

Chapitre 6 :
Le calcul du sens

Introduction

L'une des difficultés principales de la modélisation sémantique du langage vient du fait que les situations qu'il permet de décrire appartiennent à un monde perçu comme continu : deux pommes ne sont jamais identiques. Or, il emploie pour cela un code qui, par de nombreux aspects, est un code combinatoire digital¹. On retrouve ce caractère digital dans la phonologie, la morphologie et la syntaxe. Les représentations sémantiques sont elles de type digital, à l'image du reste du langage, ou sont elles analogiques, à l'image de notre interaction avec l'environnement sur laquelle portent les énoncés ? L'existence de phénomènes systématiques a conduit de nombreux auteurs à choisir la première option, en considérant que les représentations mentales évoquées par les mots étaient, elles aussi, organisées de manière à former un système combinatoire digital.

Parmi les aptitudes cognitives humaines, la capacité de donner un sens aux énoncés de notre langue est au centre de ce qui nous distingue des autres espèces animales. Peut-elle se modéliser comme le fonctionnement d'une machine de Turing ?

6.1. Le code du langage

Les êtres humains utilisent le langage, instancié dans les différentes langues, comme un code leur permettant de communiquer. Ce code présente la propriété remarquable d'être, par bien des aspects, un code tout-ou-rien. Dans un contexte donné, nous nommons la situation ou l'objet auquel nous sommes confrontés à l'aide d'une expression langagière. Nous pouvons ainsi désigner l'objet devant nous par le mot *table*. Ensuite, nous pouvons argumenter avec nos interlocuteurs pour démontrer que c'est bien cette expression qui est appropriée, tandis que d'autres ne le sont pas. Si l'on observe des personnes en train de discuter dans les conditions normales et spontanées, sans qu'ils soient soumis à des obligations de type psychologique ou sociologique, il arrive fréquemment qu'on les voie argumenter sur la bonne désignation d'un objet ou d'une situation. Est-ce bien une table, ou doit-on parler de bureau ?

Certes le code langagier est extrêmement flexible en fonction du contexte dans lequel il est utilisé et de l'individu qui l'utilise. Les objets qu'une étudiante en sciences cognitives nommera par le mot *table* ne sont pas exactement les mêmes que ceux qu'un antiquaire désignera par ce mot. L'antiquaire dispose de nombreux autres mots, comme *guéridon*, *console*, ou *desserte*, pour désigner les choses que l'étudiante confond sous le vocable *table*. De plus, si l'étudiante n'est pas francophone d'origine, elle pourra utiliser le mot *table* pour nommer ce qui, pour un Français, ne peut être désigné que par le mot *bureau*, voire par les mots *pupitre* ou *secrétaire*, simplement parce que cette distinction n'est pas habituelle dans sa langue maternelle. Pourtant, les mêmes personnes, lors d'un pique-nique, n'auront aucun problème pour trouver l'accord tacite qui consiste à désigner un tronc d'arbre scié par le mot *table*.

Cette variabilité dans l'emploi des mots ne signifie pas que le langage est intrinsèquement flou. Lorsque les interlocuteurs se mettent d'accord, dans chacun de ces contextes, l'objet perçu est, ou n'est pas, une table. Il ne devient pas une table à 76%. Le langage nous offre un code de communication qui peut être arbitrairement précis, dans chaque contexte, pour satisfaire aux besoins de la communication. Cette précision nous permet

¹ À titre de comparaison, le code de communication des abeilles est également combinatoire, mais analogique. L'angle que la danse produit avec la verticale dans l'obscurité de la ruche reproduit, de manière graduelle, l'angle entre la direction du soleil et la direction à prendre pour trouver la source de nourriture.

notamment d'amener notre interlocuteur à trouver la bonne référence. On l'obtient à l'aide de mots appropriés, mais aussi, lorsque ceux-ci font défaut, en construisant des syntagmes comme *table à écrire*, ou *table à dessiner des plans*. Ainsi, le langage apparaît comme un code par lequel nous pouvons communiquer à propos d'un nombre incalculable de situations, avec le plus souvent une précision qui satisfait aux besoins de l'interaction. Ce code est construit à partir d'un ensemble d'éléments de base, le lexique, qui est arbitraire, discret, dispersé, et donc digital.

Le caractère arbitraire du lexique apparaît de manière manifeste lorsque l'on compare les langues entre elles. Il est difficile de trouver de ressemblance entre les mots *pierre*, *sangue*, *stone*, *hajar*, désignant les mêmes objets perçus en français, en persan, en anglais et en arabe, et une perception quelconque que l'on peut avoir d'une pierre. Noter que le caractère arbitraire du lien symbolique n'est pas systématique dans tous les codes de communication. Dans les langues signées, certaines unités lexicales sont construites à partir de la ressemblance entre le geste et l'objet ou la situation qu'il désigne (MCNEILL 1992 [72]). Les individus entendant ne parlant pas la même langue utilisent spontanément des gestes iconiques pour se comprendre. Mais même dans les langues signées, l'iconicité se perd par le jeu de l'évolution linguistique qui restaure l'arbitraire du signe (MCNEILL 1992 [72]).

Le lexique présente en outre la caractéristique d'être discret. Cette propriété est remarquable si on la rapproche de la variation des données de la perception. Lorsque nous nommons certains objets perçus par les mots *gravillon*, *caillou*, *pierre*, *rocher*, nous imposons une échelle discrète là où notre perception nous permet de discerner une variété beaucoup plus riche.

Le lexique est non seulement arbitraire et discret, mais également dispersé. Une série de mots comme *gravillon*, *galet*, *caillou*, *pavé*, *moellon*, *pierre*, *rocher* ne réalise pas un échantillonnage uniforme de la continuité perceptive. Ces mots renvoient typiquement à des objets perçus qui ne se différencient pas seulement par leur taille, mais aussi par tout un ensemble de caractéristiques, si bien qu'ils se trouvent dispersés dans un ensemble de haute dimension. Le gravillon est typiquement anguleux, au contraire du galet qui est lisse ; le galet est associé à la rivière, le gravillon à la route ; *et cætera*. Du fait de cette dispersion, le lexique réalise un échantillonnage dépourvu de toute régularité de l'espace de la perception. Du fait de la dispersion du code, le langage ne conserve pas les rapports de voisinage fournis par la perception. Il n'existe pas de dépendance analogique, continue ou discrète, entre les éléments du lexique et les entités perçues qu'elles désignent.

Par ses aspects discrets non analogiques, le langage apparaît comme un code digital. Aucune mesure de distance entre les mots ne reproduit la distance entre les objets correspondants de la perception. Ce caractère digital s'ajoute à un aspect combinatoire pour former ce qui ressemble à une mécanique.

6.2. Le jugement sémantique

Le langage ne consiste pas en une suite de mots décorrelés. À partir du lexique et de la grammaire, nous créons des expressions qui, elles-mêmes, peuvent renvoyer à une entité ou une situation perçue. Face à une collection de pierres précieuses, l'individu qui est à court de mots utilisera des syntagmes comme *l'émeraude qui ressemble à une olive*, ou *le rubis en forme de triangle*. Alors que le lexique est constitué d'une liste finie d'éléments, l'ensemble des phrases possibles est infini. Ce qui nous intéresse ici est comment déterminer l'ensemble des phrases "sensées", autrement dit les phrases auxquelles un locuteur compétent pourra accorder une signification. Cette question présuppose l'existence d'un jugement d'acceptabilité sémantique. La démarche est habituelle dans les sciences du langage.

Le même type de problème se pose aux niveaux phonologique, morphologique, syntaxique et pragmatique, car il existe un jugement d'acceptabilité pour chacun de ces niveaux : un locuteur compétent sait reconnaître une prononciation acceptable, un mot composé possible, une phrase bien formée et une intervention pertinente. Le jugement d'acceptabilité divise, idéalement, l'ensemble des énoncés en deux classes, les énoncés corrects et ceux qui ne le sont pas. Le travail de modélisation consiste alors à identifier ces classes de la manière la plus économique possible. Dans le cas du jugement sémantique, la tâche de modélisation va consister à tenter de découvrir un nombre réduit de principes permettant de prédire qu'un énoncé aura un sens ou n'en aura pas. On pourrait reprocher au jugement d'acceptabilité sémantique d'être très peu binaire, en tout cas beaucoup moins binaire que le jugement syntaxique. Même une phrase comme le jardin de la porte mange le ciel peut, à condition qu'on y mette un peu de temps, recevoir un sens. Dans ce cas, le modèle devra expliquer le caractère graduel de l'acceptabilité.

L'ensemble Σ des expressions sémantiquement acceptables possède certaines propriétés fondamentales qu'il s'agira d'expliquer. Tout d'abord, il est potentiellement infini. Il ne semble pas exister de borne supérieure, hormis celles qui sont liées aux contraintes matérielles, au nombre de significations différentes qu'un être humain peut former et exprimer par le langage. Ensuite, il possède des régularités systématiques. Une première forme de systématique est la systématique de substitution. Si la phrase Jean aime Marie est interprétable, alors la phrase Paul aime Marie l'est aussi. Pour les défenseurs de la systématique sémantique, on ne trouvera pas d'être humain qui serait capable de comprendre l'une de ces deux phrases en étant strictement incapable de comprendre l'autre (FODOR & PYLYSHYN 1988 [41]). L'ensemble Σ des expressions interprétables semble être clos pour certaines opérations de substitution qui préservent la "classe sémantique" des objets substitués. Un deuxième aspect de la systématique concerne la systématique d'extraction. Il existe un lien obligatoire entre l'interprétation d'une expression et celle de ses composants. Si nous pouvons comprendre correctement la phrase le frère de Jacques aime Marie, alors nous pouvons comprendre le syntagme le frère de Jacques, de même que le mot frère (FODOR & PYLYSHYN 1988 [41]). Ajoutons à cela une systématique de négation, précisant que si une phrase appartient à Σ , alors sa négation y figure également, de manière systématique.

Le caractère infini de Σ et ses aspects systématiques constituent un défi pour la modélisation. Il est exclu que Σ résulte d'un apprentissage par cœur. La production des expressions sensées ne peut donc que résulter d'un calcul. Pour déterminer la nature de ce calcul, il est utile d'examiner comment les propriétés de Σ peuvent être reproduites.

L'infinité de Σ requiert que la génération du sens soit due à un mécanisme productif. La productivité d'un mécanisme signifie qu'il peut engendrer un ensemble non borné de formes différentes. La productivité est rendue possible, notamment, grâce à des mécanismes récursifs. L'intérêt de cette notion est de souligner la possibilité que des ensembles potentiellement infinis puissent être engendrés par des moyens finis.

Pour expliquer la systématique d'extraction, on suppose généralement que les mécanismes qui engendrent Σ sont compositionnels. Selon le principe de compositionnalité, le sens attaché à une expression peut être entièrement calculé à partir du sens des mots qu'elle contient (FODOR & PYLYSHYN 1988 [41]). On peut exprimer cette idée par l'équation suivante.

$$\mathbf{s}(m_1, m_2, \dots, m_k) = \mathbf{f}(s_1, s_2, \dots, s_k), \text{ où } s_i = \mathbf{s}(m_i)$$

La fonction $\mathbf{s}()$ est la fonction qui attache un sens à une expression langagière. Notamment, elle associe un sens s_i à un mot m_i . La fonction $\mathbf{f}()$ produit le sens d'une

expression en fonction, entre autres, de sa structure grammaticale. On comprend ainsi que si une expression possède une signification, autrement dit si elle appartient à Σ , alors ses éléments appartiennent nécessairement à Σ , puisque la fonction $\mathbf{s}()$ est définie sur eux². La compositionnalité peut fournir ainsi une explication naturelle à la systématique d'extraction.

Une conception peu plausible de la construction sémantique voudrait que la fonction $\mathbf{f}()$ ne commence à être calculée qu'une fois que tous ses arguments sont disponibles. Il est manifeste que dans la plupart des situations, les sujets commencent l'élaboration d'un sens bien avant d'atteindre la fin d'une phrase, comme on peut sans doute le vérifier dans le présent paragraphe. De plus, l'agencement hiérarchique des syntagmes laisse supposer que le calcul du sens est effectué de proche en proche (ce qui n'exclut pas des retours en arrière éventuels), chaque syntagme étant évalué pour lui-même. Par exemple, pour une structure syntaxique simple comme $(m_1 m_2) m_3$, on pourrait avoir idéalement deux étapes de calcul suggérées par l'écriture suivante.

$$\begin{aligned} \mathbf{s}(m_1, m_2, m_3) &= \mathbf{f}(\mathbf{s}(m_1, m_2), \mathbf{s}(m_3)) \\ &= \mathbf{f}(\mathbf{f}(s_1, s_2), s_3) \end{aligned}$$

Ce genre d'écriture confère un certain caractère sériel à la construction du sens, hérité de la sérialité structurelle du canal linguistique³. Si l'on accepte que le sens est ainsi, dans une certaine mesure, construit par couches successives, on comprend mieux la possibilité même de la systématique d'extraction. Cette propriété peut être re-formulée d'une manière un peu différente, en écrivant l'expression $\mathbf{f}(s_1, s_2, \dots, s_k)$ comme une fonction de la variable s_k .

$$\mathbf{s}(m_1, m_2, \dots, m_k) = \mathbf{g}_{s_1, s_2, \dots, s_{k-1}}(s_k)$$

Du fait du cloisonnement interne au calcul du sens, une substitution lexicale du mot m_k par le mot n_k n'affectera qu'une partie des arguments de la fonction $\mathbf{f}()$ et ses effets resteront, dans beaucoup de cas, circonscrits. On peut s'attendre, de ce fait, à de nombreux phénomènes de systématique, car le sens construit, sa validité et les inférences qu'il déclenche ont des chances de persister en partie après la substitution. Par exemple, le sens de la phrase Jean aime une artiste est présenté comme le résultat de l'application d'une fonction sémantique au concept ARTISTE. Cette fonction est une construction associée à l'expression linguistique Jean aime ... qui prend le concept ARTISTE comme argument. L'idée, selon laquelle le sens est construit par étapes, se traduit ici par le fait que l'expression $\mathbf{g}_{s_1, s_2, \dots, s_{k-1}}()$, en tant que fonction, est un objet dont le résultat de l'application à la variable s_k n'est pas erratique. Autrement dit, une substitution affectant la variable s_k aura des conséquences limitées, ce qui engendre des phénomènes de systématique. On peut établir une analogie entre ce phénomène de régularité, due à l'imbrication du traitement sémantique, et la propriété mathématique de continuité des fonctions. Lorsque l'on réécrit une fonction $\mathbf{f}(x, y)$ sous la forme $\mathbf{g}_x(y)$, transformant l'un des arguments en paramètre, on passe de l'équation d'une surface à celle d'une famille de courbes. Ce faisant, on privilégie une direction en chaque point de la surface. Si les courbes $\mathbf{g}_x(y)$ sont dérivables, on peut localement prédire leur comportement. C'est cette idée du comportement prédictif de la fonction $\mathbf{f}(x, y)$ par rapport à l'un de ses arguments qu'on retrouve dans l'idée de systématique de substitution.

Le fait que Σ respecte la systématique de substitution peut s'expliquer par le fait que la fonction $\mathbf{f}()$ porte sur des classes de significations et non les significations elles-mêmes. Si

² L'hypothèse implicite est que $\mathbf{f}(m_i) = s_i$, puisque pour un mot isolé la fonction \mathbf{f} s'instancie par l'équation $\mathbf{f}(m_i) = \mathbf{s}(m_i)$.

³ Cet aspect sériel est certainement loin d'être strict. L'idée est plutôt de critiquer la possibilité d'un parallélisme radical. La sérialité provient non seulement du traitement successif des syntagmes imbriqués, mais également et plus simplement de l'ordre des mots dans l'énoncé.

l'on substitue un élément par un élément de la même classe, le mécanisme continue de s'appliquer. Lorsque les chimistes ont découvert des régularités comme celle représentée par l'équation *acide + base* → *sel + eau*, ils étaient en droit de l'expliquer par le fait que le mécanisme sous-jacent n'était pas sensible aux réactifs eux-mêmes, mais seulement à leur caractère acide ou basique. De même, les régularités de substitution laissent supposer que les mécanismes compositionnels ne sont sensibles qu'à la classe sémantique des mots. Cette idée peut être illustrée par la réécriture suivante de l'expression $\mathbf{f}(s_1, s_2, \dots, s_k)$.

$$\mathbf{s}(m_1, m_2, \dots, m_k) = \mathbf{g}_{s_1, s_2, \dots, s_{k-1}}(\mathbf{t}[u]), \text{ où } \mathbf{s}(m_k) = \mathbf{t}[u] \text{ pour } u \in D_{\mathbf{t}[\]}$$

Selon cette équation, la fonction $\mathbf{g}_{s_1, s_2, \dots, s_{k-1}}(\)$ accepte comme argument des éléments qui partagent une certaine propriété. Cela revient à postuler l'existence d'une forme de typage⁴, que nous représentons à l'aide de la fonction $\mathbf{t}[\]$. Le typage permet de prédire l'interprétation systématique de l'adjectif foncé lorsqu'il s'applique à des noms de couleur comme *vert, jaune, émeraude, et cætera*. Dans ce cas, la fonction $\mathbf{t}[\]$ filtre les seuls concepts de couleur, en laissant sa variable libre de parcourir tout l'ensemble des couleurs. De même, on comprend pourquoi un lien systématique existe entre le fait que la phrase *Jean aime Marie* soit interprétable et le fait que la phrase *Paul aime Marie* le soit également. Les entités désignées par le prénom *Jean* et par le prénom *Paul* partagent un ensemble de propriétés, notamment celle d'être humain, qui sont censées garantir l'existence d'une interprétation. Tout se passe comme s'il existait un patron de pensée, du genre <humain> aime Marie, permettant à n'importe quelle instanciation du type <humain> de conduire à une pensée possible. Pour illustrer ce phénomène lié à la compositionnalité, il est utile de considérer un contre-exemple. Le sens d'une expression idiomatique comme *retourner sa veste*, dans l'esprit de la plupart des locuteurs, n'est pas lié au sens du mot *veste*. Ce défaut de compositionnalité se traduit par une lacune dans la systématisme : le fait de pouvoir comprendre l'expression *retourner son manteau* (dans son sens matériel, par exemple pour un manteau réversible) n'est pas systématiquement lié au fait de comprendre l'expression *retourner sa veste* (dans son sens figé de brusque changement d'opinion ou de camp). La nature de ce genre d'exception permet de comprendre le cas général, qui est que la compositionnalité est un excellent moyen d'expliquer les phénomènes de systématisme. Pour prendre un autre contre-exemple, l'ensemble des recettes gastronomiques est engendré par un mécanisme productif, la combinaison d'ingrédients. Cependant, ce mécanisme n'est pas connu pour être compositionnel : la production des recettes gastronomiques, à l'exclusion des autres, procède par essais - erreurs, car on ne connaît pas de fonction $\mathbf{f}(\)$ permettant de prédire le goût du résultat à partir de celui des ingrédients.

Le principe de compositionnalité est parfois jugé insuffisant pour expliquer les phénomènes de systématisme. Pour assurer systématiquement le lien entre la signification s_i d'un mot m_i et une représentation calculée comme $\mathbf{f}(s_1, \dots, s_i, \dots, s_k)$, on peut considérer que cette représentation doit conserver une mémoire de l'élément s_i . Le moyen le plus simple est d'exiger que cette représentation contienne, dans un sens matériel, ses composantes. Ainsi, selon le principe de constituance, l'accès systématique à la représentation sémantique associée au prénom *Jean* à partir de la représentation sémantique associée à la phrase *Jean aime Marie* est assurée par le fait que la première fait partie intégrante de la seconde (FODOR & PYLYSHYN 1988 [41]). La présence effective des constituants au sein des représentations garantit leur accessibilité systématique. On peut traduire ce principe par le fait qu'il s'agit de représentations récursives : des représentations qui contiennent d'autres

⁴ De manière étonnante, comme nous le verrons, certains auteurs qui mettent en avant ce type de systématisme refusent un tel typage de concepts, sans proposer d'explication alternative.

représentations de même nature. Le principe de constituance est une exigence forte, qui pose le problème du rapport entre le composé et ses constituants. Dans le cas de représentations résultant de l'interprétation des expressions du langage, les principes de compositionnalité et de constituance supposent que la structure de l'expression langagière se transporte dans la structure de la représentation mentale qui lui correspond. C'est ainsi que ces deux principes, constituance et compositionnalité, parviennent à rendre compte des phénomènes de systématique sémantique.

6.3. La combinatoire du langage

La question devient de savoir quel type de calcul est capable, pour engendrer un ensemble Σ doté des propriétés d'infinité et de systématique, d'être productif et compositionnel en produisant des représentations respectant le principe de constituance. Dans une certaine acception du mot "calcul", il s'agit de modéliser le comportement d'un système, le cerveau, à l'aide des mathématiques, en conservant une distinction radicale entre le modèle et le système. Le physicien ne prétend pas qu'une planète résout une équation différentielle pour déterminer sa prochaine position. De même, le chercheur en sciences cognitives qui s'inscrit dans la tradition cybernétique suppose que le système qui engendre les combinaisons langagières est de nature dynamique, régi par des équations d'évolution et des principes de minimisation d'énergie, sans croire que les calculs de son modèle sont réellement effectués par le système lui-même (□AN GELDER 1998 [105]). Les mécanismes qui sous-tendent le langage, comme tous les processus naturels, peuvent en principe être décrits comme des processus dynamiques, ne serait-ce qu'au niveau de leur implémentation neuronale. Cependant, la nature du langage permet d'espérer mieux. Le langage est caractérisé par des changements d'états discontinus et des unités hiérarchisées : phonèmes, mots, syntagmes. Il passe par un canal sériel, et son traitement présente de ce fait des aspects séquentiels. La syllabe et les syntagmes se prêtent à des analyses qui produisent des structures arborescentes par enchâssement. Tout ceci suggère fortement la possibilité d'un niveau de description qui se situe au-dessus du niveau dynamique. En ce qui concerne les aspects sémantiques, le fait que les phénomènes systématiques concernent des substitutions ou des extractions d'éléments plaide également en faveur d'un niveau de description que l'on peut qualifier de symbolique ou de computationnel, consistant en un calcul de type combinatoire.

Les techniques combinatoires, en mathématiques, regroupent les méthodes d'énumération et de génération d'objets finis discrets. Contrairement aux modèles dynamiques, les mécanismes symboliques qui expliquent un phénomène cognitif sont supposés réalisés dans le système qu'ils modélisent (□AN GELDER 1998 [105]). Le calcul de gradient ne prétend pas constituer une image réaliste du comportement de la bille qui roule sur une pente. Son mouvement peut être aussi modélisé par un calculateur analogique utilisant des courants électriques pour représenter les paramètres physiques. En revanche, tous les systèmes connus capables d'effectuer des calculs combinatoires sur des structures discrètes sont des variantes ou des spécialisations de la machine de Turing. En conséquence, si un aspect de la cognition se laisse modéliser par un mécanisme computationnel combinatoire, il est naturel de faire l'hypothèse que le cerveau effectue une combinatoire du même type.

Cette question de la nature combinatoire du calcul sémantique s'est trouvée au centre d'une polémique (FODOR & PYLYSHYN 1988 [41] ; FODOR & McLAUGHLIN 1990 [40] ; SMOLENSKY 1988 [97] ; SMOLENSKY 1990 [98]). Pour prétendre représenter le sens des expressions du langage, des mécanismes non combinatoires comme le connexionnisme devraient eux aussi pouvoir engendrer la systématique, ce que cette polémique mettait en discussion. Le point crucial est que la systématique ne peut être due au hasard des connexions

neuronales, si ces connexions sont supposées être directement responsables des liens qui relient le support matériel des différentes représentations mentales (FODOR & PYLYSHYN 1988 [41] ; FODOR & MCLAUGHLIN 1990 [40]). La propriété de systématisme exclut toute idée de lacune ou de panne partielle. Si la pensée correspondant à la phrase Jean aime Marie peut être produite, alors il y a 100% de chances que la pensée correspondant à la phrase Jacques aime Marie et la pensée correspondant au prénom Marie puissent être produites, non 99% ou 99.9%. Or, un dispositif qui représente les liens entre les constituants de ces représentations par des connexions matérielles ne peut assurer ces 100%. La raison fondamentale qui empêcherait les dispositifs connexionnistes d'assurer la systématisme serait leur manque de constituance. Sans constituance, les représentations ne peuvent que se déclencher les unes les autres, de manière causale, par les liens physiques qui les relient, alors que dans un dispositif doté de la propriété de constituance, une représentation a accès, de manière certaine, à ses propres composants. Le débat peut alors porter sur la question de savoir si le connexionnisme est en mesure d'assurer la propriété de constituance (CHALMERS 1993 [12]).

Dans cette partie, nous voulons explorer les conséquences de l'option combinatoire. Si le calcul du sens est un calcul combinatoire et que le fonctionnement du cerveau est, à cet égard, le même que celui d'une machine de Turing particulière, alors les propriétés de productivité, de systématisme, de compositionnalité et de constituance peuvent être garanties sans problème. Ce calcul combinatoire et les représentations symboliques sur lesquelles il porte constituent un authentique langage mental.

6.4. Le langage mental

Le projet d'un modèle combinatoire de la sémantique est de reproduire le jugement d'acceptabilité sémantique d'un énoncé quelconque d'après sa structure et la nature de ses éléments. Or, nous avons eu l'occasion de montrer qu'un projet analogue, concernant le jugement épistémique, pouvait être critiqué. La critique portait sur l'existence d'énoncés tautologiques non triviaux, dont la validité puisse être à coup sûr déterminée d'après les éléments de l'énoncé (CF. CHAPITRE 4). Si le jugement épistémique à propos de tout énoncé nécessite un accès à des connaissances externes, ne risque-t-on pas d'avoir une situation similaire dans le cas du jugement sémantique ? Par exemple, pourquoi, dans la plupart des contextes, la phrase Jean mange la pomme est sensée, alors que la phrase Jean mange la porte ne l'est pas ? Est-ce la signification du verbe manger qui constitue ici l'élément crucial, ou est-ce le fait contingent que les individus se nourrissent plus souvent de pommes que de portes ? Certaines approches logiques tentent de ramener le jugement sémantique à une question épistémique en notant que le premier énoncé, contrairement au deuxième, peut être vrai dans certains domaines d'interprétation (CF. CHAPITRE 5). Ce faisant, elles n'offrent en guise de signification qu'un critère relationnel contingent. Il est plus plausible de dire que, pour juger de l'acceptabilité sémantique des énoncés, nous effectuons un calcul combinatoire sur la signification des mots. Autrement dit, le jugement sémantique ne se réduit pas à une simple vérification dans une base de données. Lorsque nous jugeons les phrases Jean entre dans la pièce ou le calme entre dans la pièce comme sémantiquement acceptables, alors que, dans la plupart des contextes, la phrase le salon entre dans la pièce ne l'est pas, nous ne vérifions pas dans une base de données les occurrences des entités qui peuvent "entrer" dans d'autres entités ; nous effectuons un calcul sur la signification des mots.

L'approche combinatoire du calcul du sens ne se limite pas à reproduire la performance sémantique à l'aide d'un calcul formel. Elle prend ce calcul "au sérieux" et s'interroge sur sa nature. Si l'on adopte cette approche, il devient essentiel de préciser les propriétés de ce calcul

et les représentations sur lesquelles il porte. Si la production et la compréhension des énoncés se réalisent par une combinatoire de symboles, quelle est la nature de ces symboles, correspondent-ils à des pensées élémentaires, et ces pensées sont-elles en bijection avec les mots de la langue ?

Si la construction du sens s'effectue par un calcul combinatoire, nous avons affaire à un langage mental, parfois dénommé mentalais (FODOR 1975 [34]). L'existence du mentalais est une hypothèse forte sur le plan cognitif. Il ne s'agit pas simplement de dire que l'état du monde décrit par le langage naturel peut être aussi décrit par un langage formel plus précis, mais bien que l'esprit humain manipule des représentations structurées en utilisant des mécanismes combinatoires systématiques. L'hypothèse suppose que les opérations qui agissent sur ces représentations, par exemple pour produire des inférences, opèrent uniquement d'après les caractéristiques structurelles des représentations. On retrouve donc, au niveau du mentalais, ce qui, au niveau syntaxique, fait d'une langue comme le français ou le persan un mécanisme combinatoire productif et systématique. Contrairement aux langues, toutefois, le mentalais est supposé universel, ce qui permet de traduire un énoncé d'une langue à l'autre en préservant son sens. L'hypothèse de l'existence du mentalais se situe dans le cadre de l'approche fonctionnaliste en sciences cognitives. Si le calcul du sens est réellement effectué sur des bases combinatoires, alors peu importe le fait que ce soit un cerveau qui serve de substrat à ces computations (FODOR 1975 [34]). L'important est d'étudier la computation pour elle-même. Selon cette approche, l'essentiel pour comprendre l'élaboration du sens des énoncés est d'étudier le mentalais. Une bonne compréhension des lois de ce langage de la pensée, pour les partisans de l'hypothèse du mentalais, devrait permettre d'envisager sa reproduction matérielle à l'aide d'ordinateurs, qui seraient ainsi capables de produire exactement les mêmes interprétations que les êtres humains.

Le langage mental a pour fonction, entre autres, d'offrir une interprétation sémantique des expressions langagières, que nous avons appelée, dans la deuxième partie, une représentation conceptuelle. L'hypothèse du langage mental décrit la sémantique comme une double interprétation. Les expressions langagières sont interprétées dans le langage mental, sous forme de représentations conceptuelles. Mais le langage mental est lui-même un système symbolique, dont les constructions doivent être à leur tour interprétées. Cette interprétation se trouve dans le monde perçu. Ainsi, le contenu d'une expression langagière est une représentation conceptuelle ; le contenu d'une représentation conceptuelle, quant à lui, demeure dans notre interaction avec le monde.

Si le mentalais existe, ce que nous aurons l'occasion de discuter plus loin, alors une question essentielle est de déterminer la nature des représentations sur lesquelles il opère. Le modèle combinatoire suppose que ces représentations sont construites par la combinaison d'unités élémentaires, appelés concepts. Les concepts peuvent être définis par leurs propriétés représentationnelles et causales (CF. CHAPITRE 3). Les propriétés représentationnelles d'un concept constituent son contenu. Les propriétés causales d'un concept, quant à elles, sont liées aux mécanismes du calcul dans lesquels elle est impliquée, et lui confèrent une forme.

The representational properties of a creature's mental state are supposed to be sensitive to, and hence carry information about, the character of its environment. The causal properties of a creature's mental state are supposed to determine the course of its mental processes, and, eventually, the character of its behavior. [...] Concepts are the least complex mental entities that exhibit both representational and causal properties; all the others (including, particularly, beliefs, desires and the rest of the "propositional attitudes") are assumed to be complexes whose constituents are concepts, and whose representational and causal properties are determined, wholly or in part, by those of the concepts they're constructed from. (FODOR 1994 [36] p. 96)

Si les représentations conceptuelles sont constituées par les concepts qui les composent, comment définir la nature des concepts eux-mêmes ? Comme le système conceptuel ne peut, par définition, être observé directement, le problème peut être considéré à partir de ses trois interfaces : raisonnement, perception et langage.

On peut partir de l'idée selon laquelle les concepts sont les représentations conceptuelles les plus simples, sachant que les représentations conceptuelles sont définies comme les objets du langage mental qui ne peuvent pas recevoir de valeur de vérité. Il est habituel d'opérer une distinction entre les représentations propositionnelles qui peuvent recevoir une valeur de vérité, comme l'interprétation de la phrase le chat marche sur le mur, et les représentations conceptuelles proprement dites, qui ne le peuvent pas, comme l'interprétation du syntagme le chat noir du voisin. Cette distinction tire sa pertinence de l'hypothèse, admise dans certaines théories, selon laquelle seules les représentations propositionnelles peuvent être objet des attitudes propositionnelles, comme le désir ou la croyance, qui sont à l'œuvre dans les mécanismes de raisonnement (FODOR 1994 [36]). Notons que la distinction entre représentations propositionnelles et représentations conceptuelles est prédite par un critère linguistique : les syntagmes nominaux renvoient à des représentations conceptuelles, alors que les phrases sont interprétées comme des représentations propositionnelles. Cette distinction, cependant, ne nous aide pas à déterminer la nature des concepts en tant qu'éléments constitutifs des représentations conceptuelles, car elle n'a pas de conséquence sur la complexité structurelle des représentations correspondantes. Les deux formes de représentation, propositionnelle et conceptuelle, peuvent être incluses l'une dans l'autre. Par exemple, l'interprétation du syntagme l'homme qui croit que le chat marche sur le mur sera construite par la composition récursive des deux types de représentation. L'interprétation d'un syntagme nominal comme le chat noir du voisin, au même titre que celle de la phrase le chat marche sur le mur, doit nécessairement être une représentation conceptuelle composée, en vertu du principe de constituance, qui veut que le sens d'un syntagme contienne le sens des mots qui y apparaissent.

Si l'on cherche à caractériser les concepts à partir de leur contenu, on peut postuler que les concepts correspondent à des catégories perceptives. Dans le cadre d'un langage mental combinatoire, le problème peut être simplifié en partant de l'opposition entre concepts primitifs et concepts complexes. Dans certains modèles, les concepts primitifs sont ceux qui sont directement liés aux données de la perception, comme les couleurs, ou aux catégories de base de la cognition, comme l'espace ou le temps. Du fait de leur lien privilégié avec l'expérience, l'acquisition de ces concepts primitifs est souvent considérée comme un problème à part. Tous les autres concepts seraient des concepts complexes, ce qui signifie qu'ils sont acquis par des mécanismes inductifs généraux. Notons que le caractère primitif d'un concept peut être considéré comme différent de sa structure simple ou complexe. Rien n'interdit qu'un concept réputé primitif comme ROUGE corresponde à une représentation conceptuelle composée, alors qu'un concept considéré comme appris, par exemple CARBURATEUR, ne se révèle doté d'aucune structure.

Pour définir la nature des concepts élémentaires, il est tout aussi légitime de partir de la troisième interface, en considérant leur correspondance avec le langage. Nous sommes confrontés à deux systèmes symboliques, le langage proprement dit et le langage mental. Le langage mental, en tant que système symbolique, ressemble sur certains points au langage proprement dit. Il possède sa grammaire : les règles de combinaison conceptuelles. Il possède ses phrases, les constructions conceptuelles, et ses mots, les concepts. Les mots du langage naturel sont les plus petites unités susceptibles d'être à la fois interprétées et manipulées par la combinatoire grammaticale. Par analogie, les concepts apparaissent comme les plus petites unités du langage mental ayant un contenu et pouvant entrer dans des combinaisons conceptuelles. Il s'agit alors d'établir la correspondance entre ces deux systèmes. Selon cette

idée de correspondance, chaque fois que nous faisons un jugement sémantique, c'est-à-dire que nous donnons un sens à une expression langagière, nous lui associons une représentation conceptuelle⁵. La question de la nature des concepts élémentaires peut alors se poser ainsi : les mots du langage naturel, dans cette correspondance, sont-ils associés aux "mots" du langage mental ? Si les concepts lexicaux, associés aux mots du langage, sont définis de manière indépendante par leurs propriétés causales ou représentationnelles, cela n'a rien d'évident.

Le fait que les expressions langagières composés soient interprétés comme des représentations conceptuelles composées ne signifie pas que les mots qui les constituent correspondent à des concepts simples. Pour de nombreux auteurs, les concepts lexicaux correspondent, à leur tour, à des représentations conceptuelles composées possédant une structure récursive, puisque consistant en la combinaison d'autres représentations conceptuelles, un peu comme les molécules du vivant qui sont décrites comme des assemblages de molécules plus simples. Le dilemme est donc de savoir s'il existe une correspondance entre les mots et les concepts simples, ou si les concepts lexicaux sont pour la plupart des représentations conceptuelles composées. Dans le premier cas, les concepts verbalisés par les mots seront les éléments les plus simples du langage mental ; dans le deuxième cas, il faut admettre l'existence de concepts ineffables, indicibles.

Dans les deux chapitres qui suivent, nous allons tour à tour traiter les deux options. Nous considérerons tout d'abord l'approche moléculariste, qui voit les concepts lexicaux non comme des entités minimales, mais comme des représentations conceptuelles composées, récursives. Ensuite, nous aborderons l'option selon laquelle les concepts lexicaux sont d'authentiques représentations minimales du langage mental, des concepts atomiques.

Conclusion

L'approche symbolique computationnelle semble présenter un cadre approprié et prometteur pour expliquer les phénomènes systématiques de la sémantique du langage naturel. Le calcul du sens, effectué au sein d'un système conceptuel de ce type, consiste à former des expressions dans un langage mental. Dans le reste de cette partie, nous explorons les conséquences de cette hypothèse de manière à en cerner les atouts et les limites. Notre analyse se focalise sur une question précise. Si le système conceptuel possède la forme d'un langage mental, quelle est la nature des représentations de ce langage qui sont associées aux entités lexicales, dans une langue humaine quelconque ? Ces représentations, les concepts lexicaux, peuvent être soit des représentations moléculaires, possédant une structure récursive, soit des représentations atomiques, dépourvues de toute structure interne. Après avoir détaillé les différents moyens qui ont pu être proposés pour représenter les concepts lexicaux, que ce soit sous forme moléculaire ou sous forme atomique, nous montrerons qu'un certain nombre de paradoxes se présentent dans chaque cas. La liste de ces paradoxes nous conduira à penser que l'approche symbolique computationnelle, introduite dans le présent chapitre, conduit à une impasse.

⁵ Le fait que ces deux systèmes symboliques sont supposés coexister s'accorde bien avec l'hypothèse de modularité.