

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 moléculaire, 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.

```

tuer (e1, e2) ⇒déf
    séquence:  étape1:  thème:  e2,
                cause:  e1,
                étape2:  événement, thème: e2,
                étape3:  nul,
    argument1: e1,
    argument2: e2, animé.

```

À la lecture du trait *séquence*, l'interpréteur considère tout d'abord la structure *tuer* comme étant du type *prédicat*. Mais en parvenant au trait *argument1*, il doit spécifier ce type comme étant le type *binaires*. L'interpréteur vérifie que les valeurs données pour les traits respectent les contraintes de typage. Par exemple, la structure *tuer* n'aurait pas pu posséder en même temps les traits *argument* et *argument1*, car ces deux traits appartiennent à deux types différents, *unaire* et *binaires*, qui n'héritent pas l'un de l'autre. Les contraintes peuvent porter sur les types des traits de la structure. Ainsi, le trait *séquence* ne pourra être occupé que par un type particulier de séquence dont le trait *étape1* possède le trait *cause*, autrement dit un type *processus*. Noter que la valeur du trait *étape2* est explicitement précisée comme étant du type *événement*. L'utilisation de variables permet d'ajouter des contraintes entre les traits d'une même structure. Dans notre exemple, les structures temporelles et argumentales du prédicat sont portées par des traits différents. Le lien entre ces deux structures est assuré par le partage des variables *e1* et *e2*. Ainsi, le premier argument du prédicat, par ce liage de variable, est imposé en tant que *cause* du processus correspondant au trait *étape1*. De même, le deuxième argument, identifié par la variable *e2* et spécifié comme étant du type *animé*, apparaît comme *thème* de l'événement résultant.

Du fait de la richesse des structures de traits, la construction du sens ne nécessite qu'un mécanisme général, le mécanisme d'unification. L'unification consiste à appairer les structures conceptuelles en suivant les règles syntaxiques de la langue. Par exemple, le décodage d'une phrase de type *A tue B* pourra provoquer l'appariement de la structure *tuer* avec ces deux arguments. L'opération d'unification teste les contraintes, par exemple le fait que les deux arguments du verbe *tuer* sont de type *entité* et que le deuxième est de type *animé*. L'itération du mécanisme d'unification réalise la même tâche que la dérivation dans le formalisme fonctionnel. Ainsi, la construction du sens de la phrase *A tue B* conduira, en cas de réussite de l'appariement, à une structure du type suivant.

```

séquence:  étape1:  thème:  B,
                cause:  A,
    étape2:  événement, thème:  A,
    étape3:  nul,
argument1: A,
argument2: B.

```

La structure de traits que nous avons donnée pour le concept *TUER* laisse implicite le lien qui existe avec le concept *MOURIR*.

```

mourir (e) ⇒déf
    séquence:  étape1:  événement, thème: e,
                étape2:  nul,
                étape3:  nul,
    argument:  e, animé.

```

Il serait préférable d'indiquer de manière explicite que le résultat du concept *TUER* est identique à la situation décrite par le concept *MOURIR*. La valeur du trait *étape1* dans la

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

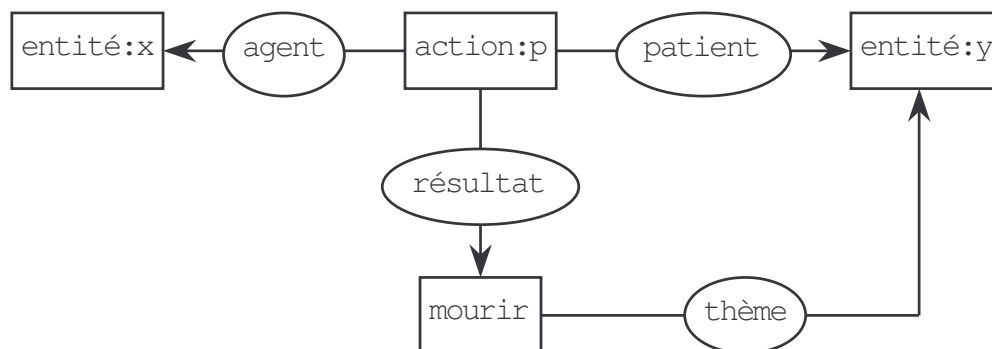
de tout un répertoire de relations entre concepts que la logique n'offre pas⁶. L'idée à la base des réseaux sémantiques consiste à représenter explicitement la mise en relation des concepts. Dans sa forme la plus immédiate, cette approche place les concepts aux nœuds d'un graphe dont les arêtes sont 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.

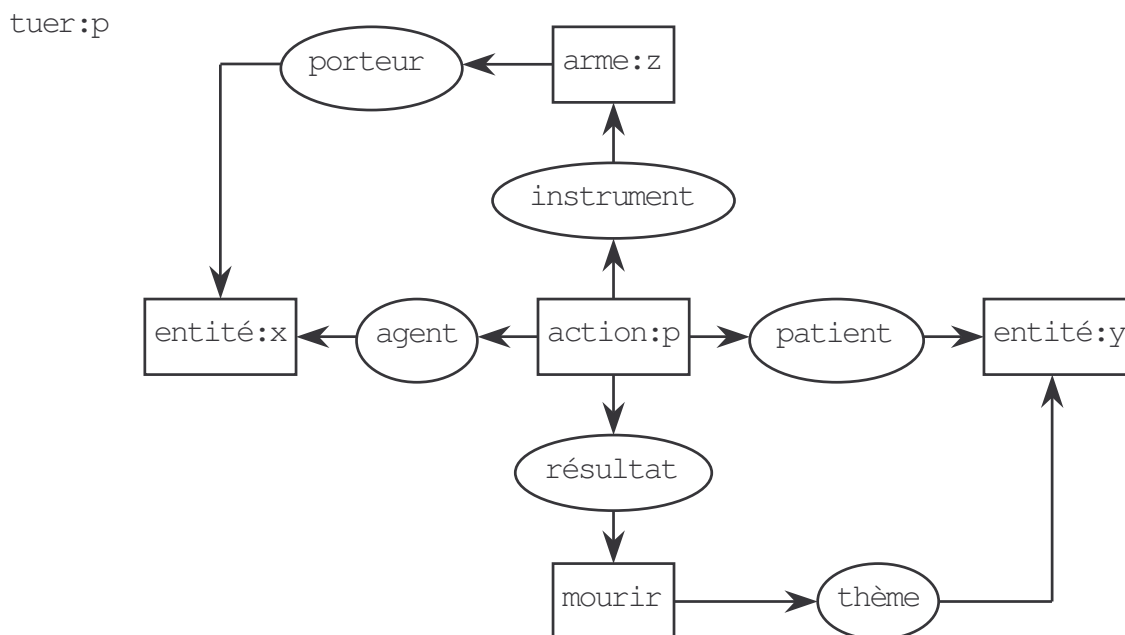
⁷ Le mot graphe conceptuel est parfois utilisé dans un sens différent, pour parler de la représentation d'une phrase.

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.

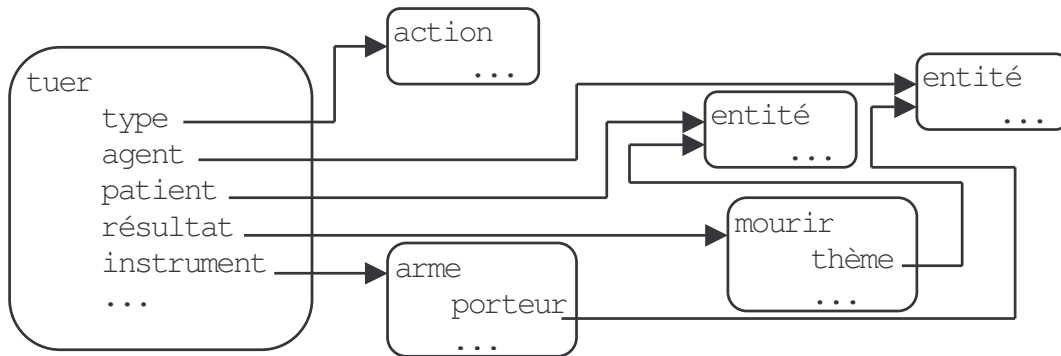
[mourir:p]: [événement:p] & [entité:x] & [vie:e₁] & [mort:e₂];
thème(x,p) & antérieur(e₁,p) & ultérieur(e₂,p).

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 ramification, 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 précedence 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.

$$\mathbf{s}(m_1, m_2, \dots, m_k) = \mathbf{g}_{S_i}(s_2, s_3, \dots, s_k)$$

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 moléculaire - 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 moléculaire 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 moléculaire

L'intérêt, tant pratique que théorique, de l'approche moléculaire 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 précédence 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 systématisme.
- 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 systématisme.

- 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.