaux ne sont pas non plus équivalents à des réseaux d'interrelations entre représentations. À défaut de ne présenter qu'une re-formulation de la notion de la définition, une telle conception conduit à une vision holistique du système conceptuel, sans plausibilité cognitive. Ces arguments nous conduisent à considérer avec sérieux l'éventualité selon laquelle les concepts lexicaux seraient des atomes dépourvus de structure conceptuelle interne. Cette position constitue un important retour en arrière sur bien des aspects, notamment en ce qui concerne la prédiction des propriétés causales des concepts, mais également en ce qui concerne leur lien représentationnel avec le monde perçu. Elle semble pourtant inévitable, et il faut donc en explorer les conséquences.

8.2. Le problème de l'acquisition

L'atomisme conceptuel fait le pari d'expliquer la propriété représentationnelle des concepts lexicaux sans utiliser aucune relation que le concept pourrait avoir avec d'autres représentations conceptuelles. L'enjeu est de doter les concepts lexicaux d'un contenu, et non simplement d'une interprétation en termes d'autres concepts. Dans une conception réaliste, il s'agit d'identifier, pour chaque concept, la loi qui le lie à une propriété du monde.

[...] content is constituted by some sort of nomic, mind–world relation. Correspondingly, having a concept (concept possession) is constituted, at least in part, by *being in* some sort of nomic, mind–world relation.

(FODOR 1998 [37] p. 146)

Le caractère intentionnel des concepts étant ainsi mis au premier plan, il s'agit d'expliquer l'origine de leur intentionnalité. Or, la solution de ce problème est loin d'être évidente. Un premier constat concerne le caractère mental des concepts. La représentation conceptuelle ARMOIRE déclenchée par la vue d'une armoire est une entité mentale, susceptible d'être impliquée dans des mécanismes linguistiques ou des mécanismes inférentiels. Ce n'est pas une propriété qui peut être déduite de l'objet armoire ou des atomes qui le constituent. La

représentation conceptuelle ARMOIRE, même si elle peut être déclenchée par des données de la perception, n'est rien d'autre qu'une émanation de notre cerveau. Si le concept ARMOIRE est une représentation interne à notre esprit, comment expliquer qu'elle se retrouve "connectée", à un moment du développement de l'individu, à certains objets du monde perçu ? On comprend que le mot armoire puisse se retrouver associé à certaines perceptions, par exemple s'il est entendu dans les contextes appropriés. En revanche, comment la représentation ARMOIRE, si elle préexiste, peut-elle être associée, soit à une perception, soit à un mot ? Et si la représentation ARMOIRE ne préexiste pas, quelle est son origine en tant qu'entité mentale ?

Le lien entre l'entité mentale ARMOIRE et l'objet armoire, tel qu'il est perçu ou, dans une perspective réaliste, tel qu'il existe dans le monde, ne peut être un lien nécessaire. Il ne peut être qu'informationnel, au sens d'une co-variation fiable entre une propriété perçue ou située dans le monde, et le concept interne au système cognitif (CF. CHAPITRE 3). Cette propriété constitue la cause fiable du déclenchement du concept. Inversement, le concept contient une information à propos de cette propriété. Comment ce lien informationnel objet - concept est-il établi ?

À première vue, on pourrait penser que ce problème ne se présente que pour la théorie atomiste. Si l'on ne s'appuie pas sur les relations inter-conceptuelles, il faut expliquer, indépendamment pour chaque concept, l'origine du lien qu'il entretient avec le monde perçu. À cet égard, l'atomisme conceptuel peut sembler contre-intuitif. Les concepts sensoriels, par exemple, semblent avoir un lien privilégié avec le monde perçu, bien plus que des concepts plus abstraits. C'est d'ailleurs sur ce genre de différence que le molécularisme s'appuie en se servant de la compositionnalité pour ancrer la plupart des concepts. Dans l'approche moléculariste, les concepts primitifs sont inanalysables conceptuellement, mais sont directement reliés au monde perçu. Les autres concepts héritent d'un ancrage par compositionnalité. Le résultat de cette compositionnalité est un lien avec le monde perçu. Si ce lien se révèle informationnel, autrement dit si le concept composé co-varie avec certaines perceptions, c'est parce qu'un mécanisme d'appariement combinatoire a fait correspondre une structure conceptuelle appropriée à la structure de l'objet réel ou perçu (CF. ANNEXE). Ainsi, si le concept PHOTON est compris comme grain de lumière, la représentation conceptuelle moléculaire GRAIN + LUMIÈRE co-variera avec les photons tels qu'ils se manifestent dans certains phénomènes comme l'effet photoélectrique. Les concepts moléculaires héritent, par un mécanisme d'appariement, leur structure de la science, ou de la langue, comme le concept CÉLIBATAIRE qui est analysé comme personne adulte non mariée.

Cette démarche, cependant, peut paraître viciée à la base (FODOR 1998 [37]). La raison invoquée est que l'ancrage des concepts primitifs, dans une théorie moléculariste, est aussi problématique que celui des concepts en général dans une théorie atomiste. Parmi les concepts candidats pour être primitifs, les mieux ancrés sont sans doute les concepts dits sensoriels. Ainsi, le concept ROUGE non seulement ne semble pas nécessiter de structure, mais son lien avec le monde perçu semble donné au départ sans qu'il soit besoin de l'établir par un quelconque apprentissage. Certes, certains aspects du concept ROUGE ne viennent pas du monde perçu et sont dus à la cognition. C'est le cas de l'aspect qualitatif de l'expérience attachée à la représentation ROUGE, qui ne peut se ramener à une description de la lumière en termes de photons (NAGEL 1974 [77] ; JACKSON 1982 [51]). Toutefois, il est facile d'accepter que le lien entre le concept ROUGE et les propriétés du monde perçu qui provoquent son activation présente un caractère arbitraire "pré-câblé". Malheureusement, cette solution ne répond pas à la question de l'ancrage des concepts sensoriels : elle repose sur une confusion entre le concept ROUGE avec la sensation de rouge, le *quale* rouge. La représentation ROUGE, en tant que concept, est susceptible de s'appliquer à des objets perçus. La sensation de rouge peut être déclenchée de manière réflexe par une perception, mais c'est le concept ROUGE qui peut être appliqué à des objets perçus pour produire un jugement sémantique. Le problème de

l'établissement d'un lien entre le concept ROUGE et les objets perçus, ou même au quale rouge, reste entier. En fin de compte, la difficulté rencontrée par la théorie atomiste pour ancrer ses concepts se retrouve avec la même intensité dans une théorie moléculariste, dans la mesure où elle repose sur l'existence de concepts primitifs dont il faut bien expliquer l'ancrage.

Nous pouvons donc envisager une version radicale de la thèse atomiste, selon laquelle non seulement la plupart des concepts lexicaux n'ont pas de structure interne, mais de plus ils ne diffèrent pas par leur mode d'ancrage (FODOR 1998 [37]). Autrement dit, les concepts sensoriels ne sont pas privilégiés à cet égard. Ainsi, les concepts ROUGE et ARMOIRE ne seraient pas différents en ce qui concerne l'établissement du lien informationnel qui les lie aux propriétés du monde perçu. Tous deux sont des entités mentales : ROUGE et ARMOIRE ne se déduisent pas des propriétés physiques d'objets perçus. Tous deux sont des entités intentionnelles : ils peuvent être appliqués à des objets perçus pour produire un jugement sémantique. Étant à la fois mentales et intentionnelles, comment les représentations conceptuelles se retrouvent-elles "connectées" au monde perçu ? Il semble n'y avoir que deux possibilités, toutes deux extrêmes. Soit le lien qui lie les concepts au monde perçu préexiste à toute expérience, ce qui voudrait dire que tous les concepts, avec leurs propriétés intentionnelles, sont innés ; soit, à l'inverse, aucun concept n'est inné.

Le premier choix, malgré son aspect invraisemblable, a déjà été considéré (FODOR 1975 [34]). L'hypothèse de l'innéité de tous les concepts a l'avantage de supprimer le problème de l'explication de leurs propriétés intentionnelles, puisque ces propriétés sont supposées données au départ. Tout ce que fait l'enfant qui entend un nouveau mot est d'associer ce mot au bon concept en utilisant le lien informationnel qui unit ce concept aux données de la perception. À proprement parler, il n'apprend donc pas le concept. Évidemment, il est difficile d'accepter que le cerveau humain soit empli d'une infinité de concepts qu'il est ou sera capable de former. Certains ont pu utiliser la métaphore du système immunitaire, qui possède dès le départ l'ensemble des formes qui lui permettront de faire face à une infinité d'antigènes (CHOMSKY 2000 [18]). Cette métaphore est quelque peu trompeuse. Le système immunitaire utilise un générateur de diversité combinatoire. Les anticorps ont donc une structure dont on peut en principe déduire leurs propriétés, en particulier leur affinité chimique pour un antigène donné. Or, c'est précisément cette structure que l'on souhaite, dans l'approche atomiste, refuser aux concepts lexicaux. Si les concepts lexicaux sont des représentations sans structure, comment expliquer qu'ils aient tous des propriétés intentionnelles différentes et appropriées ? Dans le cas du système immunitaire, la variété et le caractère approprié des affinités chimiques s'expliquent par la structure des chaînes peptidiques et par le mécanisme sélectif qui leur a permis d'exister. Rien de tel n'est disponible dans le cas des concepts lexicaux².

La critique fondamentale que l'on peut adresser à la vision innéiste des concepts est qu'elle ne résout pas le problème qui a suscité sa formulation : que les concepts préexistent à l'expérience de l'individu n'explique pas l'origine de leur intentionnalité. Le fait d'en appeler à l'histoire de l'espèce ne résout rien et ne fait que repousser le problème (FODOR 1998 [37]). On ne comprend toujours pas par quel mécanisme le concept *C* se retrouve associé par un lien informationnel à certains aspects *P* de l'expérience. Il reste à expliquer pour quelle raison

² Certains auteurs ont invoqué le mécanisme sélectif dans le cas des représentations mentales (CHANGEUX & DEHAENE 1989 [13]). La métaphore des réseaux connexionnistes suggère que les concepts puissent être des attracteurs du réseau neuronal, sélectionnés au sein d'un paysage énergétique riche de représentations potentielles dont une partie seulement est exploitée. Certains attracteurs, du fait de leur utilisation, se retrouveraient renforcés. On pourrait voir en eux des concepts préexistants à toute expérience, et dont l'intentionnalité serait elle-même innée. Une telle théorie semble conférer une plausibilité à l'innéité des concepts. Cependant, elle n'échappe pas aux critiques formulées à l'encontre des systèmes associatifs (CF. CHAPITRE 3). De fait, rien, dans un tel modèle, ne distingue les concepts des représentations perceptives.

c'est C et non C' ni C'' qui, dans l'histoire de l'espèce, s'est trouvé connecté à P . Si l'on dispose d'une telle explication, on voit mal pourquoi ne pas l'appliquer à l'ontogenèse et expliquer ainsi que les concepts, avec leurs propriétés intentionnelles, puissent être appris. En résumé, l'innéité des concepts rend leur ancrage incompréhensible, tant pour les concepts simples de l'approche moléculaire que pour l'ensemble des concepts lexicaux dans l'approche atomique.

La position opposée consiste à accepter le caractère acquis de l'intentionnalité de tous les concepts. Une possibilité classiquement évoquée consiste en la formation d'un concept à partir de la perception de groupes d'objets perçus ressemblants, selon un mécanisme général. Le nouveau concept serait ainsi lié à un prototype perceptuel, en restant distinct de celui-ci. On dispose de plusieurs modèles de tels mécanismes, par exemple les algorithmes statistiques de regroupement (CF. ANNEXE). Dans un tel schéma, le concept conserve son statut de représentation distincte des perceptions. Par exemple, un nouveau concept est "recruté" dans un ensemble de concepts potentiels pour représenter le prototype d'un groupe d'objets perçus mal représentés par les concepts existants. On obtient bien, de cette manière, un nouveau concept atomique dont les propriétés intentionnelles, quoique acquises, sont bien établies. Une manière de décrire ce type de mécanisme pour l'apprentissage des espèces naturelles par les enfants utilise la métaphore de "syndrome" (MARGOLIS 1998 [68]). L'enfant disposerait d'une disposition innée pour remarquer les caractères saillants de l'objet perçu. Il disposerait également de la prédisposition lui faisant attribuer la présence de cet ensemble de caractères saillants à l'essence de l'objet perçu, autrement dit à un concept sous-jacent, dont l'ensemble de caractères observé est une manifestation. La métaphore de syndrome est utilisée dans le sens où l'enfant mémorise les caractères saillants, mais les prend pour les résultats d'une cause qui se trouve être la propriété que possède l'objet perçu. Le concept qui se forme est lié à cette propriété essentielle cachée, et non à l'ensemble des caractères rassemblés dans le syndrome. Comme nous allons le voir, cette manière d'envisager l'origine des concepts lexicaux, si elle semble résoudre le problème de leur ancrage, ne permet pas d'avancer dans la compréhension de leurs propriétés symboliques, notamment en ce qui concerne la systématisme de leur rôle dans le langage et le raisonnement.

La question de l'acquisition des concepts atomiques révèle une difficulté qui concerne plus généralement l'ancrage de toutes les représentations qui sont manipulées symboliquement. La solution qui consiste à considérer que les symboles manipulables sont donnés au départ, munis de leur ancrage, est une position difficilement défendable, dans la mesure où cet ancrage semble miraculeux. La solution opposée, selon laquelle des concepts recrutés parmi un ensemble de concepts potentiels acquièrent leur intentionnalité par apprentissage, n'explique pas comment les concepts lexicaux obtenus se trouveront posséder les bonnes propriétés compositionnelles et inférentielles.

8.3. Les représentations et les règles

L'intérêt épistémologique de doter les concepts d'une structure endogène était d'unifier les mécanismes intra-conceptuels et inter-conceptuels. L'approche atomiste - compositionnelle (AC) renonce à considérer l'existence des premiers. Les mécanismes inter-conceptuels, en revanche, continuent dans AC à reposer sur l'existence de représentations structurées associées aux constructions du langage. C'est là le seul moyen de garantir la systématisme des significations, le fait que quiconque pense l'énoncé Jean aime Marie soit systématiquement en mesure de former la pensée associée à l'énoncé Marie aime Jean. Les représentations issues de l'interprétation des syntagmes ou des phrases restent donc moléculaires. Tout réarrangement licite de leurs atomes produit une pensée acceptable. Quelle

est la nature de ces représentations composées allouées par AC aux constructions du langage, si elle diffère radicalement de celle des concepts lexicaux ?

Le fait d'attacher des représentations conceptuelles différentes aux mots du lexique et aux syntagmes peut sembler contre-intuitif. Par exemple, la représentation conceptuelle d'un syntagme comme *employé de maison* est supposée qualitativement différente du concept lexical associé au mot *domestique*. Or, les représentations correspondantes ne diffèrent pas radicalement par les entités qui les évoquent où qu'elles évoquent, si bien qu'il est difficile d'accepter qu'elles relèvent de mécanismes mentaux d'acquisition et de manipulation qui n'ont rien à voir entre eux. De ce point de vue, l'approche moléculaireiste - compositionnelle (MC) est plus cohérente : les deux représentations y prennent sensiblement la même forme, celle d'une description qui donne les conditions d'évocation, les contraintes grammaticales et les inférences obligatoires. Dans l'approche atomiste, la similarité des concepts et de leurs propriétés causales reste inexplicée. Le problème est encore plus net quand on considère les différences entre langues. Par exemple, le mot *serein* se traduit en anglais par l'expression *evening dew*, si bien qu'un anglophone monolingue n'est pas supposé posséder le concept *SEREIN*, même s'il est ornithologue. Il est censé former une représentation complexe là où le francophone manipule un concept atomique. Inversement, l'expression *pomme de terre* renvoie pour l'un à une représentation complexe, là où le mot *potato* évoque un concept simple pour l'autre. On pourrait penser ce dernier exemple mal choisi, tant l'expression *pomme de terre* semble se comporter comme une expression figée, autrement dit comme un lexème qui devrait, selon AC, évoquer un concept simple. Or, la théorie atomique est bien en peine d'expliquer le passage d'une représentation conceptuelle complexe à un concept simple sous l'effet de l'habitude. Les problèmes théoriques liés à l'acquisition des concepts simples seraient augmentés s'il fallait imaginer un mécanisme de conversion qui rendrait atomiques des représentations qui ne le sont pas, tout en conservant leurs propriétés intentionnelles. Cette difficulté s'étend à la compréhension des descriptions. Nous pouvons comprendre une description verbale comme une *fenêtre en forme de losange* et former à partir d'elle une représentation dotée d'une intentionnalité. Dans MC, ce processus est naturel, puisque tous les concepts, hormis les quelques concepts simples, reçoivent leur intentionnalité de leur structure. L'approche AC, en revanche, est obligée de poser une différence qualitative entre l'intentionnalité produite à partir d'une description et celle d'une entité lexicale. Comment ces deux intentionnalités se rencontrent-elles pour permettre au sujet de comprendre, à l'aide de sa description, le sens d'un nouveau terme ?

La présence de deux types de représentations, simples et complexes, impose de faire cohabiter deux types de mécanismes sémantiques. On pourrait imaginer conserver dans AC certains mécanismes de MC pour opérer sur les représentations composées associées aux syntagmes. Cependant, le fait de considérer les concepts lexicaux comme atomiques oblige à concevoir un calcul compositionnel d'un nouveau type. En d'autres termes, ce calcul engendrerait une "colle conceptuelle" pour produire des représentations composées à partir des concepts atomiques (MARGOLIS 1999 [69]). Or, un tel calcul, étant supposé symbolique, opère sur des concepts d'après leur propriété causale. Si les concepts lexicaux sont dépourvus de structure, et donc de forme, accessible aux opérateurs, il devient plus difficile d'expliquer la systématisme de ce calcul. La systématisme des phénomènes sémantiques se traduit par le fait que certaines possibilités de composition sont closes pour les substitutions au sein des classes sémantiques. Ainsi, la substitution du mot *Jacques* au mot *Jean* dans la phrase *Jean aime Marie* ne doit pas changer l'acceptabilité sémantique de la phrase. Dans MC, le lien obligatoire entre l'acceptabilité de la phrase *Jean aime Marie* et celle de la phrase *Jacques aime Marie* peut être assuré par le fait que la structure du concept *AIMER* s'accorde systématiquement avec ce qu'il y a de commun aux concepts *JEAN* et *JACQUES*. Ce type d'explication n'existe plus avec AC. Nous avons vu que, dans MC, la compositionnalité est expliquée par le fait qu'un

ensemble de mots m_1, m_2, \dots, m_k , structuré syntaxiquement, est interprété par composition des concepts m_i , associés aux mots m_i . La composition des significations s'obtient par un calcul, symbolisé par la fonction $\mathbf{f}()$, qui peut s'écrire $\mathbf{f}(s_1, s_2, \dots, s_k) = \mathbf{g}_{s_1, s_2, \dots, s_{k-1}}(s_k)$ et ceci jusqu'à la fonction \mathbf{g}_{s_1} qui, formée à partir de la structure moléculaire du concept m_1 , fournit la structure qui sera à la base de celle de l'énoncé. Cette idée conduit naturellement au principe de constituance, selon lequel la structure du concept AIMER existe dans la représentation conceptuelle associée à la phrase Jean aime Marie. Le calcul symbolisé par $\mathbf{f}()$ se résume alors à une agrégation de structures. C'est ainsi que MC parvient à expliquer la systématisme des substitutions acceptables. Or, l'approche AC considérant que le concept AIMER est dépourvu de structure conceptuelle, il faut imaginer une tout autre explication de l'acceptabilité sémantique.

Les concepts lexicaux de AC n'ayant pas de structure, ils ne peuvent héberger la connaissance sémantique qui les concerne. Celle-ci doit donc être stockée de manière séparée, dans des règles. Dans l'approche MC, les règles peuvent servir à représenter des connaissances contingentes, dans la mesure où l'on introduit une différence entre ces connaissances contingentes et des connaissances qui, elles, seraient définitoires et indissociables du concept. Par exemple, la description sémantique du concept INFLATION peut être limitée à la notion de hausse des prix. Le fait que l'inflation entraîne une dépréciation de la monnaie, un renchérissement des importations et, à terme, une hausse des salaires, peut être relégué dans une base de connaissances contingentes. L'approche AC n'opère pas de distinction formelle entre connaissances définitoires et connaissances contingentes. L'ensemble des connaissances est stocké sous forme de règles. Ainsi, le fait que l'inflation se traduit par une hausse des prix sera inscrit dans une règle, de même que le fait que le concept TUER décrit un événement, que cet événement en cause un autre, que le complément d'objet doit être animé, qu'il devient inanimé à l'issue de l'événement, *et cætera*. Les bases de règles sont couramment utilisées en intelligence artificielle pour la modélisation du raisonnement. Dans leur version classique, ces règles combinent des représentations atomiques, propositions ou prédicats par exemple, à l'aide de connecteurs logiques. En utilisant des bases de règles, l'approche AC peut ainsi prétendre représenter la connaissance sémantique d'une manière beaucoup plus simple que MC en évitant tout usage de structures complexes et de mécanismes pour les apparier. La mise en œuvre de cette connaissance repose cependant sur l'utilisation d'un moteur inférentiel capable de gérer une base de connaissances de taille considérable.

Il n'y a pas de limite théorique à la quantité de connaissances que l'on peut stocker à l'aide d'une base de règles. En pratique, cependant, les tailles requises pour représenter ainsi la connaissance sémantique posent des problèmes insurmontables. En particulier, les tests permettant de savoir si la base de connaissances est logiquement cohérente croissent exponentiellement avec sa taille, ce qui les rend inutilisables. Or, une base de connaissances incohérente produit des résultats erronés. De même, la taille prohibitive de la base de règles nécessaire pour représenter la connaissance d'un être humain rend son exploitation difficile à envisager. L'évocation d'un concept plongerait le système inférentiel dans une recherche longue et fastidieuse de toutes les règles qui peuvent s'appliquer, puis dans la recherche d'autres règles qui peuvent être déclenchées par les premières. Même en supposant que l'on borne artificiellement ce processus déductif, l'audition d'une simple phrase doit produire un nombre considérable d'opérations mentales visant à trier les règles qui peuvent s'appliquer à chaque pas³. Or, aucune donnée psychologique ne vient étayer ce type de scénario.

³ On peut par exemple invoquer un principe d'économie pour limiter le nombre des déductions effectuées. L'idée est que même les premières déductions, dans le modèle de la base de règle, requièrent un nombre prohibitif d'opérations.

Mentionnons enfin que l'existence d'une base de règle suppose que ces règles soient apprises et mises à jour. Or, toute modification de la base de règle se heurte au problème du cadre, en risquant de nécessiter des mises à jour qui, de proche en proche, peuvent affecter l'ensemble de la base.

Un autre problème concernant l'utilisation de bases de règles pour représenter les connaissances sémantiques provient du caractère contingent des règles. On accepte volontiers qu'un individu puisse ignorer certaines connaissances que l'on jugera contingentes, comme l'effet de l'inflation sur les importations. En revanche, il serait étonnant qu'un individu normal ignore qu'un être tué est mort ou qu'étant mort, il le reste. Ces connaissances, que l'on qualifie de sens commun, sont mal représentées par des règles⁴. Or, en traitant de la même manière les connaissances sémantiques et les connaissances contingentes, l'approche AC n'explique pas pourquoi les connaissances de sens commun s'imposent à notre entendement. Dans MC, ces connaissances sont indissociables de la compréhension du concept. Avec AC, le concept existe par ses seules conditions d'évocation, si bien que toutes les connaissances qui lui sont attachées le sont de manière contingente. On devrait donc trouver des sujets capables de reconnaître correctement des situations dans lesquelles un être en tue un autre, sans que ces sujets soient capables d'en tirer les conséquences pour la vie de l'être occis. Le fait que le sens commun s'impose à tous exclut que la connaissance correspondante provienne d'un ensemble contingent de règles.

8.4. La monotonie de la composition

Le fait de considérer les concepts lexicaux comme atomiques, alors que les syntagmes sont, eux, associés à des représentations composées, semble conduire inévitablement à une contradiction. Dans une phrase, les syntagmes peuvent être substitués à des mots, et réciproquement. Ainsi, dans la phrase le chien de Charles est malade, on peut remplacer le syntagme le chien de Charles par le nom propre Lee. Il doit en être de même pour les représentations mentales correspondantes. Or, selon AC, les deux représentations substituées ne sont pas de même nature, l'une étant moléculaire et l'autre atomique. En conséquence, les pensées résultant de l'interprétation de la phrase le chien de Charles est malade et celle de Lee est malade ne peuvent pas être identiques. Notons que malgré les difficultés que cette distinction ne manque pas d'entraîner pour établir le lien entre les deux significations, elle a le mérite d'expliquer que certains sujets puissent considérer de bonne foi les deux syntagmes comme non synonymes.

Malheureusement, si l'on s'accommode de cette distinction, il faut accepter que les représentations provoquées par l'interprétation d'énoncés complets, par exemple la présente phrase, ont la même complexité structurelle que celle que l'on observe au niveau syntaxique. Or ceci est difficile à accepter. Noter que la reprise anaphorique par le pronom ceci, dans la phrase précédente, est censée désigner la pensée exprimée par la première phrase du paragraphe. Si l'on enchaîne des constructions syntaxiques et des reprises anaphoriques, comme cela s'observe par exemple dans les narrations, on arrive à des représentations de complexité arbitrairement grande. Or ces représentations sont censées être des objets statiques de notre entendement, susceptibles d'entrer dans la composition d'autres représentations. Un modèle qui postule l'existence de tels objets cognitifs n'est pas parcimonieux. Il est, de plus,

⁴ La recherche en intelligence artificielle a tenté d'enfermer les connaissances de sens commun dans un ensemble de règles. Le projet CYC est une tentative de grande envergure qui s'est donné cet objectif comme prioritaire (LENAT & GUHA 1990 [65]). Cependant, les succès qui ont pu être atteints, même s'ils restent intéressants d'un point de vue pratique et technique, restent dérisoires en comparaison de l'étendue des connaissances de sens commun que possède n'importe quel être humain.

démenti par l'expérience. On s'attend à ce que l'interprétation des enchaînements discursifs soit de plus en plus lente, puisqu'elle engage la manipulation de pensées dont la complexité va croissant. Or rien de tel n'est observé.

Le discours humain comporte de larges portions cohérentes. C'est le cas, nous l'espérons, du présent texte. La théorie AC n'offre, pour la construction du sens, qu'un processus monotone, strictement croissant en fonction de l'avancement du discours. Comme les parties d'un discours cohérent font référence à d'autres parties, le résultat produit est une gigantesque structure arborescente sans plausibilité cognitive. Certes, on sait intuitivement que l'on ne mémorise pas la totalité d'un discours. Les reprises anaphoriques pourraient, par exemple, faire référence à des versions partielles de ce qui a été exprimé précédemment. Malheureusement, AC n'offre aucun moyen de simplifier les structures arborescentes qu'elle engendre. Un simple mécanisme d'oubli risque de produire des structures aberrantes ou d'abandonner des portions essentielles de la structure. Ce qui manque, c'est un processus de réduction qui permette de remplacer des structures complexes par des structures plus simples. En dépit de ce que suggère l'intuition, aucun mécanisme de AC ne permet de diminuer la complexité des structures construites : la représentation conceptuelle associée au syntagme le chien de Charles n'est jamais remplacée par le concept LEE.

À titre de comparaison, le décodage syntaxique ne présente pas le même défaut de monotonie. Certes, le décodage syntaxique d'une phrase produit des structures arborescentes de taille non négligeable. L'imbrication des syntagmes fait qu'il n'est pas rare qu'on soit obligé, au milieu d'une phrase, de mémoriser une arborescence de quatre ou cinq niveaux. Cependant, la structure interne d'un syntagme peut être oubliée dès que le syntagme a été reconnu. Ainsi, il suffit de mémoriser la partie de l'arbre syntaxique qui se trouve au-dessus du point où l'on se trouve, ce qui limite singulièrement la charge cognitive. En particulier, la mémoire syntaxique est remise à zéro lors du changement de phrase. Le décodage sémantique que nous offre AC ne présente pas cette bonne propriété.

Ce paradoxe de la monotonie existait dans MC, quoique de manière moins flagrante. Dans MC, un mécanisme comme l'unification des structures lors de la composition évite à ces structures de croître trop rapidement. Ainsi, une phrase peut recevoir une structure dont la complexité peut ne pas excéder de beaucoup celle de son verbe. Il n'en reste pas moins que MC, comme AC, n'offre aucun moyen de diminuer la complexité des structures sémantiques engendrées. Celle-ci ne peut donc qu'augmenter jusqu'à des niveaux qui dépassent de beaucoup ce qui est cognitivement plausible.

Conclusion

Le problème principal auquel l'approche atomiste se trouve confronté est celui de l'acquisition des concepts. Ce problème pose, avec la même intensité pour les concepts simples, comme les concepts associés aux sensations de base, de l'approche moléculariste. Le postulat d'un caractère inné pour tous les concepts étant écarté, la seule option est de supposer que les concepts ne sont qu'un ensemble de fonctions binaires recrutées pour représenter les données de notre perception.

The model, to repeat, is *being red* : all that's required for us to get locked to *redness* is that red things should reliably seem to us as they do, in fact, reliably seem to the visually unimpaired. Correspondingly, all that needs to be innate for RED to be acquired is whatever the mechanisms are that determine that red things strike us as they do; which is to say that all that needs to be innate is the sensorium. Ditto, *mutatis mutandis*, for DOORKNOB if *being a doorknob* is like *being red*: what has to be innately given to get us locked to *doorknobhood* is whatever mechanisms required for doorknobs to come to strike us as such. [...] the kind of nativism about DOORKNOB that an informational atomist has to put up with it perhaps not one of *concepts* but of *mechanisms*.

(FODOR 1998 [37] p. 142)

Si on peut imaginer que les concepts atomiques, en tant que des entités intentionnelles, peuvent être acquis, par exemple par un mécanisme de recrutement, on peut également imaginer que les concepts sont construits, de manière éphémère, pour les besoins de l'interprétation des énoncés. Nous retenons donc l'idée que ce sont les mécanismes d'élaboration des concepts, non les concepts eux-mêmes, qui jouent un rôle essentiel dans l'élaboration du sens. Ainsi, le fait que les concepts soient des représentations atomiques ne pose plus de problème, ni pour l'apprentissage, ni pour l'explication de la systématique. Nous étendrons cette idée au-delà des concepts lexicaux, jusqu'à considérer que les représentations produites par l'interprétation des constructions linguistiques sont, elles-mêmes, atomiques. De cette manière, le processus de construction du sens cesse d'être cumulatif et ne produit plus de représentations de complexité invraisemblable. Cet atomisme radical sera rendu possible par le fait que les concepts sont considérés comme des entités transitoires, élaborés pour les besoins de la compréhension de l'énoncé.

L'approche que nous proposons, qui est celle d'un atomisme sémantique intégral, semble nous obliger à renoncer à la systématique des constructions sémantiques. Si les pensées sont, comme les concepts, atomiques, comment garantir la production de toutes les représentations possibles, de manière licite, à partir des mêmes ingrédients ? La réponse consistera à placer la systématique dans la procédure de construction du sens, non dans son résultat.

Conclusion

Au point où nous sommes parvenus, il semble que tant l'approche moléculariste que l'approche atomiste échouent à expliquer la compositionnalité. Pourtant, il semblait au départ que l'on fût obligé d'opter pour l'une ou l'autre des deux approches. À partir du moment où l'on considère que les principes de compositionnalité et de constituance sont seuls à même d'expliquer les aspects systématiques de la construction du sens, l'hypothèse du mentalais semble incontournable. Le calcul du sens d'une combinaison grammaticale et la production des inférences sont ainsi ramenés à des processus symboliques opérant sur un lexique mental. Or, lorsque l'on cherche à préciser la nature des éléments de ce lexique, on tombe, comme nous l'avons vu, sur une série de contradictions.

La situation ressemble à une impasse. Renoncer à l'idée d'un calcul symbolique revient à renoncer à expliquer la productivité et la systématisme, qui semblent être indiscutablement des propriétés de notre compétence sémantique. Accepter un calcul symbolique portant sur les significations exige que ces significations, en tant que symboles, aient une structure récursive ou soient atomiques. Or, aucune de ces deux options ne mène à une solution cohérente.

Une hypothèse communément envisagée consiste à imaginer que la construction du sens n'a rien de symbolique. Par exemple, la métaphore du connexionnisme a conduit bon nombre d'auteurs à considérer que les représentations associées au langage pouvaient être activées par un réseau d'associations (SMOLENSKY 1988 [97] ; SMOLENSKY 1990 [98]). La critique adressée à ce type de modèle sub-symbolique associatif est qu'il n'est ni productif, ni systématique (FODOR & PYLYSHYN 1988 [41] ; FODOR & MCLAUGHLIN 1990 [40]). Il n'en reste pas moins que les mécanismes symboliques ne peuvent prétendre, à eux seuls, modéliser l'ensemble des processus sémantiques. De nombreux phénomènes, omniprésents dans le langage, semblent hors de portée des modèles combinatoires (LAKOFF 1990 [63]). Par exemple, les emplois métaphoriques des mots sont mal modélisés par les procédures symboliques. La frontière entre l'emploi métaphorique d'un mot donné et un emploi qui pourrait être qualifié de littéral est même, dans la plupart des cas, floue. Si l'on peut imaginer des procédures non symboliques, par exemple des sortes de fusions de significations, pour modéliser les métaphores, pourquoi ne pas considérer que l'hypothèse du calcul symbolique est un échec et que la totalité du calcul du sens est dû à des procédures d'un autre type, même si elles ne sont pas encore modélisées ?

Le problème est que le calcul sémantique doit assurer deux interfaces avec des systèmes symboliques : le système syntaxique et le système pragmatique. Le système syntaxique, selon la plupart des modélisations, manipule des structures symboliques arborescentes selon des procédures récursives. Le système pragmatique met en œuvre des procédures de raisonnement, elles aussi récursives, portant sur des représentations symboliques qui peuvent être évaluées et niées. Il n'est donc pas possible de renoncer aussi facilement au caractère symbolique des représentations sémantiques, sous peine de perdre tout espoir de pouvoir modéliser ces deux interfaces.

Nous devons souligner que bon nombre de problèmes liés au mentalais sont dus au fait qu'il s'agit d'un catalogue de représentations statiques. Le mentalais est, avant tout autre chose, un lexique mental. Dans le cas de l'approche moléculariste, il s'agit même d'un dictionnaire mental dans lequel chaque entrée correspond à une description. Le caractère statique de ces représentations oblige à une précision absolue. Les concepts ne peuvent pas être ambigus. Les descriptions de l'approche moléculariste - compositionnelle doivent être parfaites et les conditions d'évocation, dans l'approche atomiste - compositionnelle, doivent être telles qu'elles seront correctes dans tous les contextes. Or, de telles idéalizations ne

constituent pas un cas limite. Elles dénotent une vision naïve et erronée du fonctionnement du langage. Les mots ne sont pas ambigus du fait de l'incapacité des humains à former un lexique précis. Au contraire, c'est parce que les mots sont ambigus que la communication peut fonctionner. Le langage ne dispose que de quelques dizaines de milliers de mots et de la possibilité de former un ensemble dénombrable de combinaisons. Or, le nombre de signifiés potentiels est incommensurablement plus grand. Il ne faut pas en conclure que le langage est intrinsèquement flou. Les interlocuteurs obtiennent presque toujours la précision requise pour que la communication soit un succès. Ils parviennent sans difficulté, par exemple, à se désigner un objet parmi d'autres à l'aide de mots souvent inappropriés. Ils parviennent, de même, à décrire des scènes absentes d'une manière suffisamment précise pour que les effets pragmatiques soient atteints. Un tel résultat serait impossible à obtenir avec un langage mental fixe. Les occasions dans lesquelles les entités à décrire coïncideraient avec ces représentations fixes seraient trop rares, et il en résulterait presque toujours des erreurs irréductibles. On ne peut appréhender l'infini continu des signifiés avec un langage mental dont le lexique serait connecté de manière fixe au monde perçu. Ce qui fait la force des langues, c'est précisément que leur lexique est ambigu. En refusant cette ambiguïté au mentalais, on le rend impropre à exprimer le sens. L'hypothèse du mentalais traite la construction du sens comme une traduction, d'une langue naturelle dans une pseudo-langue mentale. C'est ce qui lui confère sa force apparente. Mais les limites d'un tel modèle ne tardent pas à se révéler, comme nous l'avons constaté dans cette partie.

Le mentalais comporte un lexique, l'ensemble des concepts, mais il produit également des "phrases", les représentations associées aux constructions du langage. Dans cette conception de la compositionnalité, la représentation mentale créée par l'interprétation d'un agencement de mots est censée porter la trace, dans sa structure, de son processus d'élaboration. L'objet contient toute la complexité, alors que la procédure qui l'élabore se contente d'ajouter des éléments à l'édifice. Nous proposerons, à l'inverse, de considérer que la complexité réside dans la procédure de construction du sens, non dans la structure de la signification résultante. Il ne s'agit pas de renoncer au caractère symbolique des représentations sémantiques et des mécanismes qui les manipulent. En revanche, nous remettons en question le caractère statique des représentations conceptuelles. Autrement dit, nous pouvons nier l'existence du mentalais sans renoncer pour autant à un traitement symbolique du sens. L'hypothèse que nous défendrons dans la prochaine partie consiste à dire que les concepts, en tant que représentations stables de notre entendement, sont une illusion. Entre d'une part les partisans d'un mentalais, et d'autre part les défenseurs d'une solution entièrement sub-symbolique, dans laquelle il n'y a pas de place pour des mécanismes symboliques, nous allons opter pour une troisième voie. Les concepts, dans notre modèle, ont bien une réalité cognitive, mais seulement d'une manière éphémère. Leur existence dépend du contexte dans lequel le sens d'une phrase est élaboré, et elle ne dure pas plus longtemps. En renonçant à toute idée de catalogue préétabli de significations, nous voulons faire de la construction du sens une authentique procédure dont tous les produits, y compris les concepts eux-mêmes, sont éphémères. Cette manière d'aborder la construction du sens s'accorde bien avec la possibilité d'atteindre une certaine précision dans la communication. Cette précision ne réside pas dans la finesse de la structure construite, mais dans le détail de la procédure de construction. Ainsi, tout ce que nous pouvons faire pour amener un individu à avoir une pensée analogue à la nôtre consiste à guider sa procédure d'élaboration du sens d'une manière que nous jugeons suffisante pour la précision requise, sans garantie de succès. Les erreurs éventuelles sont révélées par la suite de l'interaction, lorsque l'individu ne produit pas les inférences souhaitées.

Partie 4 :
**Procédure de construction du
sens**

Introduction

Dans cette quatrième partie, nous allons mettre en question une hypothèse fondamentale qui a servi de base aux développements des deuxième et troisième parties. Jusqu'ici, nous avons supposé qu'il existait un système conceptuel indépendant, hébergeant des représentations particulières, les concepts, et doté de mécanismes cognitifs propres. Le cahier des charges, pour un tel système, comprend trois interfaces. Or, nous avons observé en détail les difficultés considérables qui se posent lorsque l'on cherche simultanément à ancrer les concepts dans la perception, à les doter des moyens de déclencher des inférences, et à les combiner en suivant les instructions données par la syntaxe de la langue.

Les solutions que nous avons envisagées de la manière la plus détaillée étaient de nature purement symbolique : les concepts peuvent être conçus comme des structures récursives ou comme des atomes de sens impliqués dans des règles. Nous avons amplement discuté les difficultés internes de ces solutions purement symboliques, ce que nous avons appelé le dilemme du mentalais. Cependant, la difficulté principale des solutions purement symboliques est qu'elles rendent très mal compte de l'interface avec la perception : on ne sait ancrer les symboles de manière satisfaisante.

Dans cette partie, nous allons revenir sur une solution évoquée brièvement à propos de l'ancrage. Nous avons considéré la solution empiriste qui consiste à ne pas poser de différence de nature entre concepts et percepts. Évidemment, on évite de cette manière d'avoir à traiter la question de l'ancrage des concepts dans la perception, puisque les concepts sont des représentations perceptuelles plus ou moins généralisées. Le prix à payer est l'impossibilité de rendre compte des phénomènes de systématicité que les solutions symboliques résolvent si naturellement.

Ce que nous proposons ici rejoint la solution empiriste sur un point. La plupart des mots du lexique sont supposés renvoyer à des représentations qualitatives, de même nature cognitive que les représentations perceptuelles. En revanche, nous conservons l'objectif d'expliquer la systématicité par l'intervention de représentations symboliques. La grande différence est que ces représentations symboliques ne sont pas supposées permanentes. En d'autres termes, nous allons dénier au système conceptuel le statut de système autonome, doté de représentations stockées à long terme. Le système conceptuel devient une simple interface entre la perception, le langage et le raisonnement. Cette interface fonctionne en élaborant des représentations symboliques éphémères en fonction du contexte.

Chapitre 9 :
Vers un modèle de
représentations conceptuelles
éphémères

Introduction

Dans la deuxième et la troisième partie de ce travail, nous avons, pour l'essentiel, considéré les modèles symboliques du système conceptuel. Leur avantage est de garantir le caractère systématique du calcul sémantique. En les critiquant, nous avons montré que ces avantages résultaient du fait qu'au lieu de proposer un véritable calcul du sens, ces modèles effectuent une simple traduction, un transcodage d'un système dans un autre. Nous avons souligné par ailleurs leurs défauts, que nous avons présentés comme rédhibitoires. Nous pouvons les résumer ici en disant que les entités symboliques postulées au niveau du système conceptuel sont trop rigides et ne peuvent pas être apprises. En particulier, les descriptions imaginées pour caractériser les concepts lexicaux ne peuvent pas s'adapter à la variété des contextes et elles comportent des primitives dont l'ancrage au cours de l'ontogenèse reste inexplicable.

Dans ce chapitre, notre objectif est de jeter les bases d'une solution qui échappe aux critiques des systèmes conceptuels vus comme des catalogues d'entités symboliques permanentes, sans renoncer à rendre compte de la systématisme qui caractérise certains phénomènes de sens. Le système conceptuel sera présenté comme un mécanisme d'interface dont les représentations sont construites "à la volée".

9.1. Concepts *versus* percepts

L'un des arguments les plus forts en défaveur d'un système conceptuel autonome hébergeant des représentations symboliques spécifiques vient de l'omnipotence de la faculté de conceptualisation. Nous pouvons nommer tout ce que nous pouvons discerner consciemment. Les contractions stomacales n'étant pas conscientes, il n'est pas étonnant qu'elles n'aient pas été conceptualisées avant que la science révèle leur existence. En revanche, tout être humain peut verbaliser les nuances et les contrastes de ce qu'il perçoit et de ce qu'il ressent¹. Il ne s'agit pas de dire qu'il réussit à communiquer efficacement les concepts correspondants, mais qu'il peut tenter de le faire. Au moment de parler, il doit avoir à sa disposition, sinon les mots adéquats, au moins les concepts qui lui permettent d'exprimer les discriminations qu'il a effectuées. Or, ce constat n'est pas trivial. On pourrait imaginer un champ de l'expérience consciente qui correspondrait à une cécité conceptuelle. Si tel était le cas, il existerait une impossibilité de principe rendant ces événements conscients strictement inaccessibles à la conceptualisation, et donc à la verbalisation. Cette situation, qui semble absurde, est pourtant celle que prédit un système conceptuel autonome. Si un tel système existe, comment garantir qu'il contient l'ensemble des concepts requis pour permettre de parler à tout moment de toutes les dimensions de l'expérience consciente ? Les concepteurs de tels modèles n'offrent aucune justification expliquant la présence des concepts adéquats ou l'existence d'un mécanisme capable de les acquérir au moment où leur emploi est requis. Même si l'on supposait que ce problème pût être contourné, il en résulterait une duplication invraisemblable de l'ensemble de l'information perceptuelle et sensorielle qui devrait être recopiée dans le système conceptuel de manière à être disponible pour la verbalisation.

Pour éviter ce genre d'invraisemblance, et échapper en même temps à l'ensemble des paradoxes liés au postulat de représentations conceptuelles symboliques permanentes, une

¹ Le cas particulier des odeurs, qui semblent plus difficiles à décrire et à opposer verbalement, ne sera pas commenté ici. L'existence de cette difficulté démontre que les questions qui suivent ne sont pas triviales.

solution est d'adopter une vision dans laquelle les concepts ne sont pas différents des données de l'expérience consciente. Un tel point de vue rejoint les modèles empiristes de la cognition, que nous avons brièvement considérés (CF. CHAPITRE 3). Non seulement le problème de l'omnipotence de la faculté de conceptualisation ne se pose plus, mais celui de l'ancrage des concepts dans la perception devient lui aussi sans objet, puisque concepts et percepts peuvent être identifiés. Cette solution serait retenue par tous les auteurs si elle ne présentait, elle aussi, son lot d'inconvénients apparemment insurmontables. Ces inconvénients sont liés à l'impossibilité apparente d'établir des interfaces systématiques avec le langage et le raisonnement. Les catégories perceptuelles sont par nature graduelles, et les opérations auxquelles elles se prêtent naturellement ne se laissent pas guider par la syntaxe du langage ou les opérations logiques du raisonnement.

Les représentations de type perceptuel peuvent être associées entre elles par leur contenu, et elles peuvent subir des transformations complexes. Par exemple, les représentations visuelles peuvent être parcourues mentalement ou subir des rotations mentales (KOSSLYN 1994 [61]). Elles peuvent être également synthétisées à partir de leurs composants (KOSSLYN 1994 [61]). Cependant, de telles opérations s'interfacent mal avec les relations données par la syntaxe et le raisonnement. Par exemple, si les syntagmes chiens de ferme et ferme de chiens évoquent des représentations perceptuelles, celles-ci seront différentes : peut-être, dans un cas, un chien attaché dans une cour de ferme typique, et dans l'autre une sorte de chenil avec de nombreux chiens en cage. Cette différence semble imposée par la syntaxe : dans un cas le mot ferme est le modifieur, dans l'autre c'est le modifié. Le problème, pour un défenseur de l'identité concept - percept, est d'expliquer comment cette relation syntaxique peut guider la génération d'image au point d'avoir une telle influence. Pour illustrer la difficulté de concevoir une interface avec le raisonnement, il suffit de mentionner le cas de la négation. Comment définir, en termes de percepts, la signification de l'expression non chien ? Est-ce l'image d'un chat, d'un lion, d'une table ou d'un tourbillon ? Les images, ou les percepts en général, n'admettent pas de négation. Ces problèmes d'interfaces n'ont rien d'anecdotique. On attend de telles interfaces qu'elles soient systématiques : tout concept peut être nié dans un contexte approprié. Ainsi, le syntagme non chien peut s'appliquer à un chat ou à un perroquet, mais pas à un lion ou à une table, dans un contexte où quelqu'un cherche à acquérir un animal domestique qui ne soit pas un chien. Or, aucune des transformations habituellement définies dans le domaine perceptuel ou imagé ne permet d'obtenir un tel résultat.

Malgré les difficultés qui viennent d'être mentionnées, nous allons néanmoins partir de la représentation perceptuelle pour fonder notre modèle de la représentation du sens des énoncés langagiers. En d'autres termes, nous renonçons aux concepts en tant qu'entités symboliques permanentes, considérant que les paradoxes mentionnés dans les chapitres précédents sont insurmontables. Pour que notre choix soit crédible, nous devons indiquer, au moins dans le principe, comment un domaine perceptuel peut être interfacé de manière systématique avec le langage et le raisonnement.

9.2. Espace qualitatif

Partant de l'omnipotence du langage, nous allons supposer que la plupart des mots peuvent évoquer des "scènes" de même nature que celles dont l'individu peut avoir l'expérience. Autrement dit, il n'est pas question de postuler l'existence d'un médium de représentation parallèle à celui qui héberge les scènes que nous percevons, celles dont nous avons gardé un souvenir, et celles dont nous avons la capacité d'imaginer. En particulier, un mot de la classe ouverte du lexique ne renvoie en aucun cas vers une structure conceptuelle de

type symbolique². La plupart des mots évoquent une scène dont la nature est la même que celle d'une perception remémorée. Hors contexte, l'évocation peut être faible et ne produire que des scènes typiques. Ainsi, le mot pomme peut évoquer l'image d'une pomme standard ; le mot jalousie peut renvoyer à une scène typique à trois personnages liés par des relations amoureuses. La présence d'un contexte rend les évocations beaucoup plus vives : on se remémore visuellement la pomme qu'on a laissée sur le buffet ; on repense à une pièce de théâtre que l'on vient de voir où, par empathie, on a senti la jalousie de l'un des personnages.

La représentation des scènes peut reposer sur les espaces qualitatifs comme ceux que la psychologie a mis en évidence pour rendre compte du fonctionnement de la perception et de la mémoire. Par exemple, on peut montrer que les propriétés spectrales du signal lumineux sur une surface uniforme, telles qu'elles sont ressenties, se répartissent dans un espace à trois dimensions (CLARK 1993 [20])³.

L'espace qualitatif (dénommé espace conceptuel dans la citation suivante) apparaît comme plus fondamental que les constructions symboliques postulées pour rendre compte de la construction du sens.

Humans and other animals can represent the qualities of objects, for example when planning an action, without presuming an internal language or another symbolic system in which these qualities are expressed. As a consequence, I claim that the quality dimensions of conceptual spaces are *independent* of symbolic representations and more fundamental than these. (GÄRDENFORS 2000 [42] p. 43)

L'hypothèse est donc que la plupart des mots sont associés, de manière bidirectionnelle, avec des scènes de cet espace qualitatif, compris comme incluant toutes les situations que nous pouvons reconnaître dans notre environnement et dans la perception de nos états internes. Ce choix présente trois avantages décisifs. Premièrement, comme dans le cas des approches empiristes, les problèmes d'ancrage se retrouvent résolus : le sens d'un mot est compris quand il évoque une situation que nous pourrions reconnaître directement. Deuxièmement, la bidirectionnalité de l'évocation permet d'expliquer que le même système fonctionne en compréhension et en génération. Troisièmement, le caractère associatif, donc non rigide, de l'évocation, permet d'éviter tous les paradoxes des descriptions conceptuelles symboliques. Il explique du même coup de nombreux effets de polysémie, d'amorçage, de gradualité et de typicité, choses que les approches symboliques du système conceptuel étaient incapables de faire. Ainsi, le mot chien peut, sans problème, dans le contexte d'un catalogue de jouets, évoquer la photographie d'un chien en peluche. Or, ce simple test suffit à faire échouer bien des approches symboliques du système conceptuel, car la photographie du chien ne possède que très peu de caractéristiques en commun avec la description d'un chien : ce n'est pas un mammifère, ce n'est pas vivant, cela n'a pas d'épaisseur, cela n'est pas chaud, *et cætera*. En revanche, par le fait que l'association permet de suivre des liens de similarité, le mot chien peut évoquer la photographie du chien en peluche parce que cette photographie est suffisamment similaire, par ses propriétés qualitatives de forme, à un exemplaire de chien.

² Nous verrons, cependant, que certains mots de la classe fermée provoquent des mécanismes de nature symbolique. Par exemple, un mot comme pourtant déclenche la mise en place d'une incompatibilité logique de type $(x \uparrow y)$.

³ L'ensemble des stimuli physiques, pour la couleur, constitue une structure, de dimension infinie, de fonctions (temporelles ou spectrales). L'espace qualitatif correspondant est réduit à trois dimensions, comme le montrent les expériences de reconstitution à partir de jugements de distances effectuées sur des paires de stimuli. Noter que l'oreille effectue une analyse spectrale beaucoup plus fine, conduisant à un espace qualitatif de plusieurs centaines de dimensions (CLARK 1993 [20]). La pauvreté de notre perception qualitative du spectre lumineux est compensée par une résolution spatiale très fine : nous percevons consciemment la présence d'un pixel de couleur différente sur un écran informatique.

Il est important de noter que l'hypothèse d'une connexion entre le lexique et l'espace qualitatif se distingue des principes des différentes théories fondées sur la notion de typicité. Plusieurs modèles acceptent qu'un prototype soit caractérisé par des traits, affectés d'une plage d'admissibilité (CF. CHAPITRE 4). Or, le fait que ces traits puissent eux-mêmes être nommés fait de ces prototypes des représentations moléculaires, avec les paradoxes qui en résultent (CF. CHAPITRE 7). L'hypothèse qui sous-tend le présent chapitre laisse toute sa place aux effets de typicité, tout en renonçant à la possibilité que les catégories évoquées par les mots puissent se laisser analyser en traits nommables. Le mot spirale évoque une image pour de nombreux sujets, sans qu'ils puissent spontanément ou facilement en nommer les parties ou la structure. En d'autres termes, les mots renvoient à des représentations qui, à un moment donné, sont holistiques. Il ne s'agit pas de dire qu'une scène évoquée par un mot comme mariage, par exemple un mariage champêtre typique, est inanalysable. Le principe est que le mot mariage produit une évocation avant qu'une telle analyse soit effectuée. Autrement dit, la compréhension du mot mariage ne nécessite aucune décomposition structurelle de nature conceptuelle. Ce principe rejoint l'idée que les représentations associées aux entités lexicales sont, par nature, atomiques (CF. CHAPITRE 8).

Le caractère holistique des scènes évoquées par les mots ne signifie pas que ces scènes soient dépourvues de structure. Une scène de mariage champêtre peut être riche en détails. Même l'image la plus simple que le mot pomme peut évoquer est structurée. Nous savons, par exemple, que la pomme est à l'endroit, qu'elle est vue de côté et non de dessus, qu'elle est posée sur une coupelle ou qu'elle ne se trouve pas dans la pénombre. La nature précise de la structure des représentations qualitatives ne nous concernera pas ici. Il est possible qu'un niveau de schématisation soit construit à partir des données détaillées de la perception en isolant certaines caractéristiques qui seraient mémorisées à long terme. Lors du rappel, l'image serait reconstituée à partir de cette représentation schématique épurée par un processus de simulation (BARSALOU 1999 [3]). Les représentations schématiques structurées faciliteraient certaines opérations de compositions comme l'opération de synthèse que nous évoquerons plus loin.

Nous ne prendrons pas parti sur ces questions, en nous contentant de postuler l'existence de structures qualitatives (schématiques ou non) beaucoup plus riches que les constructions conceptuelles transitoires que nous allons être amenée à postuler. La restriction fondamentale que nous nous imposons, pour respecter le principe de l'atomisme des évocations lexicales, est que l'évocation d'une scène ne nécessite pas que le système linguistique ait accès à la structure de la scène. Nous allons même poser une exigence plus forte, avec le principe d'opacité structurelle. Ce principe stipule que la structure des scènes est inaccessible hors de l'espace qualitatif. Une telle exigence peut surprendre. Il nous faudra expliquer comment les opérations syntaxiques et logiques peuvent agir de manière systématique sur le sens si elles n'ont pas accès à la forme des scènes évoquées par les mots.

Dans une perspective technique de modélisation, l'idée à retenir est celle d'un espace qualitatif de grande dimensionnalité sur lequel se projettent les perceptions ou leur mémorisation. Dans cet espace sont possibles plusieurs opérations, comme l'association par similarité et certaines manipulations ou compositions comme celles qui sont prêtées aux images mentales. Nous ne prenons pas parti sur la nature digitale ou analogique de cet espace. Un tel espace est nécessairement fortement structuré, ne serait-ce que par la diversité des modalités qui en forment ses dimensions. Le principe d'opacité structurelle se traduit par le fait que les opérations conceptuelles n'ont pas accès à la structure interne de l'espace qualitatif. En particulier, un mot ne peut pas faire référence à une dimension isolée de l'espace qualitatif. Un mot peut évoquer une zone de l'espace qualitatif, par exemple "une région convexe". Il ne peut pas évoquer la composante de saturation du pixel 235119. Cet interdit permet d'échapper aux paradoxes des représentations conceptuelles symboliques

permanentes. La connexion des mots aux représentations qualitatives est une connexion souple tant par le lien, qui est simplement associatif, que par la cible qui est par exemple une zone centrée sur un percept typique. Cette cible ne saurait constituer une description du mot.

Ayant fait le choix d'associer les mots du lexique, non à des concepts symboliques, mais à des représentations de l'espace qualitatif, nous devons montrer que les difficultés qui pourraient résulter de ce choix peuvent être contournées. Au premier rang de ces difficultés, on trouve la question de la compositionnalité guidée par la syntaxe et l'application des opérations logiques comme la négation.

9.3. Interface symbolique

À partir du moment où les significations se situent dans l'espace qualitatif, il faut expliquer comment le sens d'un énoncé peut être obtenu, dans cet espace, à partir de la signification des mots qui le composent. Comme les relations syntaxiques qui structurent les agencements de mots sont, par nature, symboliques, nous devons définir une interface plausible grâce à laquelle ces relations guident la combinaison des représentations de l'espace qualitatif évoquées par les mots. De même, nous devons indiquer comment interfacer un espace de qualités continues avec le domaine du raisonnement qui comporte des opérations de type tout-ou-rien comme la négation.

Dans le cas du temps, nous avons supposé l'existence d'une interface symbolique dont la fonction consistait à appliquer une grille topologique sur le temps qualitatif. L'échelle de la projection dépendait des événements concernés et de leur connaissance par le sujet. Grâce à cette projection, certaines opérations sur le temps devenaient possibles, comme le zoom ou la séparation (CF. CHAPITRE 2). Notre objectif est de transposer ce modèle à la procédure de construction sémantique dans son ensemble. La transposition ne s'opère malheureusement pas aussi facilement. Dans le cas du temps, on pouvait supposer l'existence de mécanismes spécialisés permettant d'appliquer systématiquement les grilles temporelles sur le temps qualitatif. Dans le cas général de l'espace qualitatif dans son ensemble, il est nécessaire de changer d'optique : il s'agit d'éviter de postuler des mécanismes spécialisés appliquant les grilles sur chacune des très nombreuses variations que nous pouvons percevoir. Nous allons, pour cela, postuler l'existence d'opérations générales, capables d'agir indifféremment sur tous les aspects de la perception.

La synthèse de percepts

Le premier problème à résoudre, lorsque l'on cherche à faire correspondre, à un énoncé, une représentation qui se situe dans l'espace qualitatif, consiste à associer un percept à des combinaisons inédites de mots. Comment associer une image à l'expression *téléphone à roulettes* ? Un système associatif est généralement cantonné à l'évocation de situations déjà rencontrées, car elles constituent les attracteurs susceptibles d'être atteints lorsque les éléments sont activés. Par exemple, c'est le souvenir d'un téléphone à cadran rotatif qui pourra être ainsi associée à l'énoncé, dans la mesure où le cadran rotatif est, plus ou moins indirectement, associé au mot *roulette*. Un modèle de ce type de combinaison associative nous est donné par les moteurs de recherche sur la toile⁴.

⁴ Une recherche d'image sur Altavista avec la suite de mots *goat garden* donne successivement une chèvre dans un pré à proximité d'une maison, l'affiche d'un café accompagnée de la mention *Baked Goat Cheese with Garden Lettuces*, une image de zoo avec un enclos où sont visible quelques chèvres, et ainsi de suite.

Les êtres humains ne semblent pas être limités à ce fonctionnement purement associatif. Ils possèdent la capacité de combiner des percepts, produisant par exemple l'image d'un téléphone monté sur quatre roulettes, même s'ils n'ont jamais rien rencontré de tel.

[...] one of the reasons that imagery is useful is that we can combine objects in novel ways. For example, one can imagine Charlie Chaplin riding a zebra, and “see” whether he would have been able to peer over the top of the zebra’s head. A theory of imagery must explain how familiar components can be arranged in novel ways in images.

(KOSSLYN 1994 [61] p. 286)

Cette capacité de synthèse d'image ne consiste pas en une simple superposition. Il s'agit d'une authentique composition, dont les imperfections sont masquées par le fait que les détails ne sont pas accessibles en même temps que la forme globale.

Specifically, multipart images could be formed in the following way. Computationally, it makes sense to form a *global* image of the pattern first [...], which can then be elaborated if need be. [...] For example, if asked to image a bear, and then asked whether it has curved front claws, most people report that the initial image did not contain that information. Instead, they started with a rather vague image of the overall shape, and added the part as necessary.

(KOSSLYN 1994 [61] p. 292)

Cette capacité de combiner les images peut servir de base sémantique à une capacité protolangagière. Le protolangage, défini comme une forme de communication simplifiée caractérisée par l'absence de syntaxe, a été présenté comme un stade de la communication pré-humaine (BICKERTON 1990 [5]). La capacité à combiner librement les scènes aurait conféré un pouvoir d'évocation aux énoncés du protolangage (DESSALLES 2000 [30]).

Ce type de combinaison libre est toutefois caractérisé par son manque de systématisme. La simple juxtaposition, non syntaxique, des mots *jardin* et *chèvre* peut conduire à l'image d'une chèvre typique en train de paître dans un jardin typique. Cependant, rien n'interdit que le résultat soit un jardin en forme de chèvre. Si l'on augmente le rôle de l'association, on peut obtenir un jardin escarpé où seules des chèvres peuvent paître, ou encore un jardin où sont entreposés des fromages de chèvre. Même si le processus de synthèse de percepts est requis pour l'interprétation des énoncés comportant une syntaxe, il reste à expliquer comment les relations syntaxiques parviennent à le guider de manière à le rendre plus systématique.

Pour la suite, il est important de noter que notre capacité de synthèse de percepts inclut des aspects cinématiques. Nous pouvons faire évoluer une scène, par exemple imaginer la transformation d'une surface qui est progressivement peinte en rouge ou la transformation d'un enfant qui grandit en quelques secondes jusqu'à une taille adulte. Nous noterons la superposition d'un percept à un autre à l'aide du signe + et l'action d'une transformation ϕ par une notation fonctionnelle.

$$A = B + C$$

$$A = \phi(B)$$

De manière métaphorique, nous pouvons comparer la superposition de percepts à une opération de “copier - coller”, alors qu'une transformation se compare à une opération de “*morphing*”. La différence de notation ne doit pas laisser supposer que les mécanismes à l'œuvre sont fondamentalement différents. De manière idéale, la seule distinction réside dans le caractère cinématique ou non de la transformation. Certes, dans le premier cas, le percept C est supposé être identifié en tant que une représentation bien distincte, tandis que la transformation ϕ correspond à un aspect perceptif lié au percept A, comme le mouvement peut

être lié à un objet qui bouge⁵. On pourrait donc songer à singulariser les transformations par le fait qu'elles possèdent une intentionnalité, qu'elles sont nécessairement à propos de l'objet qui se transforme. Ainsi, une homothétie n'est une transformation que si un objet est imaginé ou perçu en train de subir cette homothétie. Cette propriété, toutefois, n'est pas propre aux transformations. La plupart des données qualitatives de l'expérience possèdent cette intentionnalité. Par exemple, le mot *rayures* renvoie à une qualité de l'expérience qui demande à s'instancier sur une scène, par exemple un zèbre, une pomme ou au moins une surface, pour donner lieu à une visualisation. Nous ne retenons donc des transformations que le fait qu'elles impliquent un aspect cinématique. Au-delà de cette différence, les représentations C et φ restent des entités perceptives.

L'opérateur de contraste

L'hypothèse fondamentale que nous proposons pour construire un modèle de construction du sens sur l'espace des qualités sensibles est que les êtres humains disposent de la capacité inverse de celle qui leur permet de combiner les percepts. La transposition de notre travail sur le temps à l'ensemble des qualités sensibles nous suggère une voie qui repose, fondamentalement, sur l'existence d'un opérateur de contraste.

Contraster deux percepts consiste à identifier une différence. Lorsqu'il s'agit de deux objets ressemblants, comme deux pommes sur un étale, nous constatons spontanément et sans effort que l'une est plus volumineuse que l'autre, ou que l'une est plus verte que l'autre. Lorsqu'un seul objet est présent, nous le contrastons spontanément avec la version typique du même objet. Ainsi, nous trouvons la pomme devant nous particulièrement brillante ou, au contraire, fripée. En l'absence de prototype, comme lorsque nous observons une forme abstraite, nous pouvons la contraster avec la "bonne forme" la plus proche⁶. Les prototypes ou les bonnes formes se comportent comme des attracteurs dans l'espace qualitatif. Leur activation, consécutive à celle d'un percept, permet au contraste d'être effectué. En résumé, lorsque nous portons notre attention sur un objet ou une situation, nous sommes susceptibles d'opérer un contraste avec un étalon, qui peut être un autre objet présent dans la scène, un objet récemment rencontré, un prototype ou un exemplaire mémorisé, ou encore une bonne forme.

La manière dont le contraste fonctionne sur le plan psychologique est hors de notre propos. Notons simplement qu'il pourrait être lié au phénomène de l'attention sélective.

[...] consider the role of selective attention in isolating features. During a perceptual experience, the cognitive system can focus attention on a meaningful, coherent aspect of perception. On perceiving an array of objects, attention can focus on the shape of one object, filtering out its color, texture, and position, as well as the surrounding objects.

(BARSALOU 1999 [3] p. 583)

Dans le cas le plus simple, le contraste produit une différence sous la forme d'un objet identifiable. Si nous observons un téléphone muni de deux antennes, le contraste avec un téléphone typique suggère l'addition d'une antenne, comme si la scène résultait de la superposition de l'antenne surnuméraire à l'image du téléphone typique. Nous noterons le contraste entre deux représentations A et B de la manière suivante.

$$A - B = C$$

⁵ L'adjectif "lié" est employé à dessein pour évoquer le liage perceptif (*binding*) qui permet à plusieurs caractéristiques, analysées dans des endroits différents du cerveau, de se rapporter à la même représentation.

⁶ Noter que le prototype et la bonne forme ne coïncident pas toujours, comme le montre l'exemple du rocher.

Dans certains cas, l'opération de contraste peut suggérer une transformation continue, un *morphing*. L'exemple le plus simple est celui d'une comparaison sur la taille de l'objet qui peut suggérer une transformation homothétique, un zoom. D'autres contrastes suggèrent des transformations continues que nous pouvons inverser mentalement (LEYTON 1992 [67])⁷. Le contraste avec une bonne forme peut ainsi suggérer une "histoire", comme si la transformation reconstruite s'était déroulée dans le temps (LEYTON 1992 [67]). Par exemple, s'il manque un morceau de la pomme que nous observons, nous imaginons qu'elle a été mordue. Nous noterons ce type de contraste de la manière suivante.

$$A - B = \varphi$$

Le résultat du contraste peut, dans chaque cas, être combiné à l'étalon pour redonner l'objet de départ.

$$A - B = C \quad \rightarrow \quad A = B + C$$

$$A - B = \varphi \quad \rightarrow \quad A = \varphi(B)$$

Encore une fois, le fait que le résultat du contraste φ s'interprète comme une transformation, ce que précédemment nous avons appelé une "histoire", ne signifie pas que le percept C et la transformation φ sont de nature différente. Le fait que la couleur grise d'un mur soit analysée comme l'action d'un noircissement, ou que la grande taille d'un enfant soit attribuée à sa croissance, n'enlève rien à la nature du contraste, qui est avant tout un décalage entre percepts.

Les propriétés de l'opérateur de contraste et son lien avec la théorie de la typicalité et la théorie de la Gestalt sont du ressort de la psychologie. Pour la suite, nous pouvons nous contenter des propriétés résumées dans les formules précédentes. Les écritures + et - ne doivent toutefois pas occulter le fait que les opérations correspondantes ne sont pas rudimentaires. Deux scènes peuvent différer par de nombreux aspects, mais l'opérateur de contraste va isoler une différence saillante. Lorsque nous contrastons Rousseau avec Hobbes, nous allons, à un moment donné, nous attacher à la différence d'époque alors que dans un autre contexte, nous allons concentrer le contraste sur leur conception différente du contrat social.

Techniquement, l'opérateur de contraste peut être modélisé par un calcul sur les parties instanciées des percepts. Si l'on représente l'espace qualitatif comme un espace de haute dimensionnalité, dans lequel chaque percept est représenté par un point, le contraste entre deux percepts est calculé sur les dimensions comportant des valeurs significatives pour au moins l'un des deux percepts. Le contraste réussit lorsque la différence est proche d'un percept reconnaissable ou d'une transformation. Le premier cas, de la forme $A - B = C$, se produira plus probablement lorsque des dimensions sont instanciées dans le percept contrasté sans l'être dans l'étalon. Le deuxième cas, de la forme $A - B = \varphi$, a plus de chances d'émerger lorsque certaines dimensions sont instanciées de part et d'autre avec des valeurs différentes, ce qui évoque une transformation φ faisant passer des valeurs de l'étalon à celles du percept. On comprend pourquoi ce type de contraste a plus de chance de réussir si l'étalon B est le prototype du percept A. Dans ce cas, de nombreuses coordonnées du prototype B ne sont pas instanciées, et celles qui le sont ont souvent une valeur proche dans le percept A, si bien que le contraste porte sur le nombre réduit des dimensions restantes, donnant un résultat lui-même très peu instancié. De ce fait, le résultat du contraste a des chances de ne pas être éloigné d'un percept connu, autrement dit d'être identifiable.

⁷ Les transformations minimales qui permettent de retrouver une bonne forme peuvent être, dans certains cas, calculées. Le calcul consiste par exemple à augmenter graduellement la convexité de l'objet (LEYTON 1992 [67]).

Dans certains cas, le contraste pourra être infructueux. Il n'existe aucune garantie quant au résultat, par exemple lorsque les deux objets sont trop dissemblables ou, au contraire, trop ressemblants⁸. De même, l'écriture $A = B + C$ suppose que l'on puisse superposer deux scènes B et C. Or, le résultat, là non plus, n'est pas garanti. L'opération semble plus facile si l'une des deux scènes, au moins, est peu instanciée. Il est cependant important que les opérateurs de contraste et de combinaison soient des opérateurs généraux, qui peuvent être systématiquement appliqués. Le principe que nous retenons est que les opérations de contraste et de combinaison peuvent être systématiques dans leur application, mais non dans leur résultat.

Avant d'aller plus loin, nous devons dissiper un malentendu potentiel. Le résultat de l'opérateur de contraste est une représentation, non d'une transition entre représentations. Dans les modèles dynamiques de la cognition, le passage d'une représentation à une autre suppose une trajectoire entre deux attracteurs. L'idée que nous proposons n'est pas de même nature. Le contraste est une représentation, construite *a posteriori* à partir de deux représentations perceptuelles. Le fait qu'un contraste puisse être compris comme une transformation explicite est une construction cognitive *ex post*, qui n'a pas nécessairement de rapport avec le cheminement d'une quelconque transition continue entre états neuronaux. En d'autres termes, le fait que nous analysions l'aspect d'un enfant comme homothétique par rapport au souvenir que nous avons de lui ne signifie pas que les états neuronaux soient dans un rapport d'homothétie !

Nous venons de définir un opérateur de contraste susceptible de donner une différence entre deux représentations de l'espace qualitatif. Cette différence peut, dans certains cas, être perçue sous forme graduelle, c'est-à-dire sous la forme d'une transformation explicite. Nous allons étudier maintenant comment l'opérateur de contraste peut constituer un début de réponse pour la question d'une interface systématique avec le langage et le raisonnement.

L'interface avec le langage

Tout modèle qui place la signification des éléments lexicaux dans l'espace des qualités sensibles se doit d'expliquer comment les sujets, d'une part, parviennent à verbaliser les situations perçues (production) et d'autre part, réussissent à constituer une scène qui ressemble à celle que leur interlocuteur cherche à évoquer (compréhension). Nous allons suggérer que les opérations de contraste et de synthèse que nous avons précédemment définies servent respectivement ces deux processus.

L'intérêt premier de la notion de contraste est de nous permettre d'expliquer la manière dont nous parvenons à verbaliser des situations. La principale difficulté rencontrée par les modèles empiristes de la cognition est qu'ils n'offrent pas de mécanisme satisfaisant permettant de construire une interface avec le langage. En particulier, le caractère incontrôlable des mécanismes associatifs, qui sont sensibles à l'expérience passée de l'individu, rend toute idée de systématisme illusoire. La vue d'une valise dotée de roulettes peut être associée à une charge élevée, à un voyage, à un chariot d'enfant, *et cætera*, autant qu'aux notions isolées de valise et de roulette. La probabilité pour qu'un système associatif puisse produire le syntagme valise à roulettes en percevant un tel objet est faible.

⁸ Certaines formes de plaisanteries du genre "Quelle est la différence entre ..." sont fondées sur l'échec d'un contraste entre deux objets dissemblables. À l'inverse, le jeu des sept erreurs repose sur la difficulté à contraster deux images complexes qui semblent identiques. Noter que, pour que le résultat du contraste dépasse le constat d'une différence et apparaisse sous une forme identifiable, il faut que la différence soit accessible consciemment. Or, ce n'est parfois pas le cas pour certaines expériences qualitatives, par exemple lorsqu'il s'agit de comparer des saveurs.

Dans ce qui suit, nous allons nous limiter à commenter l'opération linguistique de modification. L'opérateur de contraste permet non seulement d'expliquer la production d'une expression langagière comme celle qui consiste à modifier valise par roulettes, mais il permet également de prédire les aspects systématiques de la modification, dans la mesure où l'opérateur de contraste est un mécanisme général systématiquement applicable. Ainsi, en contrastant la pomme qui est devant nous avec une pomme prototypique, nous pouvons la désigner par le syntagme pomme fripée parce que le contraste entre les deux images de pomme permet de réaliser l'association avec le mot fripée. De même, en comparant deux valises, nous parvenons à l'expression la valise avec des roulettes ou, de manière équivalente, à l'expression la valise sans roulettes. Ainsi, le contraste que nous notons $A - B = C$ peut conduire à une expression langagière dans laquelle le percept A est désigné en modifiant le mot associé au percept B par le mot associé au percept C. L'association entre le produit d'un contraste et un mot approprié est facilitée par le fait que le contraste entre deux scènes est, généralement, peu instancié. De plus, la plupart des mots servant dans des contextes variés, ils sont généralement associés à des percepts eux-mêmes peu instanciés. Dans les deux cas, ce sont donc des zones étendues de l'espace qualitatif qui sont concernées, ce qui facilite l'association et donc la sélection, quand cela est possible, d'un mot approprié qui pourra être compris par l'interlocuteur. Le cas $A - B = \varnothing$ où le contraste produit une transformation \varnothing n'est pas différent, si ce n'est qu'il pourra donner lieu à une forme linguistique différente, comme la modification par un adjectif.

Considérons maintenant l'opération inverse, qui consiste à passer d'un énoncé langagier à l'évocation d'une scène. Un système associatif, doté d'un dispositif de synthèse de percepts, produira des résultats sans aucune systématisme, comme l'illustrent nos exemples d'association téléphone-roulettes ou chèvre-jardin. Les définitions que nous avons dégagées dans cette section permettent d'envisager, dans son principe, la synthèse en partie systématique d'une scène à partir d'un énoncé verbal. Nous ne considérerons, encore une fois, que le cas simple de la modification, lorsque les scènes associées aux mots sont des images concrètes.

Il s'agit de comprendre comment la présence d'indications de nature syntaxique permet de guider le processus de combinaison de scènes, en le contraignant jusqu'à le doter d'aspects systématiques. Les rapports syntaxiques introduisent toujours une hiérarchie permettant de savoir quel est le modifieur et quel est le modifié. Un modèle de la compositionnalité doit tenir compte de cette asymétrie pour calculer une représentation de l'énoncé dans l'espace qualitatif. Dans les termes que nous avons adoptés, il s'agit d'inverser l'opérateur de contraste, autrement dit d'effectuer une synthèse de type $B + C$ ou $\varnothing(B)$, lorsque les mots d'un énoncé évoquent, dans un cas, les percepts B et C, dans l'autre cas, le percept B et la transformation \varnothing . La manière dont ces opérations sont contrôlées reste mal connue. Dans le cas d'une modification du percept B par le percept C ou la transformation \varnothing , la hiérarchie syntaxique prescrit que le résultat doit être un percept susceptible d'évoquer le mot associé au percept B. En d'autres termes, une valise à roulettes reste une valise, et un grand enfant reste un enfant. Le cas de la combinaison $B + C$ peut sembler relativement plus simple, dans la mesure où la séquentialité de l'attention peut faciliter la création d'une scène composite.

For example, if one wanted to image the head of a bear, its representation in associative memory would be looked up, which might specify its location as "attached at the top front of the global shape." This spatial relation representation in turn is used to direct the attention window to that location relative to the global image [...].

(KOSSLYN 1994 [61] p. 293)

On parvient ainsi à cerner le type de mécanisme qui peut permettre à un individu de visualiser ce que peut être une valise à roulettes ou, pour prendre un exemple irréaliste, un arbre à tasses. Dans une certaine mesure, l'opération mentale réalise un "copier - coller" partiel. Un tel mécanisme de combinaison par copier - coller peut suffire, dans certains cas, à permettre à l'association de prendre le relais. Ainsi, une description comme la tasse à la feuille peut suffire à rappeler à l'interlocuteur le souvenir d'une tasse précise portant le dessin d'une feuille. Dans d'autres cas, l'opération de copier - coller suffit à la réalisation des inférences requises. L'expression une chaussette trouée permet par exemple à l'interlocuteur de comprendre que la chaussette est impropre à la situation, même s'il n'est pas en mesure de décider si le trou est à l'avant ou à l'arrière de ladite chaussette. Il est important de noter, cependant, que le mécanisme de copier - coller peut échouer. L'expression la tasse à la feuille peut très bien, dans certains contextes, laisser l'interlocuteur en panne de visualisation satisfaisante. Le choix de la préposition, en français, peut contraindre la combinaison des images. Alors qu'une expression comme l'homme avec la chemise semble désigner naturellement un homme portant une chemise, l'expression inverse la chemise avec l'homme semble bizarre. Elle peut néanmoins avoir un sens, par exemple si une chemise porte un dessin d'homme. La contrainte que pose ici le mot avec dans une expression comme A avec B est que la visualisation de la scène associée au mot B nécessite un "zoom" par rapport à celle du mot A, comme dans l'exemple de la tête d'ours dans la citation précédente. L'emploi d'une préposition comme sur, comme dans le syntagme la pomme sur la table, inverse le rapport figure - fond qui sous-tend le copier - coller : c'est l'image du modifié, la pomme, qui est positionnée à la suite d'un zoom sur le modifieur, la table.

La modification ne procède pas toujours par copier - coller. Ainsi, l'expression la roulette à valise peut signifier une roulette destinée à être montée sur une valise. L'opération de copier - coller ne s'applique pas ici, puisqu'un zoom sur l'image d'une roulette typique ne permet pas de positionner une valise. Un mécanisme plausible commence par une combinaison libre, fournissant l'image d'une valise à roulettes, suivi d'un isolement de la roulette. Cette dernière opération est causée par l'exigence syntaxique qui veut que, dans une expression (non figée) où le mot C modifie le mot B, le résultat est une image qui peut encore évoquer le mot B.

Considérons maintenant la deuxième possibilité de modification, celle où le modifieur est préférentiellement associé à une transformation ϕ . Ce genre de transformation, rappelons-le, n'a pas de statut représentationnel particulier et s'exprime dans l'espace qualitatif au même titre que les images statiques. Simplement, certains aspects qualitatifs de l'expérience sont associés à des mouvements ou à des changements. Ils impliquent donc une dimension cinématique. Ces transformations, comme les autres représentations de l'espace qualitatif, sont susceptibles d'être associées à des mots. Lorsque ces mots sont employés comme modifieurs, ils évoquent les évolutions correspondantes.

Le fait qu'un mot puisse évoquer une transformation ouvre la possibilité d'une forme de modification distincte de ce que nous avons appelé copier - coller. La combinaison avec une transformation, dans l'espace qualitatif, a pour effet de faire agir la transformation sur le percept associé au terme modifié. Ainsi, la modification du mot chien par l'adjectif petit peut provoquer une homothétie de l'image d'un chien prototypique. La modification du mot maison par l'adjectif détruit peut, selon le cas, donner l'image typique d'une maison détruite, par combinaison libre, soit l'image d'une maison subissant une destruction. Ce qui caractérise le deuxième cas est le fait que plusieurs intermédiaires peuvent être synthétisés.

La modification graduelle peut être, dans certains cas, modélisée de manière précise. Le principe utilisé pour représenter des combinaisons comme carré rouge ou carré rond consiste à rapprocher la représentation qualitative associée au mot carré de la représentation associée au mot rouge ou au mot rond selon les dimensions de plus grande variation. Cette modélisation

consiste à représenter l'espace qualitatif comme un espace de grande dimensionnalité dont les points représentent les percepts. Un prototype ou un souvenir de percept est représenté par une zone de cet espace, définie par un point et un vecteur donnant les variations admissibles sur chaque dimension. La zone représentant le syntagme carré rouge est obtenue en déplaçant le prototype associé au mot carré vers le prototype associé au mot rouge le long des dimensions où le prototype associé au mot rouge est le moins variable. Dans l'exemple du syntagme carré rouge, le déplacement est facile, car le prototype associé au mot carré est fortement variable sur les dimensions de couleur. Or, cette condition est loin d'être toujours réalisée. Dans le syntagme peau rouge, le prototype associé au mot peau n'est pas totalement variable sur les dimensions de couleur ; dans le syntagme carré rond, c'est le manque de variabilité dans les dimensions de forme qui résiste à la modification. Le calcul réalise alors un compromis, en déplaçant le prototype du modifié d'une manière qui tient compte, pour chaque dimension, de la variabilité du modifié et de celle du modifieur. Noter que le résultat peut tomber en dehors des zones admissibles pour l'un ou l'autre des deux prototypes : une peau rouge n'est pas vraiment rouge, et un carré rond, s'il prend la forme d'un carré aux coins arrondis, n'est ni rond ni carré. Cette technique présente l'avantage de reproduire le phénomène de projection du cône de couleur (GÄRDENFORS 2000 [42]) : la répartition des teintes de peau, dans l'espace qualitatif, reproduit, déplacé et réduit, le cône standard des couleurs.

L'interface avec le raisonnement

L'idée selon laquelle la signification des énoncés du langage ne se situe pas dans un monde indépendant des interlocuteurs, ni dans des échafaudages complexes d'étiquettes symboliques, mais dans la capacité d'imagerie, au sens large, des individus change radicalement le cadre explicatif de la théorie sémantique. Avec cette perspective, la communication est rendue possible par le fait que le locuteur est capable d'anticiper le type de scène que son interlocuteur va élaborer à partir de son message. Cette anticipation permet au locuteur d'avoir une idée des inférences que son interlocuteur va effectuer. Malheureusement, les modèles qui situent l'interprétation dans l'espace qualitatif, ou plus généralement tous les modèles qui incorporent une forme de gradualité dans leurs représentations, se prêtent mal à la production de certains formes d'inférences, notamment les inférences de type logique, essentielles au raisonnement. Il s'agit là d'une des principales causes de leur rejet par bon nombre de courants de recherche en sémantique. Si nous voulons rendre plausible la thèse défendue dans ce chapitre, nous devons montrer à quoi peut ressembler une interface systématique entre l'espace qualitatif et le raisonnement. Nous allons traiter en particulier du problème de la négation.

L'opération de négation est fondamentale pour le raisonnement et la gestion argumentative du discours. Or, elle semble hors de portée d'un système de représentations qualitatives. Que veut dire nier un percept ? Que voudrait dire nier un prototype de pomme ? Est-ce que la négation d'un percept est un autre percept ? Si oui lequel ? Sinon, de quel type de représentation s'agit-il ? Comment reconnaître qu'un percept donné peut évoquer la négation d'une pomme ?

Une première solution pour le problème de la négation consiste à placer un seuil de distance dans l'espace qualitatif. Ainsi, un percept ne serait pas une pomme parce que sa ressemblance globale avec la pomme typique tomberait au-delà d'un certain seuil, exprimé comme $k\sigma$, où σ mesure la variation globale admissible pour le prototype de pomme. Une telle solution est cependant grossièrement inadéquate, car elle néglige l'influence du contexte sur trois aspects cruciaux. Le premier concerne l'anisotropie de la comparaison. Par exemple, un oiseau blanc peut ressembler à tous égards à un corbeau typique, sauf pour la couleur. Une

mesure de ressemblance globale ne donnera qu'un écart faible, alors qu'un sujet humain aura tendance à nier le fait qu'il a affaire à un corbeau. En revanche, un corbeau mazouté à moitié plumé peut avoir un aspect fort éloigné du corbeau standard et néanmoins être considéré sans hésitation comme corbeau. Un deuxième problème concerne l'impossibilité, à cause de l'influence du contexte, de fixer une valeur pour le seuil k : ce qui apparaissait comme une table peut cesser de l'être chez un antiquaire pour devenir une console ou un secrétaire. Une troisième raison pour laquelle l'opération de négation ne peut se fonder sur la seule ressemblance globale est qu'une telle négation ne peut pas être étayée par une justification. Or, les sujets humains sont, le plus souvent, capables de justifier leur négation : l'oiseau observé n'est pas un corbeau parce qu'il est blanc.

Notre analyse de la négation se fonde sur l'utilisation de l'opérateur de contraste. Lorsque le contraste produit une différence non graduelle, de type $A - B = C$, la présence ou l'absence du percept C peut servir d'étiquette binaire pour fonder la négation. Par exemple, la perception d'un objet A peut déclencher un commentaire du type *ce n'est pas une valise*, lorsque le contraste avec le prototype B de valise fait apparaître une différence identifiable C comme l'absence d'une poignée. La présence constatée du percept C dans un cas et son absence constatée dans l'autre constituent un critère, dans le contexte courant, pour utiliser une négation. L'extrait suivant illustre ce principe de manière explicite⁹.

D- T'as vu comment il tient sa raquette ?

H- C'est un Chinois, les Chinois ils tiennent tous leur raquette comme ça.

D- C'est pas naturel ! Comment ils arrivent à jouer ?

H- Oh ! C'est pas un sport ça !

D- Enfin, tu transpires !

H- Un sport où tu cours pas, c'est pas un sport.

Au début de cette interaction, D s'étonne de la technique du pongiste qu'il est en train d'observer, tandis que H s'efforce à expliquer la raison de la particularité de cette technique. Les deux dernières interventions de H sont particulièrement éclairantes pour notre propos. Il nie le fait que le tennis de table soit un sport. Selon notre modèle, cette négation est le résultat d'un contraste entre sa représentation du tennis de table et la représentation typique qu'il se fait d'un sport. Ce contraste produit une image d'individus en train de courir dans un cas et non dans l'autre. Il peut donc étayer sa négation en arguant que les joueurs de tennis de table ne courent pas. Le critère de la course lui sert de moyen pour séparer ce qui est sport et ce qui n'en est pas un.

La capacité de nier semble plus difficile à expliquer en présence d'un contraste graduel. Si l'on rétrécit progressivement une voiture, il arrive un moment où elle risque de cesser d'être considérée comme une voiture, par exemple dans un contexte où il s'agit d'utiliser le véhicule pour se déplacer. Or, la négation *ce n'est pas une voiture* suppose de considérer une frontière définie entre les voitures et les non voitures sur l'échelle de la taille. Mais où placer cette frontière ? Est-ce au facteur d'homothétie 0.655957 ? Si l'on adopte une frontière floue, doit-on centrer cette frontière sur le facteur 0.655957 ? Le principe que nous proposons pour résoudre ce problème est que dans ce cas, la négation est rendue possible par l'application d'une grille topologique, comme dans le cas du temps.

Pour expliquer la possibilité de repérage temporel en l'absence de mémoire temporelle et exhaustive, nous avons proposé un modèle reposant sur un nombre limité de grilles qui pouvaient être utilisées pour manipuler le temps qualitatif telle qu'il ressort de la perception lacunaire et approximative que l'on a des situations imaginés ou mémorisés (CF. CHAPITRE 2). L'association temporaire des éléments d'une grille avec des situations en faisait une carte

⁹ L'extrait a été enregistré par des étudiants de l'ENST en 1991.

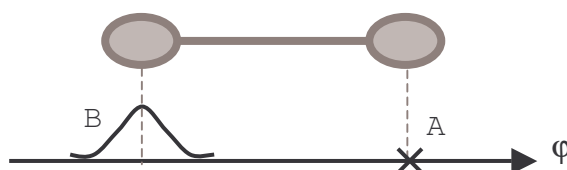
temporelle. Pour expliquer certains phénomènes comme la négation systématique en contexte, nous allons adopter une démarche parallèle, en définissant les grilles et les cartes conceptuelles. Nous supposons qu'en présence d'un contraste graduel obtenu entre deux représentations de l'espace qualitatif, nous sommes capables d'appliquer une grille conceptuelle. Les grilles conceptuelles sont exactement les mêmes que dans le cas du temps. Parmi elles, les grilles d'inclusion et de séparation jouent des rôles centraux, respectivement, pour la catégorisation conceptuelle et la négation. Leur application au résultat d'un contraste crée une carte conceptuelle. Par exemple, l'application d'une grille de type $\circ\text{-----}\circ$ permet d'opérer une séparation topologique dans une continuité, ce qui rend la négation possible. Considérons l'exemple spontané suivant.

L- Ça n'a jamais été des gambas, ce sont des crevettes. C'est trop petit pour être des gambas.

Ici, la locutrice se base sur la taille pour séparer les gambas des crevettes et pour en conclure que les crustacés qu'elle observe ne sont pas de la première sorte. Selon notre modèle, la locutrice a contrasté les crustacés, vendus en tant que gambas, avec le prototype correspondant. Ce contraste étant de la forme $A - B = \varphi$ avec $A = \varphi(B)$, elle a appliqué une grille conceptuelle de type $\circ\text{-----}\circ$ sur la variation graduelle des tailles. Une fois cette opération effectuée, les représentations A et B sont, pour la durée de son argumentation, des catégories strictement séparées, c'est-à-dire incompatibles. En d'autres termes, dans le contexte, les représentations A et B sont la négation l'une de l'autre. Ce résultat est obtenu ici grâce à une séparation par la taille. Ce n'est pas la couleur, la forme ou la possession de telle ou telle caractéristique qui distingue les deux catégories, ni une moyenne de tout cela. L'opérateur de contraste, dans le contexte, réduit la différence à une transformation φ , et c'est par cette transformation φ que les deux catégories vont être séparées.

L'opération de négation peut ainsi être décrite en deux temps : contraste, puis séparation topologique. Lorsque le contraste produit une différence de type tout-ou-rien, la séparation utilise cette différence pour engendrer la négation. Lorsque le contraste produit une différence graduelle, l'application d'une grille de séparation agit comme une quantification grossière pour séparer les deux catégories et justifier la négation.

Techniquement, l'application de la grille conceptuelle sur un contraste graduel peut être réalisée de manière simple. Notons tout d'abord qu'un contraste graduel $A = \varphi(B)$ crée un ordre total dans l'ensemble des intermédiaires, le long de la transformation φ , entre les représentations A et B. La séparation entre un percept A et un prototype B peut se fonder sur la comparaison du percept avec les variations du prototype. Si le percept A est supérieur, pour l'ordre lié à la transformation φ , à la plus grande variation connue du prototype B, alors les représentations A et B sont séparables. Le processus précis par lequel nous décidons de projeter une grille de séparation doit être étudié sur le plan psychologique, notamment pour cerner le rôle des exemplaires ou celui des catégories concurrentes (comme les crevettes dans l'exemple). L'important est de noter, comme dans le cas du temps, qu'il n'est pas nécessaire de postuler que l'ensemble parcouru par la transformation φ entre les représentations A et B possède d'autres propriétés que l'ordre total, ni même que les éléments de cet ensemble soient potentiellement accessibles. En particulier, l'opération de séparation ne nécessite pas de métrique.



L'intérêt évident du modèle est qu'il ne présente aucun des trois défauts de la méthode de seuillage de la distance globale. L'anisotropie, qui faisait défaut dans cette méthode, nous est fournie ici par la transformation φ ; la séparation n'utilise pas de seuil $k\sigma$; enfin, la justification argumentative de la négation est également donnée par la transformation φ , comme l'illustre l'exemple des gambas.

Nous constatons, sur le cas particulier de la négation, que l'opérateur de contraste, appliqué aux représentations de l'espace qualitatif, permet d'interfacer cet espace avec le raisonnement. Un point essentiel est que cette interface est systématique. Toute représentation qualitative peut être contrastée avec une autre représentation, et de ce fait être niée.

L'interface avec le raisonnement inclut d'autres dispositifs similaires, liés à l'utilisation d'autres grilles conceptuelles. Par exemple, dans le cas où le percept A est intérieur à la zone de variations du prototype B dans la comparaison sur l'ordre total créé par la transformation φ , c'est une grille d'inclusion qui s'applique, ce qui permet, par exemple, de réfuter une négation. Cette opération conceptuelle d'inclusion a souvent été réduite, à tort, à une simple opération de catégorisation. Il existe selon nous une différence fondamentale entre la catégorisation holistique, telle qu'un réseau de neurones entraîné est capable de la faire, et l'opération conceptuelle d'inclusion qui vient à la suite d'un contraste. Contrairement à la première qui intervient constamment dans la perception, la seconde intervient surtout dans l'argumentation, elle donne lieu à des justifications et peut servir à contrer une négation.

Dans ce qui précède, nous avons voulu décrire à grand traits ce que pourrait être un mécanisme de combinaison, dans le cas particulier de la modification linguistique opérant sur des images concrètes. L'objectif était de suggérer qu'une théorie de l'interface entre le langage et l'espace qualitatif ne pose pas de problème de faisabilité *a priori*. Pour étayer ce type de thèse, il faudrait ré-interpréter certains travaux de psychologie et de linguistique pour montrer qu'ils sont compatibles avec le cadre d'hypothèses que nous avons esquissé. Nous évoquerons plus loin la question de la systématisme d'une telle interface, et son extension à d'autres modes de construction syntaxiques, comme le groupe verbal.

9.4. Des structures géométriques aux formes logiques

Les espaces conceptuels

Le cadre théorique esquissé au début de ce chapitre présente de nombreuses convergences avec celui de la théorie des espaces conceptuels (GÄRDENFORS 2000 [42])¹⁰. La notion d'espace conceptuel correspond en grande partie à celle d'espace qualitatif, avec quelques nuances toutefois. L'hypothèse de base est qu'un espace conceptuel est doté d'une structure géométrique. Cette structure suppose au moins une relation d'entrelacement (*betweenness*) et une relation d'équidistance (GÄRDENFORS 2000 [42]). La première relation permet de fonder la notion de ligne; la seconde est une notion qualitative de distance. Dans certains sous-espaces de l'espace conceptuel, une mesure de distance est supposée permettre de calculer l'éloignement de tout couple de points.

Les dimensions de l'espace conceptuel incluent les dimensions qualitatives de l'expérience, comme les dimensions de couleur. Elles incluent en outre des dimensions plus abstraites. Par exemple, l'espace conceptuel peut inclure des connaissances sur un arbre

¹⁰ Dans cette section, la mention "espace conceptuel" désignera systématiquement son acception au sein de la théorie développée dans (GÄRDENFORS 2000 [42]).

généalogique sous forme d'une structure de graphe (GÄRDENFORS 2000 [42]). Dans ce cas, la distance entre points est discrète. Certaines dimensions de l'espace conceptuel peuvent correspondre à des actions élémentaires, mesurées par exemple par les mouvements relatifs simplifiés de parties du corps. Un exemple de dimension non perceptive est donné par les propriétés fonctionnelles affordantes des objets (GÄRDENFORS 2000 [42]). Ainsi, la propension d'une chaise à suggérer de s'asseoir compterait parmi les dimensions de l'espace conceptuel. L'une des dimensions est constituée par la représentation du temps. Comme d'autres dimensions, sa nature dépendrait de la culture. Par exemple, la dimension temporelle de sujet vivant dans certaines cultures aurait une géométrie circulaire (GÄRDENFORS 2000 [42]). L'espace conceptuel apparaît ainsi comme résultant de la manière dont on structure le monde.

Le caractère inclusif de l'espace conceptuel, tel qu'il est présenté dans ce cadre, le différencie de la notion générale de l'espace qualitatif définie dans le cadre de la psychologie cognitive. Notre choix est de nous restreindre aux dimensions qualitatives de l'expérience, telles qu'elles peuvent être mises en évidence par les études psychologiques. Les autres "dimensions", comme la relation de parenté dans un arbre généalogique, l'action motrice consistant à tirer sur une corde ou l'affordance fonctionnelle d'une chaise présentent un caractère contingent. Il s'agit pour nous de représentations qui n'existent de manière permanente qu'en tant que prototypes. Les dimensions qui nous permettent de percevoir les actions sont plus générales, liées entre autres aux perceptions kinesthésique et visuelle. L'aspect graduel lié à la plus ou moins grande affordance d'une chaise n'est pas, selon nous, une dimension fondamentale de l'espace qualitatif, mais le résultat d'un contraste graduel opéré en contexte. Le fait de nous restreindre aux dimensions qualitatives de base a pour conséquence que les aspects culturels ont une influence négligeable. En ce qui concerne le temps, l'existence d'une structure globale n'est pas requise. Il ressort de la première partie de notre travail que la seule hypothèse nécessaire est une perception locale, au sens de l'échelle considérée, de l'antériorité. Pour cette dimension comme pour d'autres, nous ne faisons pas l'hypothèse que les éléments sont systématiquement accessibles pour des comparaisons.

La richesse des dimensions de l'espace conceptuel s'accompagne avec une forte structuration. Ainsi, un domaine est défini comme un sous-espace dont les dimensions s'instancient simultanément (GÄRDENFORS 2000 [42]). L'exemple par excellence est le domaine des couleurs : la teinte et la saturation doivent être choisies ensemble. Au sein d'un domaine, une région est définie comme un ensemble de points pouvant posséder des propriétés géométriques comme la convexité et entretenir des relations topologiques comme la connexité ou l'inclusion (GÄRDENFORS 2000 [42]). C'est à partir de cette notion de région que sont définis deux types de représentations conceptuelles.

A natural property is a convex region of a domain in a conceptual space.

(GÄRDENFORS 2000 [42] p. 71)

A natural concept is represented as a set of regions in a number of domains together with an assignment of salience weights to the domains and information about how the regions in different domains are correlated.

(GÄRDENFORS 2000 [42] p. 105)

Ces définitions sont censées regrouper les caractéristiques intéressantes des schémas et des prototypes avec les avantages présentés par la structure géométrique. Un exemple de propriété naturelle nous est donné par les adjectifs de couleur. Un concept naturel est par exemple ce qui est évoqué par le nom pomme. Ce concept introduit des corrélations entre domaines, par exemple entre la couleur rouge et le goût sucré. Les pondérations de saillance (*salience weights*) sont supposées affecter le calcul de distance entre une perception et le

concept. Ces pondérations représentent la pertinence de chaque domaine dans le contexte (GÄRDENFORS 2000 [42]). Ainsi, le goût d'une pomme est plus pertinent si elle doit être dégustée que si on s'en sert comme d'une balle.

Les relations topologiques sur lesquelles ces définitions se fondent sont d'une pertinence indéniable en ce qui concerne la caractérisation des structures de l'espace qualitatif. Toutefois, la distinction entre propriétés et concepts, motivée essentiellement par la distinction adjectif - nom, et la réification des domaines, qui conduit à considérer un grand nombre de domaines contingents, peuvent, nous semble-t-il, pouvoir être évitées dans le cadre de la modélisation des mécanismes sémantiques. Le fait que certains domaines puissent être isolés dans l'espace qualitatif est une propriété de notre constitution biologique. Un prototype qui est restreint dans une dimension qualitative de couleur a donc plus de chances d'être restreint dans les autres dimensions de couleur que dans une dimension de goût ou de mouvement. La notion de domaine s'étend, cependant, à tous les groupes de dimensions, y compris les "dimensions" contingentes que nous avons mentionnées plus haut. Il en résulte des domaines eux-mêmes contingents, qui peuvent être modifiés par l'apprentissage ou la culture. L'éducation musicale, par exemple, permet à un mélomane de séparer des "dimensions" comme la note et le timbre, ou la mélodie et l'harmonie, alors que ces "dimensions" restent groupées au sein de domaines pour le néophyte. Or, la notion de domaine, malgré son caractère contingent, est utilisée pour fonder l'opposition propriété - concept, qui se reflète au niveau du langage dans l'opposition adjectif - nom (GÄRDENFORS 2000 [42]). Dans le cadre de notre modèle, l'opposition entre concepts et propriétés perd sa pertinence. Elle joue sur le nombre de dimensionsinstanciées dans le prototype, qui est souvent moins grand pour ce qu'on appelle couramment des propriétés. C'est le souhait d'établir un parallèle entre l'opposition théorique propriété - concept et la distinction linguistique adjectif - nom qui constitue, semble-t-il, la motivation essentielle pour inclure de nombreuses "dimensions" contingentes et de nombreux "domaines" contingents dans l'espace conceptuel. Si un adjectif renvoie à une propriété et qu'une propriété est une région d'un domaine, alors la beauté, les fables et les machines doivent constituer des domaines de l'espace conceptuel puisque beau, fabuleux et machinal sont des adjectifs ! Nous préférons renoncer à la notion de domaine et considérer que la distinction entre noms et adjectifs est avant tout syntaxique, liée à l'opération de modification, et fort peu sémantique¹¹. Ce qui, dans la théorie des espaces conceptuels, apparaît comme des régions dans des dimensions ou des domaines prédéfinis revient, dans notre modèle, à une délimitation contingente dans la gradualité produite par l'opérateur de contraste. La différence entre les deux approches peut se constater dans le test d'omnipotence du langage. Selon notre modèle, tout contraste est susceptible de donner lieu à une verbalisation. On pourra ainsi dire de deux voitures que l'une est plus féminine que l'autre. Dans la théorie des espaces conceptuels, il reste à expliquer comment une opposition qui ne s'inscrit pas dans une région convexe d'un domaine prédéfini peut être verbalisée sous forme de propriété. Cela nous semble difficile, sauf à étendre de manière incontrôlée le nombre de domaines à considérer.

Les pondérations qui, dans la théorie des espaces conceptuels, permettent de tenir compte du contexte correspondent au souci d'anisotropie de la comparaison que nous avons signalé plus haut. Dans notre modèle, l'anisotropie est assurée par l'opérateur de contraste. Noter que notre opérateur de contraste ne nécessite pas la donnée d'une distance globale. Le fait que certaines dimensions deviennent saillantes est dû au contraste lui-même, autrement dit à ce qui fait la différence entre les représentations. L'influence du contexte, dans notre

¹¹ Jusqu'à la fin du Moyen Âge, les grammairiens ne faisaient pas des adjectifs une classe séparée de celle des noms, alors qu'ils distinguaient nettement le verbe conjugué et son participe passé.

modèle, ne s'exerce pas sur l'opérateur, mais est prise en compte par les effets d'amorçage exercé sur les représentations elles-mêmes.

L'explication du mécanisme de modification langagière constitue une autre divergence entre la théorie des espaces conceptuels et le modèle que nous proposons.

The combination CD of two concepts C and D is determined by letting the regions for the domains of C , confined to the contrast class defined by D , replace the values of the corresponding regions for D . (GÄRDENFORS 2000 [42] p. 122)

Cette définition regroupe plusieurs cas. Le cas le plus simple se présente lorsque C est une propriété et D un concept. Le mécanisme consiste à recopier la région spécifiée par C dans le domaine correspondant de D (GÄRDENFORS 2000 [42]). Dans le syntagme pomme verte, la propriété associée au mot vert remplace la région du concept POMME qui se trouve dans le domaine de couleur. Dans cet exemple, le remplacement agit comme une spécification, car le domaine de couleur du concept POMME, avant la combinaison, était supposé instancié avec une liste comme [VERT, JAUNE, ROUGE]. Dans le syntagme livre vert, le remplacement agit comme une instanciation, puisque le domaine de couleur du concept livre était supposé non instancié. Dans certains cas, la modification se traduit par une révision du domaine du modifié, comme dans le syntagme éléphant vert où il y a conflit entre la région de couleur du concept ÉLÉPHANT et la région de la propriété VERT.

L'analyse de la modification ne peut se limiter à un tel mécanisme, car l'action du modifieur dépend souvent du modifié. Le concept ROUGE évoqué par le syntagme peau rouge est différent de celui qui est évoqué par le syntagme pomme rouge. On comprend avec cet exemple, et plus encore avec un adjectif comme grand, que la modification d'un objet tient compte de la comparaison avec les autres membres de sa classe. Le mécanisme, dans ce cas, consiste à projeter la structure du domaine de couleur sur la région correspondant au domaine de couleur du concept PEAU (GÄRDENFORS 2000 [42]).

Le cas d'une modification d'un nom par un nom semble moins immédiat. Le mécanisme consiste alors à réduire C à une région de l'un de ses domaines, ce qui nous ramène au cas précédent (GÄRDENFORS 2000 [42]). Dans le syntagme lion de pierre (stone lion), le concept associé au modifieur pierre confère la région correspondant à sa matière au concept LION. En revanche, il ne confère par la région qui correspond à sa forme, si bien que le concept lion conserve la sienne. On notera que l'effet sur le modifié dépasse largement le remplacement d'une région dans un seul domaine. Ainsi, le syntagme lion de pierre évoque un objet qui n'est pas vivant, qui ne peut pas bouger, *et cætera*, contrairement à la représentation typique évoquée par le mot lion.

Ces suggestions concernant le processus de modification comportent des aspects fort intéressants. En particulier, la projection de la structure d'un domaine sur un sous-ensemble de lui-même, comme dans l'exemple des couleurs de peaux, peut être généralisée. Ainsi, dans les termes de la théorie des espaces conceptuels, on peut comprendre l'expression grand théorème comme la projection de la structure liée au domaine des tailles physiques sur un domaine quelconque du modifié, ici le domaine des importances relatives des théorèmes. Ce mécanisme suffirait donc à expliquer les cas de modification métonymique et métaphorique.

Toutefois, le rôle important attribué à la notion de domaine dans cette explication du mécanisme de modification peut être évité dans le cadre de notre modèle. Le problème principal concerne la sélection du domaine pertinent pour une modification donnée. Que ce soit pour le syntagme lion de pierre ou pour celui de grand théorème, il faut expliquer comment le domaine pertinent est sélectionné. Une fois le domaine sélectionné, il faut prédire quel mécanisme s'applique. Pourquoi le percept évoqué par le syntagme éléphant vert est-il franchement vert et n'est pas d'une couleur grise qui tire sur le vert, alors que le syntagme

peau rouge évoque une peau dont la teinte typique rose - beige tire sur le rouge ? Dans la théorie des espaces conceptuels, le résultat d'une modification dépend de la pertinence relative du domaine associé au modifieur au sein du concept modifié (GÄRDENFORS 2000 [42]). Ainsi, la couleur serait d'une importance marginale dans le concept ÉLÉPHANT, ce qui lui permettrait d'être remplacée sans problème par la région de couleur apportée par le modifieur vert. Dans notre modèle, la sélection du "domaine" et celle du mécanisme de modification sont réalisées d'une autre manière. Dans le cas standard, le modifieur renvoie soit à un prototype C qui peut être "collé" à la scène évoquée par le modifié, soit à une transformation ϕ qui peut agir sur cette scène. Le processus permettant à la transformation ϕ d'agir prédit, comme nous l'avons expliqué plus haut, le phénomène de projection décrit à propos de l'exemple des peaux. La situation est différente lorsque ces mécanismes de modification sont en échec. Contrairement à ce qui est proposé par la théorie des espaces conceptuels, nous ne postulons pas de mécanisme additionnel spécifique pour sélectionner un domaine approprié, le domaine de matière dans le cas du syntagme lion de pierre ou le domaine de l'importance dans le cas du syntagme grand théorème. L'échec du processus de transformation par ϕ peut conduire à préférer le mécanisme de collage, comme dans l'exemple du syntagme éléphant vert. Lorsque ces deux mécanismes sont en échec, le processus d'interprétation déclenche une phase d'association opérant sur le modifié et sur le modifieur. Cette phase s'arrête dès que la modification devient possible. Ainsi, dans le cas du syntagme lion de pierre, la phase d'association nous conduit à la combinaison libre d'une statue de lion en pierre. Celui qui n'aurait jamais rencontré de telle statue aurait peut-être réalisé une autre modification, celle d'un lion aux muscles durs comme de la pierre. Dans ce cas, le passage de la pierre à la dureté, et celui du lion aux muscles seraient dus, encore une fois, à des associations. L'intervention d'une phase associative s'accompagne, on le voit, d'une perte de systématisme. L'anticipation du résultat de telles modifications dépend entièrement de la capacité du locuteur de prévoir les associations que son interlocuteur réalisera préférentiellement. Dans le cas du syntagme grand théorème, la phase d'association consiste par exemple à passer, de manière métonymique, du théorème lui-même à son importance dans le panthéon des mathématiques. Noter que la phase d'association, par un effet d'amorçage mutuel entre le modifieur et le modifié, peut précéder, et donc éviter, un échec éventuel des deux mécanismes à notre disposition, le collage et la transformation. Ainsi, dans le cas du syntagme livre épais, l'adjectif peut activer la prise en compte de caractéristiques physiques, alors que dans celui de livre obscur, c'est plutôt le contenu qui est pré-activé.

En résumé, notre modèle de modification ne suppose pas que certains mots, comme les adjectifs, renvoient à des dimensions ou des domaines prédéterminés. Comme les autres mots du lexique, ils renvoient de manière associative à des prototypes. Contrairement à la théorie des espaces conceptuels, qui doit gérer la sélection des domaines pertinents et différents cas de conflit entre propriétés incompatibles, nous ne faisons appel qu'à nos deux opérateurs de modification, le collage et la transformation, et au mécanisme général d'association. L'avantage de notre modèle est qu'il gagne en généralité et reste compatible avec le phénomène d'omnipotence du langage. Cette omnipotence sera en défaut dans la théorie des espaces conceptuels, chaque fois que l'on cherche à verbaliser autre chose que l'un des domaines prédéfinis.

Un dernier problème concerne la notion de propriété. Si une propriété renvoie à une région convexe aux frontières bien délimitée de l'espace conceptuel, comment expliquer les effets de gradualité ? Peut-on dire que le syntagme adjectival plus ou moins rouge délimite une zone convexe ? On pourrait considérer qu'une telle expression renvoie à une mesure de la distance avec le centre de la région rouge. Dans ce cas, on perd l'idée de propriété en tant que région convexe, et surtout on perd la possibilité de la négation ceci n'est pas rouge. Dans le cas où une propriété renvoie à un domaine multidimensionnel, comme la propriété coloré,

comment interpréter l'expression plus coloré ? Le centre du domaine coloré correspond à un gris moyen, ce qui n'est certainement pas un sens acceptable pour la limite du syntagme plus coloré. Si l'expression s'interprète comme soulignant la saturation des couleurs, l'espace obtenu n'est pas convexe au sens de la métrique habituelle. Si l'expression renvoie à la quantité de couleurs présentes, on est amené à créer un nouveau domaine pour expliquer la propriété d'être coloré. Notre modèle permet de répondre à ces difficultés. L'expression plus ou moins rouge renvoie à une gradualité dans la modification, conçue comme une transformation ϕ . Il en est de même pour l'expression plus coloré, comprise comme portant sur la saturation des couleurs. Le modèle permet en outre l'intervention d'un mécanisme d'association, de manière à rendre possible un contraste graduel menant à une surface multicolore.

Prédicats, discours et cohésion sémantique

L'interprétation des phrases du langage ne peut se résoudre dans les seuls espaces qualitatifs. Les phénomènes systématiques qui conduisent aux aspects logiques du raisonnement nécessitent un autre niveau d'explication. Nous avons montré comment, dans notre modèle, l'opérateur de contraste permet l'application de grilles conceptuelles et la formation de cartes conceptuelles qui assurent certaines opérations logiques, au premier rang desquelles la négation. Il s'agit ici d'indiquer, à grands traits, comment notre modèle peut s'interfacer avec les couches supérieures du traitement du langage, notamment la structure du discours.

L'introduction de représentations de type logique, des prédicats constitués à partir des expressions langagières, est au centre de la théorie de la représentation du discours (*Discourse Representation Theory*, DRT) (KAMP & REYLE 1993 [56]). Ce cadre théorique, qui s'inscrit dans le cadre de la modélisation logique du langage naturel, postule une distinction entre le sens et l'interprétation d'une expression langagière : un modèle de la sémantique doit inclure une composante qui rend compte du phénomène de l'interprétation des fragments discursifs (KAMP & REYLE 1993 [56])¹². Ce processus d'interprétation est en charge de produire les représentations impliquées dans les mécanismes de raisonnement, au premier rang desquels se situent les mécanismes déductifs.

Not only does it seem impossible to give a satisfactory account of logical deduction without accounting also for meaning; conversely, if a theory of meaning is to include – as we think it should – a theory of interpretation, an account of meaning includes in its turn an account of deduction – at least, it will have to include all those forms of deduction that enter into the process or processes of interpretation; but those forms cover [...] a very substantial part of deduction in general – and for all that is known at present, they may cover all.

(KAMP & REYLE 1993 [56] p. 24)

L'interprétation d'un fragment de discours, dans le cadre théorique de la DRT, consiste à former une représentation structurée en appliquant un ensemble de règles de construction aux énoncés qui constituent ce fragment (KAMP & REYLE 1993 [56]). Cette théorie utilise un formalisme proche de la logique des prédicats avec une sémantique ancrée dans le cadre de la théorie des modèles. Une représentation structurée contient des variables qui représentent les référents du discours, et un ensemble de formules qui imposent des conditions sur ces variables (KAMP & REYLE 1993 [56]). Cette forme de représentation logique n'est pas astreinte à la contrainte d'isomorphisme. L'interprétation est dirigée par la composition

¹² Dans cette section, la mention "interprétation" désignera systématiquement son acception au sein de la théorie développée dans (KAMP & REYLE 1993 [56]).

syntactique, mais les représentations qui en résultent sont “plates”, c’est-à-dire qu’elles ne conservent pas la structure syntaxique. La construction de la structure est réalisée de manière cumulative, énoncé par énoncé. L’interprétation de chaque énoncé contribue à la construction en transformant la structure déjà existante. Cet aspect incrémental de l’interprétation est lié au phénomène important de la cohésion (*cohesiveness*) sémantique.

The connection between cohesiveness and incremental discourse processing is, in rough outline, this: to understand what information is added by the next sentence of a discourse to what he has learned already from the sentences preceding it, the interpreter must *relate* that sentence to the information structure he has already obtained from those preceding sentences. (KAMP & REYLE 1993 [56] p. 59)

L’idée principale est que dans un fragment cohérent de discours, différents énoncés sont liés entre eux par des références croisées. Un exemple de ce phénomène nous est donné par l’aspect anaphorique des pronoms. Prenons l’exemple du fragment Pierre connaît Marie. Elle le fascine. L’interprétation de la première phrase fait intervenir deux fois la règle de construction concernant les noms propres, ce qui introduit deux nouveaux référents x et y , renvoyant respectivement aux individus Pierre et Marie. L’analyse produit ensuite les conditions Pierre(x), Marie(y), et x connaît y . La règle de construction concernant le pronom, déclenchée par la deuxième phrase, introduit deux nouveaux référents w et z et produit la nouvelle condition w fascine z . La même règle produit deux autres conditions qui lient les référents introduits dans les deux phrases, produisant les contraintes $w = y$ et $z = x$. Dans le cas de cet exemple, c’est l’accord du genre qui détermine l’identification des référents. Dans le cas général, la règle peut utiliser d’autres traits d’unification, par exemple des traits sémantiques comme humain/non humain. À partir de la phrase Pierre connaît Marie qui le fascine, d’autres règles vont produire une représentation contenant trois référents x , y , et z , et les conditions Pierre(x), Marie(y), x connaît y , y fascine z , et $z = x$. À partir de ces structures construites en suivant le discours pas à pas, d’autres processus inférentiels de type logique peuvent s’appliquer, de manière à effectuer les déductions requises.

Nous avons le même souci, dans notre modèle, de pouvoir représenter le sens d’un énoncé pour le rendre accessible aux mécanismes logiques. Nous avons mentionné ce souci en parlant du principe de liaison sémantique (CF. CHAPITRE 5). Le principe de liaison sémantique désigne le moyen principal par lequel les différents syntagmes d’un énoncé se trouvent reliés au niveau sémantique (DESSALLES 2000 [30]). Le mécanisme peut être décrit comme un partage de variables entre prédicats, comme dans le cadre de la DRT. Ainsi, dans une écriture à la Prolog, un syntagme comme le livre qui est sur la table du salon pourra être décrit par livre(L), sur(L, T), table(T), dans(T, S), salon(S). Dans ce type d’écriture, le fait de donner le même nom aux variables de différents prédicats établit la liaison sémantique.

Les principes sous-jacents à la DRT ou le principe de liaison sémantique peuvent contribuer à résoudre simplement l’interface entre le discours et le raisonnement. Cependant, ils le font au prix d’une hypothèse lourde : la correspondance entre les mots du discours et des prédicats du formalisme logique. Nous avons longuement critiqué l’idée d’une telle correspondance préétablie (CF. CHAPITRES 7 & 8). C’est pourquoi notre modèle est construit autour d’une procédure de construction du sens qui produit des représentations symboliques éphémères. Lorsque l’opérateur de contraste est enchaîné avec une grille de séparation, le résultat est un prédicat négatif. Son expression langagière dépend des mots auxquels les ingrédients du contraste parviennent à s’associer dans le contexte. Le résultat peut être une expression positive ou négative. Ainsi, si l’opération cognitive consiste à séparer le véhicule perçu du prototype de voiture sur un contraste de nationalité, le résultat verbal peut être

indifféremment positif, comme dans la phrase *c'est une voiture de marque étrangère*, ou négatif, comme dans *ce n'est pas une voiture de marque française*. Notre modèle, de manière assez naturelle, permet ainsi d'expliquer la production, à la volée, de prédicats logiques. On notera que l'existence cognitive du prédicat n'est pas strictement liée à la disponibilité de mots capables de l'exprimer.

La production des prédicats à la demande est un pas important pour expliquer l'interprétation du discours. Il nous faut cependant expliquer par quel mécanisme ces prédicats éphémères peuvent être liés entre eux, comme le demande le principe de liaison sémantique.

La réponse que notre modèle est susceptible d'apporter à ce problème est un prolongement de la procédure de localisation que nous avons décrite à propos du temps (CF. CHAPITRE 2). En prononçant la suite de phrases *hier, avant le repas, elle est sortie pour acheter des cigarettes ; elle a rencontré son professeur, nous ne nous contentons pas de faire mention de l'événement de la rencontre*. Nous le localisons dans le temps à des fins pragmatiques¹³. Pour cela, nous utilisons la procédure de localisation temporelle de manière réursive. L'idée que nous voulons suggérer ici est que le mécanisme à l'œuvre pour la production du syntagme *le livre qui est sur la table du salon* est similaire. Il s'agit de donner les moyens à l'interlocuteur, pour des raisons pragmatiques (par exemple si celui-ci est censé aller le chercher), de localiser le livre. Pour cela, le locuteur imagine l'objet dans son espace qualitatif. Par association, il produit l'image de l'objet dans un décor plus ou moins précis. En réalisant un contraste de type figure - fond entre l'objet et le décor, il en vient à considérer la table et à former la relation *sur*, selon un mécanisme qui peut être celui que nous décrirons dans la prochaine section. À ce stade, le syntagme *le livre qui est sur la table* pourrait être émis. Cependant, la trop grande ambiguïté de ce message, dans le contexte, pousse le locuteur à repérer la table pour son interlocuteur. De nouveau, il effectue contraste figure - fond de l'image de la table avec son décor, ce qui l'amène à produire la deuxième prédication et, ainsi, le syntagme complet. Dans un tel schéma, l'aspect réursif de la syntaxe est une conséquence de la réursivité de la procédure de repérage. La connexion des cartes conceptuelles entre elles, comme dans le cas des cartes temporelles, réalise la liaison sémantique. L'objet qui sert de fond dans la première carte de notre exemple devient la figure dans la deuxième carte. Cette liaison permet la constitution d'une représentation, au-delà des prédicats isolés, du discours dans son ensemble.

Ce qui distingue cette approche d'une théorie comme la DRT est essentiellement la place qui est ménagée aux représentations qualitatives. Les cartes conceptuelles étant par nature éphémères, la production des inférences logiques nécessite une lecture répétée de la scène construite. Alors qu'une représentation de type DRT peut devenir assez complexe, accumulant les conditions exprimées dans les différents syntagmes, c'est la scène qui, dans notre modèle, conserve le résultat des différentes prédications. En conséquence, les entités symboliques postulées dans notre modèle restent simples, alors que la DRT est sujette au problème de la monotonie représentationnelle (CF. CHAPITRE 8).

¹³ Le fait de localiser ainsi l'événement de la rencontre dans le temps donne les moyens à l'interlocuteur de comprendre, par exemple, que la rencontre est récente et qu'elle est due au hasard, ce qui lui permettra de faire les inférences souhaitées.

9.5. Extensions du modèle

La gestion de l'abstraction

Le modèle dont nous avons jeté les bases au début du présent chapitre repose sur l'existence d'un espace qualitatif sur lequel agissent les opérateurs de synthèse et de contraste. Cela semble suggérer que la pensée conceptuelle n'est possible que lorsqu'elle est sous-tendue par des données de la perception. En d'autres termes, la pensée conceptuelle ne pourrait être que concrète. Or, le fait que les phrases de ce paragraphe puissent être comprises du lecteur prouve à l'évidence le contraire. Dans un modèle fondé sur descriptions, l'abstraction n'est pas un problème, au contraire. Il est souvent plus facile de trouver une description à peu près satisfaisante d'un terme abstrait, par exemple linguistique, que d'enfermer un objet concret dénommé par un mot comme cheval dans une description. Notre modèle est-il compatible avec le fait que nous parvenions à comprendre les expressions abstraites ?

Nous avons déjà rencontré cette situation dans le cas du temps (CF. CHAPITRE 2). On voit facilement comment les grilles temporelles s'appliquent à la mémoire épisodique : chaque élément de la grille se retrouve lié à un épisode vécu. Ainsi, en nous souvenant de la situation exprimée par la phrase Pierre est venu avant Marie, nous appliquons une grille de séparation $\circ\text{---}\circ$. Cette application crée une carte où les deux moments sont associés aux deux souvenirs, respectivement l'arrivée de Pierre et celle de Marie. Or, lorsqu'il s'est agi de représenter le raisonnement de Zénon, il n'était plus question de s'ancrer sur la mémoire épisodique. Dans ce cas là, le système des grilles tournait "à vide" : elles étaient ancrées dans un système représentatif minimal. Certaines personnes, à qui l'on demande ainsi, de manière répétée, d'imaginer un instant juste avant le déclenchement de la sonnerie, déclarent visualiser les positions des moments sous une forme spatiale fortement épurée. Les cartes temporelles qui en résultent sont dépourvues d'échelle, ce qui provoque l'entrée en boucle recherchée par Zénon.

Le raisonnement de Zénon est un exemple de raisonnement abstrait. Nous pouvons nous en inspirer pour cerner le mécanisme qui rend la pensée abstraite possible dans le cadre de notre modèle. Considérons la phrase la pensée conceptuelle ne peut être que concrète. Dans le contexte de cette section, le lecteur associera l'adjectif concret à l'idée d'espace qualitatif : est concret ce qui se situe dans un espace qualitatif. Il associe le syntagme pensée conceptuelle à différents exemplaires de pensées. La structure grammaticale de la phrase l'invite à former une grille d'inclusion $\text{---}\circ\text{---}$, le forçant ainsi à placer ces exemplaires de pensée au sein du concret des espaces qualitatifs. Cette inclusion n'a de sens que de manière métonymique, comme lorsque l'on prononce la phrase cette personne ne peut être que sur ta liste. Le lecteur de l'exemple comprend donc que certains aspects de la pensée conceptuelle héritent des propriétés de l'espace qualitatif. Or, les associations qu'il forme avec des exemplaires de pensée résistent à cette inclusion et à cet héritage. En contrastant ces pensées avec celles qui comportent des aspects qualitatifs, il produit une grille de séparation qui lui permet de réfuter l'énoncé.

Le commentaire que nous venons de faire sur notre exemple ne doit pas être pris de manière trop littérale. Nous ne disposons bien entendu pas des moyens de montrer que les choses se passent ainsi. Il n'a d'autre but que de montrer que notre modèle est compatible avec la production de pensées abstraites. Les aspects qualitatifs ne sont jamais totalement absents. Dans les cas les plus abstraits, qui se produisent par exemple lorsque l'on découvre un discours théorique dont on ne maîtrise pas les concepts, on peut avoir recours à une

représentation spatiale épurée. Tout se passe comme si les grilles engendraient leur propre aspect qualitatif. Ainsi, une contradiction peut être “visualisée” comme une séparation spatiale, alors qu'une implication apparaîtra comme une inclusion spatiale. Il est intéressant de noter que la gestuelle spontanée inconsciente des locuteurs reproduit parfois, dans le domaine spatial, la topologie de nos grilles.

The speaker metaphorically divided space into two spheres: one for the actual morality of the characters (the left) and the other for their apparent morality (not the left).

(MCNEILL 1992 [72] p. 229)

The moral ambiguity the speaker found in the film was visualized as space, then, and laid out like a map with spatial regions for the moral conflict. The absolute spatial assignments (left versus non-left or front versus right) had no significance in themselves: what mattered was the oppositions.

(MCNEILL 1992 [72] p. 156)

La traduction des relations logiques ou mathématiques sous forme spatiale peut également jouer un rôle important dans le raisonnement abstrait. Précisons, cependant, que la capacité qui nous permet d'exprimer les grilles sous forme spatiale n'entraîne en aucune façon le fait que les grilles postulées dans notre modèle soient réductibles à des métaphores spatiales. Cette capacité est due au simple fait que notre représentation qualitative de l'espace est beaucoup plus riche que les grilles conceptuelles du modèle. Ce sont les grilles qui nous permettent, dans quelques cas particuliers, d'utiliser l'espace pour raisonner logiquement, en nous permettant par exemple d'opposer ce qui est intérieur à ce qui est extérieur. Ce n'est pas l'inverse : notre capacité générale de nous repérer dans l'espace ne peut fonder le raisonnement logique. La puissance des grilles conceptuelles est due à leur extrême simplicité, propriété que ne possède pas la représentation spatiale.

Les mots et les grilles

Nous avons vu, dans le cas du temps (CF. CHAPITRE 2), que certains mots étaient associés à des instructions permettant la mise en place des grilles conceptuelles. Ainsi, le mot *avant*, dans une expression comme *X avant Y*, convoque une grille de séparation. Dans le cas général de l'interface entre l'espace qualitatif, le langage et le raisonnement, un certain nombre de mots aiguillent la formation des contrastes et l'application des grilles conceptuelles. Par exemple, certains emplois de la copule convoquent *in fine* une grille d'inclusion. Dans la phrase *ce livre est cher*, l'adjectif *cher* invite à opérer une transformation graduelle $\phi(B)$ à partir du prototype *B* de *livre*, produisant un axe qui caractérise le prix. La transformation isole une zone de “cherté” sur cet axe. Ensuite, l'intervention d'une grille d'inclusion permet de situer le livre considéré à l'intérieur de cette zone de cherté.

La synthèse et le contraste graduel, en ramenant la complexité de l'espace qualitatif à un ordre total contrôlé par la transformation ϕ , permettent de retrouver les relations topologiques que nous avons dans le cas du temps (CF. CHAPITRE 2). Considérons les deux phrases suivantes.

- (a) Ce livre est plus cher que l'autre.
- (b) La soutenance est après les vacances.

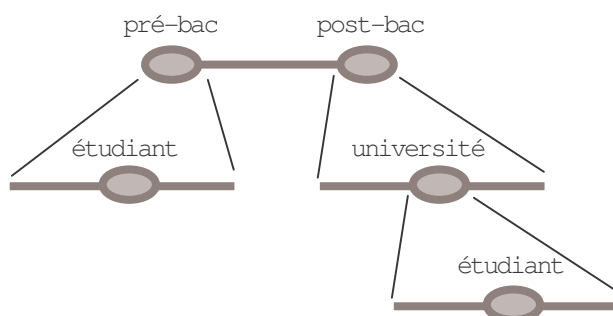
Les interprétations des phrases (a) et de (b) sont analogues. Dans les deux cas, nous disposons d'un ordre total, donné par le contraste de cherté pour les livres et par la relation de précédence temporelle pour la soutenance. La compréhension de la phrase (b) suppose l'enchaînement d'une grille de début, convoquée par le mot *après*, puis d'une grille

d'inclusion qui positionne la soutenance au sein de la période de la première grille. Le traitement de la phrase (a) est supposé procéder exactement de la même manière sur le contraste de cherté.

Au cours du traitement de la phrase (a), la projection sur l'axe correspondant à la transformation ϕ d'une grille conceptuelle produit une carte conceptuelle, analogue aux cartes temporelles. Cette carte possède une échelle, qui sera différente selon qu'il s'agit d'un livre cher ou d'une maison chère. Dans la phrase ce livre est très cher, l'indication quantitative nous éloigne du prototype, dans la direction de la transformation ϕ , d'une quantité qui dépend de l'échelle. Dans la phrase ce livre est trop cher, une grille de séparation distingue ce qui est admissible de ce qui ne l'est pas sur l'axe graduel produit par la transformation ϕ . Ensuite, comme précédemment, une grille d'inclusion permet de situer le livre en question à l'intérieur de la zone non admissible. Ainsi, la copule ou des mots comme plus, très, trop, sont associés à des types de relations sur un axe graduel. On constate, par ces exemples, que l'opérateur de contraste peut définir une gradualité unidimensionnelle sur laquelle le sens de nombreux mots grammaticaux de la langue peut être décrit.

Les opérations logiques

On attend d'un modèle sémantique qu'il soit en mesure, au moins en principe, de produire les représentations qui se prêtent à des mécanismes logiques comme la contradiction, la disjonction ou l'implication. De ce point de vue, notre modèle possède quelques atouts. Le caractère topologique des grilles conceptuelles permet de constituer les relations tout-ou-rien de la logique. Ainsi, la nécessité d'avoir le baccalauréat pour un étudiant d'université peut être représentée par un enchaînement de deux cartes, une carte de séparation et une carte d'inclusion, comme dans la paraphrase l'université, c'est après le bac. De même si l'on veut représenter la nécessité d'avoir au moins la note 10 pour valider un module d'enseignement. L'implication s'exprime donc par la combinaison de deux cartes, une carte pour séparer deux zones sur un contraste, et une carte d'inclusion pour inclure l'antécédent dans la zone associée au conséquent. La contradiction s'analyse avec une carte supplémentaire : un même objet doit être positionné dans deux zones séparées d'un contraste. Le cas de l'étudiant d'université qui n'a pas passé le baccalauréat provoque une contradiction, car il est inclus à la fois dans la zone pré-bac (sur un contraste temporel ou sur un contraste de niveau d'études) et dans la zone université qui est elle-même incluse dans la zone post-bac.



Noter que l'enchaînement de cartes d'inclusion permet de reproduire des opérations logiques comme le *modus ponens* ou le syllogisme. La lecture d'une partie du schéma peut être lue comme un syllogisme : (i) toutes personnes qui est à l'université a le baccalauréat ; (ii) cet étudiant est à l'université ; *ergo* (iii) cet étudiant a le baccalauréat.

La contradiction étant une opération de base du niveau argumentatif (DESSALLES 1993 [28]), il est important que le modèle sémantique soit en mesure de la reproduire. En revanche, certaines opérations comme la disjonction peuvent être produites de

manière entièrement symbolique. Ainsi, à partir d'une contradiction constatée entre p et q , le sujet peut exprimer la disjonction $\text{non } p$ ou $\text{non } q$ sans être obligé de former de nouvelle représentation sémantique (ou en les formant *a posteriori*).

Le propre du calcul logique est d'être systématique. En particulier, tout prédicat peut être nié. Si cette négation formelle s'accompagne chaque fois de la pensée correspondante, le système de cartes conceptuelles de notre modèle peut-il parvenir à représenter dans chaque cas la pensée niée ? Si nous reprenons les exemples de la sous-section précédente, peut-on facilement représenter à l'aides des cartes la gradation des expressions négatives (a) ce n'est pas un livre ; (b) ce livre n'est pas cher ; (c) ce livre n'est pas plus cher que l'autre ? Le cas de la phrase (a) est classique : l'objet est contrasté avec le prototype de livre, et une carte de séparation est projetée sur le contraste. Pour la phrase (b), le contraste est imposé, c'est celui de cherté, et nous avons toujours une grille de séparation. La phrase (c) peut sembler problématique. Dans la mesure où les opérations de contraste interviennent à la fois dans le traitement du comparatif plus et dans la négation, il semble que nous ayons là une opération de degré deux¹⁴. Le traitement de la phrase (c) est analogue à celui d'une expression temporelle comme (d) la soutenance n'est pas après les vacances. Dans la sous-section précédente, nous avons vu comment les versions positives de (c) et de (d) pouvaient être traitées. À quelle opération sémantique supplémentaire correspond la négation ? Nous avons associé la négation à une grille de séparation. À quoi faire correspondre les éléments d'une telle grille ? Notre modèle ne prévoit pas de traitement univoque dans un tel cas. Cette indétermination, loin d'être une faiblesse, reproduit l'indétermination de la négation argumentative. Après avoir entendu la phrase (c), on ne sait pas si la suite est (c1) ils ont le même prix ; (c2) il est moins cher ; ou (c3) il est plus cher que celui que j'ai montré tout à l'heure. De même, après la phrase (d), on peut situer la soutenance pendant les vacances, avant les vacances ou après un autre repère. Dans notre modèle, la négation argumentative s'interprète donc comme invitant à "défaire" les opérations sémantiques déclenchées par l'interprétation de la phrase positive, avec l'ambiguïté que cela suppose. Noter, cependant, que la situation reste fortement contrainte dans le cadre du modèle, ce qui permet de prédire le faible éventail des possibilités (c1)-(c3).

Les relations actantielles et les configurations

Notre modèle, sous sa forme actuelle, permet d'envisager le traitement de la modification, par exemple la modification d'un nom par un adjectif ou par un complément. La modification est généralement traduite, en logique, par un prédicat à une place. Actuellement, notre modèle ne permet pas d'expliquer la formation des relations que l'on traduit par des prédicats à deux places. Par exemple, la phrase Pierre frappe Paul évoque une scène, alors que la phrase Paul frappe Pierre évoque une autre scène. Une théorie de l'interface entre le langage, la perception et le raisonnement se doit d'indiquer comment chacune de ces deux phrases évoque systématiquement la bonne scène. Il est clair qu'un système purement associatif ne répond pas à la question. Faut-il développer un modèle radicalement différent du nôtre pour rendre compte des relations actantielles de ce type, ou peut-on envisager une extension de notre modèle qui couvre également ces relations ?

Les réflexions que nous avons eues sur cette question ne sont pas assez avancées pour donner une réponse argumentée. Nous donnons juste quelques indications sur le rôle que l'opérateur de contraste pourrait jouer dans la détermination des rôles sémantiques. Le point de départ est la capacité de repérer une configuration dans une scène donnée. Ainsi, nous savons reconnaître une scène de combat impliquant deux ou plusieurs personnes. Dans

¹⁴ Noter que l'opération de négation, en logique, est toujours une opération d'ordre supérieur, puisqu'elle porte sur des proposition ou des prédicats.

l'espace qualitatif, une telle scène est fortement structurée : les objets perçus bougent, adoptent des postures, interagissent avec l'environnement. La question est de savoir comment cette structure est traitée pour donner lieu à une expression d'une extrême simplicité comme *frapper* (X, Y), dans laquelle les objets évoquant les mots X et Y sont individualisés, ainsi que la relation qui les unit. Le problème n'est pas trivial, car la configuration est un rapport entre objets. Elle n'a pas de contrepartie tangible dans la scène et disparaît dès que l'un des objets qu'elle implique est ignoré. Notre conjecture est que c'est précisément cette propriété qui peut donner lieu à un contraste. L'individuation des actants est rendue possible par le contraste de la scène avec les actants et de la scène sans les actants. Dans les scènes correspondant à la phrase *Pierre frappe Paul* ou à celle de *Pierre regarde Marie*, le fait de supprimer Pierre de la scène fait disparaître la configuration ; de même si c'est Paul ou Marie que l'on exclut de la scène. Inversement, le fait de se focaliser à la fois sur Pierre et Paul ou sur Pierre et Marie provoque la perception de la configuration qui les unit. Ce mécanisme permet d'expliquer l'individuation des actants à partir d'une configuration, ou le repérage d'une configuration à partir des actants. En revanche, il ne permet pas d'expliquer que les actants puissent recevoir des rôles différents. Pourquoi, dans la phrase *Pierre traverse la rue*, est-ce Pierre qui traverse la rue et non la rue qui traverse Pierre ? Il faut que le mot *traverser* ait un lien d'évocation non seulement avec la configuration, considérée de manière holistique, mais aussi avec l'ordre dans lequel les actants sont pris en considération. Là encore, l'opérateur de contraste peut intervenir pour établir la dissymétrie. Il faut supposer que le contraste entre Pierre et Paul, dans lequel Pierre est la figure et Paul le fond, évoque le verbe *frapper*, ce que ne fait pas le contraste opposé.

Ces suggestions n'ont d'autre but que de montrer que l'approche par l'opérateur de contraste ne se limite pas au traitement de la modification. Elles permettent aussi de souligner la difficulté d'interfacer, de manière cognitivement plausible, le langage avec les aspects qualitatifs de la perception. Ces problèmes sont souvent ignorés dans les modèles de construction du sens, notamment lorsque ces modèles se contentent de considérer que les relations préexistent dans le monde et qu'il suffit de les lire pour les traduire sous forme de prédicat.

Perspectives générales

Le langage reste un phénomène dont la complexité résiste encore, par bien des aspects, à l'analyse. Il est peu probable que ses aspects sémantiques puissent être expliqués par un principe unique, que ce soit l'association, la combinaison d'images, la métaphore ou l'opérateur de contraste et les grilles conceptuelles. Néanmoins, l'idée de l'opérateur de contraste nous semble suffisamment féconde pour mériter d'être étendue à d'autres aspects du langage.

Une question intéressante à résoudre, dans le cadre de notre modèle, serait la coexistence de plusieurs systèmes de repérage parallèles. Rien n'indique que le repérage temporel, que nous avons décrit au début de ce travail (CF. CHAPITRE 2), puisse se mêler au repérage sémantique que nous avons brièvement décrit dans la présente section. Si ces deux systèmes de repérage opèrent en parallèle, comment s'effectue leur coordination ? Est-ce que le système de détermination constitue un troisième repérage en partie autonome ? L'existence de plusieurs systèmes de repérages plus ou moins indépendants, et néanmoins ressemblants, pose la question de leur origine. Une conjecture pourrait être que le système de détermination et le système de localisation temporelle soient tous deux dérivés du système de localisation spatiale, qui serait lui-même dérivé du système général des grilles conceptuelles. Toutes ces perspectives théoriques doivent être étudiées, mises à l'épreuve de la parcimonie et de la plausibilité cognitive, puis testées sur le plan psychologique.

9.6. Bilan

Le problème que nous avons abordé dans ce chapitre était de rendre plausible une interface entre l'espace qualitatif, le langage et le raisonnement. Pour cela, nous avons défini l'opérateur de contraste, ce qui nous a permis de comprendre comment l'opération de modification pouvait être ancrée dans l'espace qualitatif. L'opérateur de contraste permet également de ramener la complexité de l'espace qualitatif, dans un contexte donné, à une gradualité totalement ordonnée sur laquelle peuvent se projeter les grilles conceptuelles qui servent au repérage et au raisonnement.

La caractéristique essentielle de ce modèle est qu'il permet d'éviter les liens "durs" permanents que l'on trouve dans la plupart des modèles du système conceptuel. Les seuls liens permanents, dans notre modèle, sont des liens d'association. En particulier, les mots du lexique ne sont ni décrits, ni restreints par des contraintes relationnelles explicites. Ils sont associés à des représentations dans l'espace qualitatif. C'est le calcul sémantique, permis notamment par la synthèse des percepts, l'opérateur de contraste et l'application des grilles conceptuelles, qui établit l'interface entre ces représentations qualitatives, les relations syntaxiques et les opérations logiques, au premier rang desquelles la négation. De cette manière, on évite les différents paradoxes que nous avons dénoncés dans la partie précédente (CF. CHAPITRES 7 & 8). Certes, quelques mots particuliers, comme *avant*, *avec*, *trop*, la négation, la copule *et cætera* sont associés à des instructions précises qui guident la mise en œuvre du contraste et des grilles conceptuelles. L'éventail limité des instructions possibles fait que ces associations sont très stables, au point de pouvoir passer pour des liens "durs". Cependant, cette caractéristique ne concerne qu'un nombre limité de mots, de la classe dite fermée. En évitant les descriptions fixes, le modèle évite la plupart des difficultés théoriques liées à la polysémie (VICTORRI & FUCHS 1996 [108]). Au lieu d'apparaître comme un bruit affectant un système idéalement parfait, la polysémie devient une propriété nécessaire de la construction du sens. Les mots du lexique renvoient à des zones plus ou moins étendues de l'espace qualitatif. Ces zones se retrouvent transformées et réduites par le contexte et la combinaison linguistique, juste ce qui est nécessaire pour permettre aux interlocuteurs d'effectuer les inférences souhaitées. Il n'y a aucune idée de communication parfaite, de correspondance entre descriptions. La communication, considérée au niveau sémantique, n'est pas une affaire de codage - décodage (SPERBER & WILSON 1986 [101]). Les locuteurs réalisent des constructions plus ou moins précises sur les substrats généralement graduels que leur fournissent leurs espaces qualitatifs, à charge pour les interlocuteurs d'effectuer une construction similaire. Ce sont des critères pragmatiques qui permettent d'assurer une certaine efficacité dans la communication, non la conformité des mots employés à une sorte de dictionnaire de référence.

Ainsi, même si notre modèle n'existe, pour l'instant, qu'à l'état d'esquisse, il permet d'entrevoir comment des représentations ancrées dans la perception peuvent entrer dans des manipulations symboliques, ce qui offre un début de solution au problème de l'ancrage symbolique (HARNAD 1990 [46]).

Le renoncement à toute taxonomie sémantique et à toute forme de description symbolique n'est pas une idée nouvelle. Elle est présente dans les différentes traditions empiristes (CF. CHAPITRE 3). Ce que notre modèle propose de nouveau, c'est un moyen de concilier une approche dans laquelle les mots n'évoquent que des représentations de l'espace qualitatif avec la prédiction des phénomènes systématiques liés au langage et au raisonnement. Reprenons les phénomènes systématiques mentionnés au début de la partie précédente.

La substituabilité suppose que certains mots puissent être systématiquement remplacés par toute une classe d'autres mots au sein d'un énoncé, avec la garantie que l'énoncé conservera un sens. Ainsi, si le syntagme pomme mauve a un sens, on peut prédire que le syntagme pomme AC, où AC désigne un adjectif de couleur, aura systématiquement un sens. On considère généralement que cette forme de systématisme ne peut être strictement circonscrite que dans une approche symbolique de type moléculaire (CF. CHAPITRE 6). Si les adjectifs de couleur n'ont pas de structure conceptuelle, on ne peut les assigner à une classe sémantique, si bien que les contours de la substituabilité ne peuvent être délimités. Notre modèle, bien que se situant à l'interface entre le qualitatif et le symbolique, s'accommode de cette systématisme de substituabilité. Notre hypothèse est que les mots et les énoncés renvoient à des scènes dont la richesse structurelle est, potentiellement, celle de notre perception. Si les adjectifs de couleurs produisent des effets sémantiques similaires dans les combinaisons langagières, ce n'est pas en raison d'une prétendue structure conceptuelle qu'ils auraient en commun, mais parce que les représentations qualitatives auxquelles ils renvoient sont proches dans l'espace qualitatif. En conséquence, l'opérateur de synthèse qui permet la création du sens du syntagme modifié aura un effet à peu près similaire sur la représentation qualitative associée au mot pomme pour chaque adjectif de couleur qui lui est adjoint. En revanche, ce phénomène de systématisme de substitution disparaît pour un adjectif comme décontenancée qui renvoie à une représentation qualitative sans rapport avec celles des adjectifs de couleur. Noter que cette forme de systématisme est partielle, avec une gradation possible dans l'acceptabilité après substitution, ce qui est conforme à l'idée que la ressemblance entre scènes est une relation graduelle.

L'extraction constitue une autre forme de systématisme sémantique liée au langage. Celui qui peut penser l'énoncé Pierre V Marie est systématiquement capable d'avoir la pensée correspondant au mot Pierre. Il s'agit cette fois d'une systématisme stricte, qui ne souffre pas d'exceptions (si l'on exclut les expressions figées). Dans les modèles computationnels classiques, cette forme de systématisme est expliquée par la constituance des représentations (CF. CHAPITRE 6). Dans notre modèle, la prise en compte de la systématisme de l'extraction est assurée par le mécanisme qui forme les prédicats à une ou deux places. L'énoncé Pierre V Marie ne peut être émis ou compris que si la scène associée est, respectivement, analysée ou synthétisée. En d'autres termes, notre modèle reproduit la systématisme d'extraction parce qu'il respecte le principe de compositionnalité : le sens qu'il alloue à un composé est calculé à partir du sens qu'il donne aux composants. En revanche, notre modèle évite de postuler la constituance des représentations. Le syntagme pomme rouge, émise à la vue d'un fruit particulier, est produite parce que l'image du fruit, par contraste, est analysée comme un prototype modifié de pomme. L'image du prototype de pomme ne fait pas partie intégrante de l'image du fruit, mais elle intervient nécessairement dans le processus ayant conduit à la production du syntagme pomme rouge. Donc, la personne capable de former la pensée correspondant au syntagme pomme rouge est systématiquement capable de former la pensée correspondant au mot pomme.

Une autre forme de systématisme concerne les relations logiques, notamment la négation. Tout concept peut être nié. Cette propriété, qui est évidente pour les systèmes symboliques, ne l'est pas du tout pour les systèmes empiristes. L'un des atouts de notre modèle est de prédire cette systématisme, comme nous l'avons vu plus haut, grâce à l'emploi du contraste et des grilles conceptuelles.

Le modèle décrit dans ce chapitre a pour ambition, à terme, d'unifier un certain nombre de phénomènes sémantiques. Le même mécanisme de contraste et de grilles est censé expliquer la modification, la temporalité et l'interface avec le raisonnement, ainsi, peut-être, que la formation des relations prédictive à deux places, la détermination et la récursivité.

Cette unification est obtenue sans postuler l'existence de structures complexes associées aux mots du lexique, autres que celles qui interviennent déjà dans la perception et l'imagerie.

En retour, notre modèle permet de comprendre le constant désir, sans cesse déçu, éprouvé par les penseurs depuis l'antiquité, de réifier la signification des mots sous forme de descriptions conceptuelles. Selon nous, la permanence des descriptions conceptuelles dans notre esprit est une illusion. Dans notre modèle, seuls persistent à long terme les liens d'association. Certaines "descriptions" peuvent être ainsi stockées sous forme purement linguistique et non sémantique, au même titre que l'adage le carré de l'hypoténuse est égal à la somme des carrés des deux autres côtés, les comptines, l'opération cinq fois trois quinze ou la suite lundi, mardi, mercredi. Pour tout le reste, nous formons nos descriptions à la volée, à partir de nos liens associatifs. Lorsque nous "retrouvons" la définition de ce qu'est un groupe en mathématique, nous utilisons toutes les ressources de notre imagerie et de notre capacité à contraster pour retrouver les éléments de la définition officielle. Par exemple, nous revoyons le e désignant l'élément neutre ou le -1 désignant l'élément inverse, tels que nous avons l'habitude de les écrire dans le contexte où nous manipulons des groupes. Ces associations sont censées suffire, grâce aux mécanismes du modèle, pour retrouver une description à laquelle nous pouvons conférer une certaine nécessité. Si les concepts de groupe, d'anneau, de corps ou d'algèbre étaient stockés sous forme de définition dans notre mémoire, nous pourrions reproduire ces définitions instantanément. Ce que nous produisons est une reconstruction. C'est la nécessité que nous conférons à cette reconstruction qui nous donne l'illusion d'une existence permanente. Le même constat peut être fait à propos de tout concept. Nous avons l'impression que le concept de semaine peut être décrit par les jours qui la composent. Or, le lien permanent entre ce qu'évoque le mot semaine et les mots lundi, mardi, *et cætera* n'est qu'un lien d'association, renforcé par l'observation de milliers d'occurrences. Si quelqu'un essaie de nous persuader que mercredi vient avant mardi, nous lui opposons une nécessité issue de raisonnements réalisés à la volée, comme le fait que mercredi suit jeudi sur l'agenda qui se trouve sur la table. Le fait même que nous soyons capables d'argumenter démontre que la description est construite, et n'est pas constitutive du concept.

Conclusion

L'objectif de ce chapitre était d'indiquer les grandes lignes d'un modèle dans lequel la construction du sens s'effectue sur des représentations de même nature que les données de l'expérience qualitative. L'objectif n'est, certainement, pas nouveau. En revanche, nous avons inclus dans le cahier des charges la reproduction des aspects systématiques et tout-ou-rien du sens des énoncés, ce qui est plus rare dans les approches qui ancrent le sens dans l'expérience. Notre modèle se veut original, également, par les solutions qu'il propose. Nous expliquons comment des représentations, qui peuvent être vues comme symboliques, par exemple des prédicats logiques, sont produites à la volée, sur un substrat qualitatif qui, par nature, est multidimensionnel et graduel. Les notions théoriques qui sous-tendent notre modèle sont, principalement, l'opérateur de contraste et le mécanisme de projection des grilles conceptuelles. Nous nous sommes surtout concentrée sur l'opération linguistique de modification et l'opération logique de négation. Cependant, nous avons essayé de montrer que la portée de ces outils théoriques était beaucoup plus large. Notre espoir est que l'ensemble du champ de la sémantique pourra être atteint avec des mécanismes généraux et simples comme ceux que nous avons suggéré, ce qui permettra de concevoir une authentique construction du sens des énoncés, ancrée sur les données de l'expérience qualitative.

Conclusion

Cette dernière partie de ce texte explorait l'hypothèse d'un système conceptuel autonome, doté de ses propres représentations permanentes et entretenant des interface avec la perception, le langage et raisonnement. Nous nous sommes efforcée de montrer qu'une telle hypothèse conduisait à des conséquences absurdes. Dans cette quatrième partie de notre travail, nous avons exploré une hypothèse radicalement différente. Le système conceptuel serait un système minimal, une simple interface, quasiment dépourvu de représentations permanentes propres. Pour assurer l'interface entre trois systèmes *a priori* incompatibles, le système linguistique, le système logique et le système perceptif, nous avons postulé l'existence de mécanismes, produisant des représentations transitoires que sont les cartes conceptuelles.

L'enjeu de ce changement radical d'hypothèse est de rechercher une authentique construction du sens des énoncés. La seule chose qu'un système conceptuel hébergeant une collection de concepts permanents peut offrir est une traduction d'une langue naturelle dans une langue mentale qui, malgré bien des efforts, demeure totalement insaisissable. En renvoyant une partie du sens dans l'espace qualitatif, en supposant que l'interprétation d'un énoncé passe par la constitution de scènes concrètes ou abstraites, nous avons le sentiment de rendre à la question de la construction du sens sa dimension cognitive. L'interprétation n'est pas confiée à un oracle externe ou au câblage préétabli d'un réseau de concepts opportunément présent dans notre esprit. En proposant une interface à base de cartes conceptuelles, dont les représentations sont transitoires, nous avons le sentiment de rendre l'hypothèse de la construction dynamique du sens crédible.

Assurément, le modèle présenté dans cette partie reste à l'état d'ébauche. Notre objectif était surtout d'éviter le reproche que l'on oppose aux modèles ancrés, à propos du mystère qui entoure généralement l'explication des aspects systématiques et symboliques de la signification. En proposant quelques mécanismes originaux, simples et de portée générale, nous avons cherché à démystifier cette question, en nous efforçant de réconcilier une approche qualitative du sens avec son utilisation symbolique dans le langage et le raisonnement. Nous sommes convaincue qu'à terme, des systèmes mécaniques pourront utiliser des processus comme ceux que nous avons proposés pour construire une représentation sémantique leur permettant de produire des énoncés et des raisonnements similaires à ceux des humains.

Conclusion

Dans ce travail, nous avons cherché à explorer les conséquences d'une hypothèse, présente dans de nombreuses théories de la cognition, d'un système conceptuel autonome censé héberger les représentations symboliques associées au sens des expressions langagières. Nous avons essayé de montrer que, dans la mesure où un tel système contient des représentations conceptuelles permanentes, la construction du sens des énoncés couramment utilisés dans une langue quelconque ne pouvait pas être assurée d'une manière qui soit cognitivement plausible. Par exemple, la représentation des relations temporelles ne peut être réalisée à l'aide d'une structure globale contenant l'ensemble des instants ou des intervalles que nous pouvons concevoir. Plus généralement, il est impossible d'enfermer la totalité des relations sémantiques qui peuvent être potentiellement exprimées par le langage dans un ensemble fixe et permanent de représentations conceptuelles. Il ne s'agit pas d'une impossibilité théorique : les contraintes de la modélisation cognitive sont simplement incompatibles avec le postulat des structures et des procédures qu'un système conceptuel à représentations permanentes exigerait.

D'un autre côté, renoncer à toute idée de représentation symbolique pour expliquer l'interprétation des énoncés langagiers n'est pas satisfaisant. Si, comme dans une certaine tradition empiriste, les représentations sémantiques ne sont pas qualitativement différentes des perceptions, il devient impossible d'expliquer les aspects systématiques liés à la combinaison des mots et aux inférences qu'ils occasionnent.

Nous avons donc opté pour une troisième voie. Les représentations symboliques associées aux expressions langagières, dans le modèle que nous proposons, sont des représentations éphémères. Grâce à notre opérateur de contraste et à un inventaire limité de grilles, nous sommes en mesure d'indiquer, au moins pour certaines opérations cognitives comme la modification ou la négation, comment la procédure de construction du sens peut rendre de telles opérations systématiques.

La démarche qui a été la nôtre demande à être complétée sur le plan théorique. Il s'agit d'enrichir le modèle pour le rendre propre à reproduire les différentes formes de relations sémantiques, à commencer par les relations prédicatives à deux arguments et la détermination. Comment l'opérateur de contraste doit-il être modifié pour produire de telles relations ? Un autre pan de l'étude consistera à proposer des implémentations de l'opérateur de contraste, en simulant l'espace des qualités perceptives par un espace de grande dimension.

Notre modèle présente l'avantage de se prêter à la réfutation potentielle, dans la mesure où il effectue des prédictions. Par exemple, une mise à l'épreuve consiste à faire l'inventaire des types de relations prédicatives exprimés par le langage pour voir si l'inventaire limité des grilles permet de les reproduire. En d'autres termes, il s'agit de voir si le modèle offre une description parcimonieuse et cognitivement plausible des différentes relations sémantiques qui ont été identifiées dans les différentes langues. Cette ré-interprétation des résultats linguistiques pourra conduire à des révisions du modèle. Notre espoir est que ces révisions ne lui ôteront pas sa simplicité et sa plausibilité. Toujours sur le plan linguistique, il est également important de voir comment le modèle s'interface avec d'autres modèles, notamment les modèles syntaxiques et pragmatiques. Le renoncement à des échafaudages conceptuels censés représenter le sens des expressions composées, au profit d'une procédure de construction du sens, suppose que cette procédure puisse être étroitement guidée par la syntaxe. En particulier, le caractère récursif de la procédure doit correspondre à la récursivité des structures syntaxiques. De même, les opérations de modification et de détermination rendues possibles par l'opérateur de contraste et les grilles sont censées suffire pour satisfaire

les exigences pragmatiques, comme la nécessité de rendre une désignation suffisamment non ambiguë pour l'interlocuteur.

Notre modèle présente aussi l'avantage de pouvoir être testé sur le plan psychologique. Il s'agit de trouver, dans le comportement des sujets en situation de produire ou d'interpréter des énoncés langagiers, des indices de la manipulation des grilles et de l'emploi du contraste. Par exemple, nous avons supposé que dans l'application récursive des grilles, "une grille chassait l'autre", autrement dit que deux grilles ne pouvaient pas coexister sur la même carte. Il s'agit là d'une exigence forte qui doit avoir des conséquences mesurables, par exemple si l'on demande aux sujets de commenter le rapport entre deux éléments appartenant à des cartes différentes, quoique emboîtées.

Commentons enfin le fait que notre modèle est susceptible de mener à un certain nombre d'applications pratiques. L'une des limitations des systèmes de traitement du langage naturel, au niveau sémantique, est liée au caractère strict et circonscrit des ontologies employées. Le problème présente un enjeu non négligeable à l'heure où il est question d'offrir une description sémantique des contenus mis à disposition sur la toile. Si, comme nous l'espérons, notre modèle peut conduire à une interprétation souple des mots et des syntagmes en contexte, nous pouvons espérer reproduire des phénomènes comme les glissements de sens et certaines métaphores sans lesquels l'interaction humain - machine en langue naturelle restera cantonnée à des emplois utilitaristes bien délimités. Si l'on parvient à placer l'ensemble des signifiés potentiels dans un espace qualitatif doté de propriétés géométriques, les opérations de modification et de négation pourront être effectuées selon les principes décrits dans le modèle proposé, même lorsque les mots précis font défaut.

Le travail exposé dans ce document comporte une importante partie critique, dont le but était de souligner les limitations des modèles existants sur le plan cognitif. Cette critique nous a permis de motiver la recherche d'un modèle plus plausible. Pour ce faire, nous avons dû renoncer à un présupposé essentiel posé en fondement de la plupart des théories sémantiques : l'existence des concepts en tant qu'entités symboliques permanentes. Ce renoncement nous a été suggéré par la recherche d'un moyen cognitivement plausible pour représenter les relations temporelles exprimées par le langage. De même que l'hypothèse d'une structure temporelle permanente n'est pas tenable, l'existence d'un réseau de connaissances conceptuelles symboliques se révèle extrêmement problématique. En proposant le fait que les représentations symboliques que nous associons aux énoncés du langage sont des représentations éphémères, nous avons été conduit à présenter la permanence des descriptions comme une illusion cognitive. Les descriptions sont produites "à la volée" à des fins argumentatives. Le sentiment qu'elles possèdent une permanence dans notre esprit vient du fait que nous pouvons, dans la plupart des cas, les reconstituer facilement. Si l'on exclut le cas des paraphrases verbales apprises par cœur, les descriptions que nous produisons ne résultent pas d'une simple lecture de notre mémoire, mais sont le résultat d'un calcul. L'image de notre cognition qui est ainsi offerte se trouve bouleversée. Il ne s'agit plus d'imaginer l'esprit humain comme une encyclopédie bien pleine, mais comme un système de calcul efficace. Nous espérons que ce changement de point de vue sera fécond et inspirera d'autres recherches.

Annexe :
**Contraintes d'un modèle
d'apprentissage de concepts**

Les paradigmes d'apprentissage

À plusieurs reprises dans le texte de la thèse, nous faisons référence à des méthodes ou à des paradigmes d'apprentissage, tels qu'ils sont proposés dans le domaine de l'apprentissage automatique¹. Dans cette annexe, nous passons en revue quelques modèles d'apprentissage, avant de les opposer sur la question de l'innéité du biais qu'ils postulent.

L'apprentissage statistique

Les modèles statistiques de l'apprentissage reposent, pour l'essentiel, sur un mécanisme d'extraction de régularités. L'effet de l'apprentissage est de conserver les formes fréquentes en écartant les variations contingentes. Un certain nombre de modèles entrent dans cette catégorie. Citons-en quelques-uns.

- la théorie behavioriste : couplage de stimuli ou association stimulus - réponse.
- les réseaux de neurones artificiels : établissement d'un couplage continu entre une entrée et une sortie, en respectant, dans le cas supervisé, certaines associations imposées.
- l'apprentissage par renforcement : acquisition d'une fonction de coût permettant de mémoriser les meilleures actions dans chaque état.
- les systèmes de regroupement (*clustering*).

Commentons les systèmes de regroupement, car ils sont fréquemment invoqués pour expliquer l'opération de catégorisation perceptuelle, voire conceptuelle. Le principe général des techniques de regroupement consiste à constituer des classes d'objets à partir d'une mesure de similarité, de manière à minimiser la variance à l'intérieur des classes et à maximiser la variance entre les classes. La méthode des nuées dynamiques est un exemple simple d'algorithme de regroupement, dans le cas où la mesure de ressemblance est donnée par une métrique. Elle consiste à choisir n objets au hasard, qui consistent la première nuée. On fabrique n classes initiales en regroupant chaque objet avec l'élément de la nuée le plus proche. On forme ensuite une nouvelle nuée en prenant les objets les plus proches des centres de gravité des classes. On itère le processus jusqu'à ce que les classes obtenues soient stables. Les éléments de la dernière nuée obtenue peuvent être interprétés comme des prototypes.

Certains systèmes de regroupement statistiques sont de type symbolique. La distance se mesure au nombre d'attributs symboliques que les objets ont en commun. La technique de regroupement reste statistique, mais elle est astreinte à produire des classes qui possèdent une caractérisation simple. Les systèmes de type SBL (*similarity based learning*) fonctionnent typiquement en produisant des classes correspondant à des conjonctions d'attributs.

Les modèles statistiques d'apprentissage nécessitent généralement un grand nombre d'exemples pour pouvoir fonctionner. En cela, ils diffèrent des autres paradigmes d'apprentissage que nous considérons maintenant.

L'apprentissage gestaltiste

Les modèles gestaltistes de l'apprentissage privilégient l'accès aux bonnes formes. Une bonne forme est une forme, concrète ou abstraite, qui reste invariante pour de nombreuses transformations. Dans le domaine perceptif, un cercle ou un carré sont des bonnes formes ;

¹ Le lecteur pourra se reporter à (MITCHELL 1997 [74]) pour une revue des principales techniques d'apprentissage automatique.

elles sont plus symétriques qu'un patatoïde informe et qu'un polygone irrégulier. Pour un mécanisme d'apprentissage gestaltiste, les bonnes formes sont plus faciles à apprendre. On peut même considérer que le résultat de tout apprentissage est une bonne forme.

Il est intéressant de rapprocher la théorie constructiviste (PIAGET 1932 [82]) d'un système gestaltiste. Dans cette théorie, les stades de développement de l'enfant sont caractérisés par l'accès à des structures invariantes pour des groupes de transformations. Ainsi, l'enfant accède à une première version du concept de justice en considérant que toute action qui est bonne envers lui est juste. Ce concept est invariant pour tout changement d'acteur. Lorsqu'il devient capable de se décentrer, l'enfant est en mesure de substituer d'autres patients aux actions. En accédant à ces nouvelles transformations, il forme un nouveau concept de justice, invariant pour toute substitution de personnes dans les rôles d'acteur et de patient : est juste une action qu'il ressentirait comme bonne si elle était dirigée vers lui. La théorie constructiviste explique ainsi que différents enfants, confrontés à des expériences limitées, forment les mêmes concepts de justice (PIAGET 1932 [82]).

L'appariement combinatoire

Le paradigme de l'appariement regroupe un certain nombre de modèles et de techniques, caractérisés par l'existence d'une structure préalable riche. Mentionnons-en quelques-uns.

- la théorie des principes et paramètres.
- l'induction symbolique.
- l'apprentissage par analogie.
- les systèmes sélectifs.

La théorie dite des principes et paramètres a été invoquée pour expliquer l'acquisition par l'enfant de la grammaire de sa langue, sachant qu'il dispose d'un patron de grammaire universel. L'apprentissage se limite à instancier des paramètres binaires du patron au vu de quelques exemples de structures de phrases. Ainsi, un enfant iranien apprendra, en analysant quelques phrases dépourvues de sujet grammatical explicite, que pour le persan, le paramètre *pro-drop* doit être positionné à la valeur 1, alors qu'un enfant français, en observant des sujets explétifs, fixera le même paramètre à 0. Ce type d'apprentissage se rapproche du phénomène d'empreinte mis en évidence en éthologie.

Les systèmes d'induction symbolique cherchent à généraliser les exemples qui leur sont soumis. Par exemple, un programme de type ILP (*inductive logic programming*) à qui l'on présente deux instances du concept MIGNON, l'un qui est un chat en peluche, l'autre qui est un petit chien en peluche, pourra former la moins générale des généralisations, à savoir qu'un moyen d'être MIGNON est d'être un petit animal en peluche. Pour parvenir à ce résultat, le programme utilise une connaissance d'arrière-plan qui lui dit que chiens et chats sont des animaux, et qu'un chat est petit (sans cette connaissance d'arrière-plan, la conclusion aurait été que toute peluche est mignonne, ce qui est excessivement général). La technique utilisée par les programmes ILP est la "résolution inverse". La méthode consiste à partir d'un exemple et d'une règle de la connaissance d'arrière-plan, puis à former une nouvelle règle qui, enchaînée déductivement avec la règle de départ, produit l'exemple.

Les systèmes de type EBL (*explanation based learning*), utilisent également une connaissance d'arrière-plan. Ils mêlent les aspects inductifs et déductifs pour apprendre des concepts qui sont à la fois des généralisations des exemples qui leur sont soumis et des spécialisations des concepts qu'ils connaissent déjà.

Les systèmes comme ILP ou EBL sont capables de produire des représentations structurées. La comparaison entre structures est à la base de l'apprentissage par analogie. Il s'agit d'une forme d'apprentissage dans laquelle on s'intéresse plus aux rapports entre exemples qu'aux exemples eux-mêmes. Il est alors possible d'engendrer de nouveaux objets qui entretiennent un rapport connu avec un objet donné, selon les techniques du raisonnement par cas (*case-based reasoning*).

Les systèmes d'apprentissage sélectifs utilisent une mesure de la distance entre les formes qu'ils produisent et la forme à obtenir. Ils engendrent de nouvelles formes en ré-assemblant et en modifiant des formes existantes, jusqu'à approcher suffisamment la forme cible. Les algorithmes génétiques offrent un exemple de mécanisme sélectif. Un autre exemple connu d'apprentissage sélectif est donné par la modélisation du système immunitaire. Une population de lymphocytes est pré-assemblée par réarrangement d'un jeu d'éléments génétiques donnés. L'apprentissage d'une molécule étrangère inconnue se fait par sélection du lymphocyte qui présente la plus grande affinité avec la molécule. Le lymphocyte en question se multiplie, engendrant un clone lymphocytaire qui peut perdurer pendant des années. Des mécanismes sélectifs du même type ont pu être proposés pour l'apprentissage conceptuel (CHANGEUX & DEHAENE 1989 [13]).

Tous ces systèmes, que ce soit la théorie des principes et paramètres, l'induction symbolique, l'analogie ou l'apprentissage sélectif, fonctionnent par appariement entre une forme qui sert de stimulus et une forme interne au système. Cette dernière peut résulter d'un assemblage combinatoire, constitué à partir de briques de base connues dès le départ par le système. Dans le cas des systèmes sélectifs, le hasard est seul responsable de cet assemblage. Dans d'autres formes d'appariement combinatoire, l'assemblage est dirigé pour être en adéquation avec le stimulus.

Les modèles d'apprentissage par appariement sont remarquables par le fait qu'ils sont capables d'apprendre de nouvelles formes, de manière non triviale, en une fois. Contrairement aux modèles statistiques, qui nécessitent un échantillon conséquent d'instances de la forme à apprendre, les systèmes d'appariement combinatoire peuvent se satisfaire d'une exposition unique. Cette prouesse est obtenue grâce à la connaissance qui est déjà présente dans le système avant que l'apprentissage ait lieu.

De la nature des biais d'apprentissage

Les théories de l'apprentissage, en sciences cognitives, ont souvent été opposées du point de vue des hypothèses qu'elles faisaient à propos des connaissances *a priori* prêtées au sujet apprenant, autrement dit le biais d'apprentissage. Dans le cas du langage, on a pu opposer les théories behavioristes, les théories constructivistes et les théories innéistes (PIATTELLI-PALMARINI 1979 [84]). Nous mentionnons ici une classification qui s'attache davantage aux propriétés de symétrie du biais qu'à une simple mesure quantitative de ce biais. L'intérêt de cette classification pour les sciences cognitives est de permettre de caractériser le type d'apprentissage à l'œuvre à partir des propriétés des formes apprises.

Certains modèles d'apprentissage possèdent les propriétés d'isotropie et de relativité. Un mécanisme d'apprentissage est isotrope (respectivement relatif) si le fait de faire subir une rotation (respectivement une translation) à toutes ses entrées, ce qui revient à un changement de repère, ne modifie pas son comportement global après apprentissage (DESSALLES 1998 [29]). Ces propriétés signifient que le système d'apprentissage n'utilise au plus que les distances relatives entre les données. En d'autres termes, le système ne possède aucune connaissance absolue lui permettant de privilégier une direction particulière dans les données auxquelles il est exposé.

Il est intéressant de constater que, parmi les modèles couramment proposés pour rendre compte de l'apprentissage animal et humain, ces propriétés d'isotropie et de relativité distinguent clairement les systèmes statistiques et les systèmes gestaltistes d'un côté, des systèmes d'appariement combinatoire de l'autre (DESSALLES 1998 [29]). Pour prendre un exemple, un perceptron multi-couches est insensible à une permutation systématique de ses neurones d'entrée (rotation) ou à une complémentation partielle des signaux qui lui sont présentés (translation). La classification acquise par le système est rigoureusement la même lorsque les données d'apprentissage et les données de test subissent la même transformation systématique, sachant qu'il s'agit d'une isométrie (rotation ou translation) n'affectant pas la distance, et donc la ressemblance, entre les données.

Le résultat principal concernant l'isotropie des mécanismes d'apprentissage est qu'un système isotrope et relatif est nécessairement astreint à une forme de gestaltisme (DESSALLES 1998 [29]). Plus précisément, pour une classification apprise par un tel système, le produit de son harmonie (nombre d'isométries laissant les classes invariantes) par sa variété (nombre de classifications différentes qui auraient été obtenues sur des données transformées) est constant. La conséquence de ce résultat est importante pour la modélisation cognitive. Le tableau qui suit résume l'idée principale : il compare les trois paradigmes d'apprentissage mentionnés ci-dessus au regard de trois propriétés, la convergence des résultats de l'apprentissage, l'isotropie et la relativité du mécanisme considéré, et l'harmonie des formes apprises.

	<i>convergence</i>	<i>isotropie/relativité</i>	<i>harmonie</i>
statistique	non	oui	non
gestaltiste	oui	oui	oui
Appariement	oui	non	non

La convergence est la propriété selon laquelle des individus confrontés à des données différentes apprennent les mêmes formes. Cette propriété distingue d'un côté les systèmes statistiques, qui peuvent apprendre toute forme de régularité, et d'un autre côté les systèmes gestaltistes, dans lesquelles les formes apprises se comportent comme des attracteurs, et les systèmes par appariement, capables de produire des résultats fiables dans des environnements variés (il s'agit de l'argument classique de la pauvreté du stimulus). Certaines formes d'apprentissage conceptuel sont considérées comme convergentes. Ainsi, les individus d'une culture donnée acquièrent des concepts comparables, ce qui leur permet de communiquer. Lorsque l'on s'intéresse aux catégories fondamentales de la cognition que tout individu humain est supposé posséder, et qu'on les considère comme acquises, la convergence est encore plus manifeste.

Le résultat théorique mentionné plus haut contraint tout modèle de l'apprentissage cognitif à choisir entre les trois lignes du tableau. Or ce choix est fortement contraint. Si l'on s'intéresse à des acquisitions comme celle du concept de justice, il est parfaitement possible d'opter pour un système gestaltiste, puisque la forme apprise présente un certain nombre de symétries. En revanche, si la forme apprise ne peut être considérée harmonieuse, *id est* comme une bonne forme, ce qui est sans doute le cas pour la plupart des acquisitions conceptuelles quotidiennes (rappelons qu'un enfant acquiert pendant plusieurs années une dizaine de mots nouveaux par jour, avec les concepts associés), il faut choisir entre un apprentissage statistique et un mécanisme d'appariement. Or, le choix statistique, en l'absence de la propriété de convergence, n'est acceptable que si l'on est en mesure de montrer que les différents individus ont été confrontés aux mêmes expériences. Si l'on considère une telle contrainte comme inacceptable, la seule option qui demeure ouverte est celle de l'appariement combinatoire.

Bibliographie

1. SHARON L ARMSTRONG & LILA R GLEITMAN & HENRY GLEITMAN, 1983, "What Some Concepts Might Not Be", *Cognition*, 13: 263-308.
2. ALAN D BADDELEY, 1992, "Working Memory: the interface between memory and cognition", *Journal of Cognitive Neurosciences*, 4: 281-288.
3. LAWRENCE W BARSALOU, 1999, "Perceptual Symbol Systems", *Behavioral and Brain Sciences*, 22: 577-609.
4. JON BARWISE & JOHN PERRY, 1983, *Situations and Attitudes*, MIT Press.
5. DEREK BICKERTON, 1990, *Language and Species*, University of Chicago Press.
6. DEREK BICKERTON, 1995, *Language and Human Behavior*, University of Washington Press.
7. NED BLOCK, 1986, "Advertisement for a Semantics for Psychology", *Midwest Studies in Philosophy*, 10: 615-678.
8. SUSAN CAREY, 1985, *Conceptual Change in Childhood*, MIT Press.
9. RUDOLF CARNAP, 1947, *Meaning and Necessity: a study in semantics and modal logic*, University of Chicago Press.
10. BOB CARPENTER, 1997, *Type-Logical Semantics*, MIT Press.
11. PETER M CARRUTHERS, 1996, *Language, Thought and Consciousness: an essay in philosophical psychology*, Cambridge University Press.
12. DAVID J CHALMERS, 1993, "Connectionism and Compositionality: why Fodor and Pylyshyn were wrong", *Philosophical Psychology*, 6: 305-319.
13. JEAN-PIERRE CHANGEUX & STANISLAS DEHAENE, 1989, "Neuronal Models of Cognitive Functions", *Cognition*, 33: 63-109.
14. NOAM A CHOMSKY, 1972, *Studies on Semantics in Generative Grammar*, Mouton.
15. NOAM A CHOMSKY, 1975, *Reflections on Language*, Pantheon.
16. NOAM A CHOMSKY, 1981, *Lectures on Government and Binding Theory*, Foris.
17. NOAM A CHOMSKY, 1995, *The Minimalist Program*, MIT Press.
18. NOAM A CHOMSKY, 2000, *New Horizons in the Study of Language and Mind*, Cambridge University Press.
19. PAUL M CHURCHLAND, 1989, *A Neurocomputational Perspective: the nature of mind and the structure of science*, MIT Press.
20. AUSTEN CLARK, 1993, *Sensory Qualities*, Oxford University Press.
21. BERNARD COMRIE, 1976, *Aspect*, Cambridge University Press.
22. BERNARD COMRIE, 1985, *Tense*, Cambridge University Press.
23. FRANCIS CRICK & CHRISTOF KOCH, 1990, "Towards a Neurobiological Theory of Consciousness", *The Neurosciences*, 2: 263-275.
24. DONALD DAVIDSON, 1980, *Essays on Actions and Events*, Oxford University Press.

25. DONALD DAVIDSON, 1984, *Inquiries into Truth and Interpretation*, Oxford University Press.
26. MICHEL DE GLAS, 1992, "A Local Intensional Logic", *International Conference on Algebraic Logic and their Computer Science Applications*.
27. JEAN-PIERRE DESCLÈS, 1989, "State, Event, Process, and Topology", *General Linguistics*, 29: 159-200.
28. JEAN-LOUIS DESSALLES, 1993, *Modèle cognitif de la communication spontanée, appliqué à l'apprentissage des concepts*, Thèse de doctorat : ENST.
29. JEAN-LOUIS DESSALLES, 1998, "Limits of Isotropic Bias in Natural and Artificial Models of Learning", In G RITSCHARD & A BERCHTOLD & F DUC (Eds.), *Apprentissage: des principes naturels aux méthodes artificielles*, 307-319, Hermès.
30. JEAN-LOUIS DESSALLES, 2000, *Aux origines du langage : une histoire naturelle de la parole*, Hermès.
31. FRED I DRETSKE, 1981, *Knowledge and the Flow of Information*, MIT Press.
32. JEFFREY L ELMAN & ELIZABETH A BATES & MARK H JOHNSON & DOMENICO PARISI, 1996, *Rethinking Innateness*, MIT Press.
33. GILLES FAUCONNIER, 1997, *Mappings in Thought and Language*, Cambridge University Press.
34. JERRY A FODOR, 1975, *The Language of Thought*, Crowell.
35. JERRY A FODOR, 1990, *A Theory of Content and Other Essays*, MIT Press.
36. JERRY A FODOR, 1994, "Concepts: a potboiler", *Cognition*, 50: 95-113.
37. JERRY A FODOR, 1998, *Concepts: where cognitive science went wrong*, Oxford University Press.
38. JERRY A FODOR & MERRILL F GARRETT & EDWARD C T WALKER & CORNELIA H PARKES, 1980, "Against Definitions", *Cognition*, 8: 263-367.
39. JERRY A FODOR & ERNEST LEPORE, 1992, *Holism: a shopper's guide*, Blackwell.
40. JERRY A FODOR & BRIAN P MCLAUGHLIN, 1990, "Connectionism and the Problem of systematicity: why Smolensky's solution doesn't work", *Cognition*, 35: 183-204.
41. JERRY A FODOR & ZENON W PYLYCHYN, 1988, "Connectionism and Cognitive Architecture: a critical analysis", *Cognition*, 28: 3-71.
42. PETER GÄRDENFORS, 2000, *Conceptual Spaces: the geometry of thought*, MIT Press.
43. ALISON M GOPNIK & ANDREW N MELTZOFF, 1997, *Words, Thoughts, and Theories*, MIT Press.
44. LAURANT GOSSELIN, 1996, *Sémantique de la temporalité en français : un modèle calculatoire et cognitif du temps et de l'aspect*, Duculot.
45. STEPHEN J GOULD & RICHARD C LEWONTIN, 1979, "The Spandrels of San Marco and the Panglossian Program: a critique of the adaptationist program", *Proceedings of the Royal Society of London*, 205: 281-288.
46. STEVEN R HARNAD, 1990, "The Symbol Grounding Problem", *Physica D*, 42: 335-346.
47. DOUGLAS R HOFSTADTER, 1995, *Fluid Concepts and Creative Analogies: computer models of the fundamental mechanisms of thought*, Basic Books.

48. JOHN J HOPFIELD, 1982, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", *Proceedings of the National Academy of Sciences*, 79: 2554-2558.
49. RAY JACKENDOFF, 1983, *Semantics and Cognition*, MIT Press.
50. RAY JACKENDOFF, 1990, *Semantic Structures*, MIT Press.
51. FRANK JACKSON, 1982, "Epiphenomenal Qualia", *Philosophical Quarterly*, 32: 127-136.
52. MARCIA K JOHNSON, 1988, "Reality Monitoring: an experimental phenomenological approach", *Journal of Experimental Psychology: general*, 117: 390-394.
53. MARCIA K JOHNSON & MARY A FOLEY & AURORA G SUENGAS & CAROL L RAYE, 1988, "Phenomenal Characteristics of Memories for Perceived and Imagined Autobiographical Events", *Journal of Experimental Psychology: general*, 117: 370-376.
54. PHILIP N JOHNSON-LAIRD, 1977, "Procedural Semantics", *Cognition*, 5: 189-214.
55. HANS KAMP & BARBARA H PARTEE, 1995, "Prototype Theory and Compositionality", *Cognition*, 57: 129-191.
56. HANS KAMP & UWE REYLE, 1993, *From Discourse to Logic: introduction to modeltheoric semantics of natural language, formal logic and discourse representation theory*, Kluwer.
57. FRÉDÉRIC KAPLAN, 2001, *La naissance d'une langue chez les robots*, Hermès.
58. JERROLD J KATZ, 1972, *Semantic Theory*, Harper and Row.
59. DANIEL KAYSER, 1997, *La représentation des connaissances*, Hermès.
60. FRANK KEIL, 1989, *Concepts, Kinds, and Cognitive Development*, MIT Press.
61. STEPHEN M KOSSLYN, 1994, *Image and Brain*, MIT Press.
62. SAUL A KRIPKE, 1972, *Naming and Necessity*, Harvard University Press.
63. GEORGE LAKOFF, 1990, "The Invariance Hypothesis: is abstract reason based on image-schemes?", *Cognitive Linguistics*, 1: 39-74.
64. RONALD W LANGACKER, 1987, *Foundations of Cognitive Grammar: volume I, theoretical prerequisites*, Stanford University Press.
65. DOUGLAS B LENAT & RAMANATHAN V GUHA, 1990, *Building Large Knowledge Based Systems: representation and inference in the Cyc project*, Addison-Wesley.
66. BETH LEVIN, 1993, *English Verb Classes and Alternations: a preliminary investigation*, University of Chicago Press.
67. MICHAEL LEYTON, 1992, *Symmetry, Causality, Mind*, MIT Press.
68. ERIC MARGOLIS, 1998, "How to Acquire a Concept", *Mind and Language*, 13: 347-369.
69. ERIC MARGOLIS, 1999, "What is Conceptual Glue? ", *Minds and Machines*, 9: 241-255.
70. ELLEN M MARKMAN, 1990, "Constraints Children Place on Word Meanings", *Cognitive Science*, 14: 57-77.
71. JOHN MCCARTHY 1963, "Situations, Actions and Causal Laws", *Stanford Artificial Intelligence Project, Memo 2*.

72. DAVID MCNEILL, 1992, *Hand and Mind: what gestures reveal about thought*, University of Chicago Press.
73. RUTH G MILLIKAN, 1984, *Language, Thought, and other Biological Categories: new foundations for realism*, MIT Press.
74. TOM M MITCHELL, 1997, *Machine Learning*, McGraw-Hill.
75. JACQUES MONOD, 1970, *Le hasard et la nécessité*, Seuil.
76. RICHARD MONTAGUE, 1974, *Formal Philosophy*, Yale University Press.
77. THOMAS NAGEL, 1974, "What is it Like to be a Bat?", *Philosophical Review*, 83: 435-450.
78. ALLEN NEWELL, 1980, "Physical Symbol Systems", *Cognitive Sciences*, 4: 135-183.
79. DANIEL N OSHERSON & EDWARD E SMITH, 1981, "On the Adequacy of Prototype Theory as a Theory of Concepts", *Cognition*, 11: 35-58.
80. STEPHEN E PALMER, 1999, "Color, Consciousness, and the Isomorphism Constraint", *Behavioral and Brain Sciences*, 22: 923-990.
81. CHRISTOPHER PEACOCKE, 1992, *A Study of Concepts*, MIT Press.
82. JEAN PIAGET, 1932, *Le jugement moral chez l'enfant*, (1969, Presses Universitaires de France).
83. JEAN PIAGET, 1945, *La formation du symbole chez l'enfant*, (1978, Delachaux et Niestlé).
84. MASSIMO PIATTELLI-PALMARINI, 1979, *Théories du langage - Théories de l'apprentissage*, Seuil.
85. ARTHUR N PRIOR, 1967, *Past, Present, and Future*, Oxford University Press.
86. JAMES PUSTEJOVSKY, 1995, *The Generative Lexicon*, MIT Press.
87. HILARY PUTNAM, 1975, *Mind, Language and Reality*, Cambridge University Press.
88. ZENON W PYLYSHYN, 1984, *Computation and Cognition: toward a foundation for cognitive science*, MIT Press.
89. WILLARD V O QUINE, 1953, *From a Logical Point of View*, Harvard University Press.
90. WILLARD V O QUINE, 1960, *Word and Object*, MIT Press.
91. HANS REICHENBACH, 1947, *Elements of Symbolic Logic*, Macmillan.
92. ELEANOR ROSCH, 1975, "Cognitive Structures of Semantic Categories", *Journal of Experimental Psychology: general*, 104: 192-233.
93. DAVID E RUMELHART & JAMES L MCCLELLAND, 1986, *Parallel Distributed Processing: volume I, foundations*, MIT Press.
94. BARBARA A C SAUNDERS & JAAP VAN BRAKEL, 1997, "Are There Non-Trivial Constraints on Color Categorization?", *Behavioral and Brain Sciences*, 20: 167-228.
95. CHRISTINE A SKARDA & WALTER J FREEMAN, 1987, "How Brains Make Chaos in order to Make Sense of the World", *Behavioral and Brain Sciences*, 10: 161-195.
96. EDWARD E SMITH & DOUGLAS MEDIN, 1981, *Categories and Concepts*, Harvard University Press.

97. PAUL SMOLENSKI, 1988, "On the Proper Treatment of Connectionism", *Behavioral and Brain Sciences*, 11: 1-23.
98. PAUL SMOLENSKY, 1990, "Tensor Product Variable Binding and the Representation of Symbolic Structures in Connectionist Systems", *Artificial Intelligence: an international journal*, 46: 159-216.
99. JOHN F SOWA, 1984, *Conceptual Structures: information processing in mind and machine*, Addison-Wesley.
100. ELIZABETH S SPELKE, 1990, "Principles of Object Perception", *Cognitive Science*, 14: 29-56.
101. DAN SPERBER & DEIRDRE WILSON, 1986, *Relevance: communication and cognition*, Blackwell.
102. LEONARD TALMY, 2000, *Toward a Cognitive Semantics: volume I, concept structuring systems*, MIT Press.
103. BARBARA TVERSKY, 1993, "Cognitive Maps, Cognitive Collages, and Spatial Mental Models", *Conference on Spatial Information*.
104. JOHAN F A K VAN BENTHEM, 1983, *The Logic of Time: a model-theoretic investigation into the varieties of temporal ontology and temporal discourse*, Reidel.
105. TIM J VAN GELDER, 1998, "The Dynamical Hypothesis in Cognitive Science", *Behavioral and Brain Sciences*, 21: 615-628.
106. FRANCISCO J VARELA, 1988, *Invitation aux sciences cognitives*, Seuil.
107. ZENO VENDLER, 1967, *Linguistics in Philosophy*, Cornell University Press.
108. BERNARD VICTORRI & CATHERINE FUCHS, 1996, *La polysémie : construction dynamique du sens*, Hermès.

Index

A

apprentissage
 biais de, 289
 mécanismes de, 287–89
 arbitraire, 184
 Armstrong, 135
 atomisme conceptuel, 231

B

Baddeley, 67
 Barsalou, 254, 257
 Barwise, 151
 Bates, 104
 Bickerton, 144, 145, 256
 Block, 102

C

Carey, 129
 Carnap, 123, 125
 Carpenter, 155, 156
 Carruthers, 145
 carte conceptuelle, **264**, 275
 carte temporelle, **67–68**, 72, 74
 Chalmers, 189
 Changeux, 235, 289
 Chomsky, 104, 106, 142, 143, 164, 219, 235
 Churchland, 95
 Clark, 118, 253
 classe sémantique, 187
 combinatoire, 188
 compositionnalité, 185–87, 279
 Comrie, 20, 21
 concepts
 ancrage de, **89**, 93–96
 application de, 100
 atomiques *Voir* atomisme conceptuel
 caractère compositionnel de, **89**, 141–42, 196
 en tant que descriptions, 195–96
 en tant que percepts intégrés, 103–7
 en tant que représentations éphémères, **255**, 259–65
 en tant que symboles, 107–8
 évocation de, 100
 fonction de, 102
 hypothèse atomiste de *Voir* atomisme conceptuel

hypothèse moléculaireiste de *Voir* molécularisme conceptuel
lexicaux, 192
moléculaires *Voir* molécularisme conceptuel
primitifs, 216–18
rôle inférentiel de, **89**, 121–22, 195
constituance, 187–88, 279
Crick, 96

D

Davidson, 46, 47, 48, 98
de Glas, 54
définitions (théorie de), 122–27
Dehaene, 235, 289
Desclès, 50, 51, 52, 53
Dessalles, 105, 106, 107, 144, 256, 271, 275, 290
digital, 184
dilemme de la granularité, **62**
dilemme du mentalais, **243**
discret, 184
Dretske, 112

E

échelle, **68**, 275
Elman, 104
empirisme, 103, 251
époque, **68**
espace qualitatif, 252–55
état, **49**, 70
événement, **49**, 70
externalisme, 99

F

Fauconnier, 56, 57
Fodor, 98, 114, 115, 116, 117, 126, 127, 131, 132, 185, 187, 188, 189, 190, 191, 214, 216,
217, 220, 222, 226, 231, 232, 233, 234, 235, 241, 243
Foley, 66
Freeman, 96
Fuchs, 278

G

Gärdenfors, 172, 253, 262, 265, 266, 267, 268, 269
Garrett, 126, 127, 214
Gleitman H, 135
Gleitman L R, 135
Gopnik, 129, 130
Gosselin, 79, 80
Gould, 114
grain, **62**
grammaire générative (et la composition sémantique), 142–46

grammaires cognitives, 169–73
grille conceptuelle, **264**, 274
grille temporelle, **67–68**, 69, 71
Guha, 239

H

Harnad, 94, 107, 278
Hofstadter, 167, 168, 225
Hopfield, 96

I

instant, **32**, 45–46
intentionnalité
 et communication, 108
 et information, 113
internalisme, 99
internalisme conceptuel, 102
internalisme méthodologique, **99**
intervalle, **32**, 45–46

J

Jackendoff, 124, 159, 160, 161, 221
Jackson, 234
Johnson M H, 104
Johnson M K, 66, 69
Johnson-Laird, 164, 165
jugement sémantique, 184

K

Kamp, 29, 30, 31, 49, 50, 136, 137, 138, 270, 271
Kaplan, 111
Katz, 124
Kayser, 166, 167
Keil, 130
Koch, 96
Kosslyn, 72, 252, 256, 260
Kripke, 108, 109

L

Lakoff, 243
Langacker, 170, 171
langage mental, 190
Lenat, 239
Lepore, 131, 132, 226
Levin, 162
Lewontin, 114
Leyton, 258

M

Margolis, 236, 237
Markman, 218
McCarthy, 225
McClelland, 95
McLaughlin, 188, 189, 243
McNeill, 184, 274
Medin, 133, 134, 135, 222
Meltzoff, 129, 130
mémoire temporelle, **62**
 comparaison (capacité de), 62
 insertion (capacité de), 63
mentaux, 179
Millikan, 113
Mitchell, 287
moléculisme conceptuel, **197**
 méthode définitionnelle, **198**
 méthode relationnelle, **198**
moment, **68**
Monod, 118
monotonie de la composition, **239**
Montague, 149, 150, 151, 154, 155

N

Nagel, 234
Newell, 93

O

omnipotence de la conceptualisation, **251**
opacité structurelle, **254**
Osherson, 137

P

Palmer, 118
paradoxe de la connexion, **107**, 116–18, 233–36
Parisi, 104
Parkes, 126, 127, 214
Partee, 136, 137, 138
Peacocke, 102
Perry, 151
Piaget, 106, 225, 288
point de vue, **50**
Prior, 27, 29
problème de la connexion *Voir* paradoxe de la connexion
productivité, 185
prototypes (théorie de), 133–38
Pustejovsky, 163
Putnam, 109, 110
Pylyshyn, 94, 185, 187, 188, 189, 243

Q

Quine, 112, 128, 129, 132

R

Raye, 66

réalisme, 99

réalisme empirique, **99**

récurtivité

de construction du sens, 272

de repérage temporel, 72–74

référence (de mots), 100

Reichenbach, 22, 23, 24

re-présentation, 97

représentations

caractère causal de, 101

caractère intentionnel de, 96

contenu, 96

forme, 101

sémantique, 94

syntaxe, 94

Reyle, 29, 30, 31, 49, 50, 270, 271

Rosch, 133, 134

Rumelhart, 95

S

Saunders, 218

sémantique formelle, 146–59

sémantique lexicale, 159–64

sémantique procédurale, 164–69

sens (de mots), 103

Skarda, 96

Smith, 133, 134, 135, 137, 222

Smolensky, 188, 243

Sowa, 166, 222

Spelke, 218

Sperber, 278

Suengas, 66

systematicité, 185–88, 210–11, 278–79

d'extraction, **185**

de négation, **185**

de substitution, **185**

système conceptuel, **89**

en tant qu'interface, **255**, 259–65

interface avec la perception, 93–96

interface avec le langage, 141–42

interface avec le raisonnement, 121–22

trois interfaces de, **89**

T

Talmy, 169
temps qualitatif, 66–67, 255
théories (théorie de), 128–33
Tversky, 67

V

van Benthem, 32
van Brakel, 218
van Gelder, 93, 188
Varela, 95, 96
Vendler, 24, 25, 26
Victorri, 278

W

Walker, 126, 127, 214
Wilson, 278

Z

Zénon, 63, 64, 65, 73, 273