A propos de la mathématisation du langage : linguistique, informatique théorique et biologie

Alain LECOMTE
UMR-7023 CNRS et Université Paris 8 – Vincennes – Saint-Denis

1 Introduction

1.1 La mathématisation de la syntaxe dans les années cinquante

Le processus de mathématisation de la linguistique au milieu du XXème siècle a surtout eu lieu autour de Noam Chomsky, dont l'article « Trois modèles de description du langage », daté de 1956, a eu beaucoup de retentissement et, dans une moindre mesure, de Yeshoshua Bar-Hillel, qui devait jouer plus tard un rôle important dans l'évaluation des programmes de recherches en Traduction Automatique. Ces deux auteurs sont au départ de deux traditions en modélisation de la syntaxe qui, comme nous le verrons, ne cessent de se croiser et de s'entrecroiser : les « grammaires syntagmatiques » et les « grammaires catégorielles »[1]. Les modèles issus de ces deux traditions ont été beaucoup exploités, notamment en linguistique informatique où ils ont donné lieu à des réalisations intéressantes et ont été plus ou moins acceptés par les linguistes. En ce qui concerne Chomsky, bien sûr, on peut dire que la syntaxe générative a connu une expansion considérable, même si elle s'est parfois diversifiée en plusieurs théories. Néanmoins, ce développement s'est fait au prix de l'abandon d'un objectif de mathématisation qui était implicite au début de la démarche. En ce qui concerne les grammaires catégorielles, on doit avouer qu'elles n'ont guère touché, pendant longtemps, qu'une minorité de linguistes, et presque uniquement dans le monde anglo-saxon (Steedman, 1985, 1987, Dowty, 1988, Jacobson, 1996), mais elles ont connu un mouvement de formalisation beaucoup plus important, en particulier dans la lignée des travaux de Jim Lambek (cf. Lambek, 1958, 1961), sur lesquels nous reviendrons plus loin[2].

Le modèle génératif de base des années soixante et soixante-dix a connu de nombreux bouleversements provoqués par Chomsky lui-même, jusqu'à ce qu'on appelle « le Programme Minimaliste » (Chomsky, 1996) où certains chercheurs (cf. Vermaat, 1999, Lecomte & Retoré, 2001, Stabler, 1997, 2000, Retoré & Stabler, 2004) ont vu une sorte de variante de l'autre modèle, celui des grammaires catégorielles. Entre temps, certains linguistes, comme Gerald Gazdar (Gazdar, Klein, Pullum, Sag, 1985) ont prôné une sorte de retour à l'orthodoxie en proposant un modèle très formalisé des grammaires syntagmatiques. Dans les années quatre-vingt, a été ainsi créé le modèle dit « des grammaires syntagmatiques généralisées », mais au cours de la décennie suivante, cet effort lui-même s'est trouvé en

[1] Nous ne donnerons pas ici d'exposé didactique présentant ces deux modèles. Le lecteur intéressé est prié de se reporter à l'excellent article en deux parties que C . Retoré a publié dans La Gazette des Mathématiciens, en particulier le volet 1, dans le numéro 115, de janvier 2008, de cette revue.

[2] Il faut également préciser que des travaux sur des formalismes très voisins des grammaires catégorielles se faisaient en Union Soviétique dans les années soixante, en particulier sous l'influence de Sebastian Shaumyan (cf. Shaumyan, 1987), qui lui-même reprenait des idées de H. Curry (Curry, 1961). Ces formalismes sont dits « applicatifs » au sens où ils s'appuient sur des mécanismes qui ont été réalisés dans des langages de programmation comme LISP, où tout, en principe, se ramène à des définitions de fonctions et des applications d'objets à d'autres objets. En France, cette tendance a été incarnée dans les travaux de J.P. Desclées (cf. Desclées, 1990) et de son école. Les idées de Curry ont à vrai dire fait école jusqu'à aujourd'hui : c'est en se basant sur elles que P. de Groote, S. Pogodalla et M. Kanazawa ont développé le formalisme général des « Grammaires Catégorielles Abstraites » (ACG) et démontré, dans ce cadre, des théorèmes d'expressivité et de complexité d'analyse très importants (cf. de Groote, 2001, de Groote & Pogodalla, 2004, Kanazawa, 2006).

quelque sorte dévoyé, les grammaires syntagmatiques généralisées s'étant développées en « grammaires syntagmatiques dirigées par les têtes » (*head-driven phrase structure grammars,* Pollard et Sag, 1987, 1994), où, là encore, le lien avec les grammaires catégorielles s'est avéré patent (cf. C. Pollard,1988). Cela étant, l'apparition de modèles en linguistique ne semble pas avoir eu le même effet que ce qu'a pu avoir par exemple l'introduction de théories mathématisées dans le monde de la physique : les linguistes non « formels » et/ou non « computationnels » ne semblent pas en avoir été bouleversés au point de se sentir contraints à les adopter. Il y a des critiques sur ces modèles et sur les conceptions de la langue qu'ils véhiculent et c'est cela plus spécifiquement qui sera l'objet de cette contribution, avec les réponses que nous apporterons à ces critiques. Nous commencerons par évoquer deux objections majeures à la mathématisation de la linguistique. La première repose sur l'idée qu'elle serait en réalité une sorte de plaquage de concepts mathématiques sur une réalité langagière qui ne les nécessiterait pas pour être adéquatement pensée, et que, plus grave encore, une certaine notion de langue aurait été inventée à seule fin de permettre cette mathématisation. La seconde, concernant plutôt le volet informatique (ou « computationnel ») consiste en ceci que nous serions contraints, à cause d'elle, de faire dépendre une théorie du langage d'une théorie du fonctionnement de l'esprit (ici en l'occurrence assimilé à un ordinateur).

1.2 La langue est-elle une invention des linguistes à des fins de mathématisation ?

Les travaux auxquels il est fait référence ci-dessus s'inscrivent dans un tournant historique fondamental : celui qui accompagne les avancées importantes de la logique au cours de la première moitié du XXème siècle, avec les grandes découvertes de Church, Turing, Gödel, Tarski, Curry etc. lesquelles ont servi de fondements à l'éclosion de l'informatique. Le premier Chomsky, tout linguiste spécialiste de l'hébreu qu'il fut, participait de ce travail formel, comme en témoigne d'ailleurs la mise au point de la fameuse hiérarchie des langages, dite « de Chomsky-Schützenberger » que l'on enseigne toujours aux débutants en informatique. De même, Bar Hillel est considéré comme un pionnier de la traduction automatique et un spécialiste de théorie de l'information. Il n'est donc pas étonnant qu'ils aient considéré qu'une dimension de la langue relevait du même traitement que celui appliqué à un langage formel. Reste évidemment à cerner cette dimension. Selon l'importance que l'on accorde à cette composante formelle de la langue, on sera tenté d'accorder plus ou moins d'importance à leurs travaux.

Une des difficultés centrales de l'approche logico-mathématique de la langue est, bien entendu, qu'elle semble substituer un objet totalement construit et artificiel à un objet dit « naturel ». La remarque la plus fréquemment entendue dans la bouche de linguistes non formalistes est : « voyons, on ne parle pas comme ça »[3]. On peut leur objecter que les planètes ne décrivaient pas non plus de véritables ellipses au temps où Kepler a formulé ses lois (ni évidemment maintenant !) et qu'en physique, le problème des trois corps est indécidable. On peut conclure ainsi que les lois de la mécanique céleste sont aussi, d'une certaine façon, un objet construit et artificiel (un modèle mathématique) que l'on a mis à la place du réel et qui

[3] S. Auroux, dans l'ouvrage que nous citerons plus loin reprend cette critique d'une manière plus élaborée en faisant remarquer « qu'aucune langue formelle construite jusqu'ici ne se parle » (p. 57) et qu'en conséquence, si on adopte le point de vue formaliste, il faut ou bien admettre que le fait d'être parlée n'est pas une propriété pertinente des langues naturelles, ou bien dire qu'il n'y a aucun obstacle à la construction d'une langue formelle qui se parle. Le poids de l'argument réside évidemment dans ce que l'on entend par « parler ». On peut ainsi objecter que les systèmes de dialogue développés dans certains laboratoires d'informatique visent bien à créer des « langues formelles qui se parlent », en un sens purement « phonétique », et qu'il a bien existé dans le passé au moins une langue qui n'était pas « parlée », au sens d'une parole usuelle, à savoir le sanskrit.

ne permet pas de reproduire le mouvement réel des astres. Comme le rappelle opportunément S. Auroux, Chomsky revendique souvent une position qu'il qualifie de *galiléenne*, c'est bien dire la part d'idéalisation de l'objet « langue » qui sous-tend son programme de recherche.

Il semble pourtant qu'il y ait une différence entre la situation de la physique et celle de la linguistique. Cette différence nous fait penser à celle qui existe entre deux figures de style : la *métaphore* et la *catachrèse*. Dans la métaphore, pour parler d'objets qui ont un nom bien défini dans un domaine donné, on emploie les mots d'un autre domaine. Dans la catachrèse, on utilise les termes servant à désigner une réalité pour en désigner une autre, mais c'est alors tout simplement parce qu'il n'existe pas déjà de termes pour désigner cette dernière, comme cela est le cas pour *les ailes* d'un avion par exemple. Le vocabulaire utilisé dans la catachrèse fait force de loi : on doit l'utiliser puisqu'on n'a rien d'autre, alors que celui utilisé dans la métaphore n'a pas cette force, on ne l'utilise qu'aussi longtemps que « l'image nous plaît ». Ainsi les sciences physiques emprunteraient-elles aux mathématiques parce qu'il n'y a pas d'autre « langage » pour désigner les objets dont elles traitent, et que, surtout, aucun autre langage ne leur permettrait d'atteindre un degré suffisant de précision, alors que les sciences du langage au contraire pourraient se passer du langage des mathématiques. A l'appui de cette objection, on affirmera par exemple que les analyses grammaticales sont des exercices vieux de plusieurs millénaires (au moins deux !) qui n'ont pas attendu les mathématiques pour être effectuées. On peut dire cela autrement : alors qu'en physique, il est admis depuis Bachelard que les mathématiques constituent « le réel de la physique », ou que, comme aurait dit Althusser, il y a un « rapport de constitution » entre les objets des mathématiques et ceux de la physique, il n'est pas sûr, à première vue, qu'il en aille de même pour ce qui est des objets de la linguistique, comparés à ceux de la logique et/ou des mathématiques. Comme le fait remarquer justement Sylvain Auroux dans « La raison, le langage et les normes » (1998), il y a toujours déjà des entités linguistiques théorisées (que ce soit dans un savoir *épilinguistique*, au sens de Culioli[4], ou dans un savoir *métalinguistique*), et la mathématisation ne ferait que « rajouter une couche ». Plus élégamment dit :

> *Il faut comprendre que la théorie linguistique est une sorte de théorie$_2$, qui représente formellement une théorie$_1$, intrinsèque à la façon dont la langue elle-même représente ce qu'elle représente et, notamment, des structures du réel.*
>
> S. Auroux (p. 67)

Dans ce sens, il y aurait une possibilité effective que les outils mathématiques soient artificiellement plaqués sur des entités théoriques déjà existantes, ne lui apportant de ce fait, rien d'autre qu'un semblant de scientificité. Le but de cette contribution est pourtant d'essayer de prouver le contraire.

[4] Dans une interview de janvier 2003 (http://www.shs.univ-paris5.fr/html/LFHPPR4X3Y5V5YWG.shtml) recueillie par Evangelia Adamou, la linguiste Josiane Boutet déclare : « Culioli enseignait qu'il fallait postuler l'existence d'une activité épilinguistique, par définition invisible ; la seule chose visible étant des marques de cette activité. Une activité épilinguistique dont on n'a jamais de marque tangible alors que l'activité métalinguistique peut se dire dans un discours métalinguistique. Pour Culioli on ne pouvait comprendre l'acquisition du langage si on ne faisait pas l'hypothèse que le bébé soumet en permanence le flux de discours auquel il est exposé à une activité mentale. Cette activité n'est pas une activité de répétition, ni de reproduction, elle n'est pas non plus seulement une activité statistique (elle l'est un peu du point de vue phonologique), mais c'est une activité d'analyse. Comme il ne voulait pas l'appeler activité métalinguistique, puisqu'elle ne donne pas lieu à des discours métalinguistiques, il l'a appelé activité épilinguistique »

1.3 La mathématisation dans les sciences du langage est-elle tributaire d'une théorie computo-représentationnaliste ?

Sylvain Auroux (1998) tente de résumer la situation actuelle dans les sciences du langage au début de son ouvrage par ces mots (p. 3):

> *La grande question métaphysique aujourd'hui, en matière de langage, est de savoir si un modèle computationnel suffit à expliquer le comportement linguistique humain. Est-ce que la faculté de langage de chacun d'entre nous et, plus spécialement, sa compétence linguistique s'explique parce que nous avons des algorithmes implémentés dans notre tête ?*

Il s'agit bien là d'une question épistémologique fondamentale, mais est-elle bien pertinente pour notre propos ? Est-ce que la question de la mathématisation en linguistique se ramène à elle ? si nous continuons sur le chemin d'une comparaison entre le domaine de l'univers et celui du linguistique, il nous faut constater qu'au cours du XXème siècle, il ne serait jamais venu à l'esprit des savants physiciens ou des philosophes commentant leurs travaux de se demander si « le fonctionnement de l'univers s'explique parce qu'il y a des processus de calcul implémentés quelque part ». A moins bien entendu de référer à un Dieu chef d'orchestre absolu, auquel cas tous les physiciens étaient d'accord pour mettre cette question à part, et pour la déléguer aux croyances de chacun, en dehors de la science et de l'épistémologie en tout cas. Aujourd'hui, il est vrai, cette question revient, lorsque certains cosmologistes comparent l'univers à un vaste ordinateur… mais *c'est une autre histoire* et elle repose bien sûr sur une autre base.

On ne demande pas qu'il y ait un centre de calcul quelque part quand on décrit un phénomène au moyen d'un système d'équations différentielles, et si un informaticien parvient à implémenter une théorie linguistique, sous la forme par exemple d'un analyseur syntaxique ou bien d'un programme de traduction automatique, on ne demande pas *nécessairement* que « notre tête » renferme le même programme…

Disons pour clore provisoirement ce sujet qu'on ne saurait se limiter à une conception de la grammaire liée au contexte des années soixante pour critiquer la mathématisation dans les sciences du langage, c'est-à-dire une représentation datée du programme de la grammaire générative, qui considérait alors que 1) le but de la grammaire était de montrer comment les phrases sont « engendrées » - dans le sens très retreint donné à ce mot par le Chomsky des années soixante[5], pas dans le sens de (simplement) « produites » - et que 2) ce mécanisme d'engendrement était réaliste au sens où il aurait reflété la façon dont les choses se passent dans nos têtes[6]. Si on en reste à Chomsky lui-même, il y a belle lurette que le linguiste du MIT a rejeté les « règles hors-contexte » au rang d'artefacts, que l'on ne parle plus de « structure profonde » et que l'on est prudent sur le réalisme cognitif. Et bien sûr, si on dépasse la seule pensée chomskyenne pour se référer aux travaux d'autres linguistes, de Keenan à Pollard, en passant par Pullum, Szabolcsi ou Stabler, on verra que cette conception « générative » au sens étroit joue peu de rôle. Ces linguistes « formalistes » ont bien autre chose en tête que vouloir à tout prix décrire comment s'engendrent toutes les phrases d'une langue : ils veulent tout simplement, à la façon dont œuvre n'importe quel scientifique dans un domaine donné rendre compte de certains phénomènes régulièrement observés, comme les phénomènes de portée des expressions quantifiées, les phénomènes de formation d'interrogatives ou bien les mécanismes d'occurrence des items de polarité. A cette fin, ils

[5] C'est-à-dire au sens où le langage engendré par une grammaire G est l'ensemble des mots w appartenant à V_T^* (V_T étant le vocabulaire terminal) tels que S $\Rightarrow^*$ w, où S est l'axiome de la grammaire et « $\Rightarrow^*$ » la relation de dérivabilité.

[6] Autrement dit « nous aurions des arbres dans la tête »…

déploient tout l'attirail de concepts logico-mathématiques dont ils disposent. On voit alors apparaître le fait que des notions comme celles de *fonction*, de *fonction monotone*, de *paire résiduée*, de *connexion de Galois*, *d'ensemble partiellement ordonné*, de *filtre* ou d'*idéal* ont *un rôle explicatif*[7].

1.4 Y a-t-il un substrat commun entre langue et mathématiques?

On ne saurait étudier le problème de la mathématisation en linguistique sans aborder la question du rapport des signes linguistiques avec les entités mathématiques. Y aurait-il un substrat commun entre langue et mathématiques ?

Que les mathématiques et le langage soient construits sur le concept de signe est une trivialité. Il ne faut pourtant pas en déduire trop vite une identité : bien sûr le signe linguistique et le signe mathématique ne sont pas identiques. Le premier découle d'un processus « naturel », au sens où la culture est un prolongement de la nature, d'où vient d'ailleurs la désignation de « langue *naturelle*[8] », et où, même si on reconnaît le caractère arbitraire du signe depuis Saussure, une telle reconnaissance ne se confond pas avec un conventionnalisme : nous n'avons pas accès à l'acte de baptême ou au contrat qui, un jour, a pu lier un signifiant à son signifié. A l'opposé, le signe mathématique semble être de pure convention (voir la manière dont s'énonce un résultat mathématique : soit T un espace topologique, soit $(O_i)_{i\in I}$ une famille d'ouverts etc…). D'autre part la syntaxe des mathématiques est formalisée entièrement. Elle définit un langage au sens formel du terme : ce langage ignore par principe et par construction toute ambiguïté.

S'il ne saurait y avoir identité, on peut cependant émettre l'hypothèse d'un recouvrement partiel des deux types d'entités. Le fait est que les relations entre nombre et langage semblent particulièrement intriquées, comme le prouvent les recherches récentes en sciences cognitives, notamment celles de Susan Carey ou d'Elisabeth Spelke (cf. Carey, 1998, 2004, Spelke 2000, 2001). Selon les neuro-psychologues spécialistes du calcul numérique, nous avons un certain sens inné des « numérosités » (que nous ne trouvons d'ailleurs pas que chez l'homme, mais aussi chez beaucoup d'espèces animales) préalable au développement de la faculté de langage. Ce sens conduit à des calculs approximatifs. Le calcul exact viendrait, selon Susan Carey (2004), d'un processus de *bootstrapping* qui mettrait ensemble ces dispositions très précoces et les débuts de développements du langage : l'apparition et la maîtrise des quantifieurs (*tout, quelques, certains*…) servirait ici de modèle pour que puissent se développer notre usage exact du nombre.

Si nous examinons les choses dans l'autre direction, il apparaît bien vite que *les unités linguistiques* se prêtent à des opérations de comptage, dénombrement, classification qui en font, semble-t-il très tôt dans l'histoire, *l'objet d'une mathématique*. Quoi de plus fascinant par exemple que les solutions trouvées par Panini vers le Vè siècle avant l'ère chrétienne pour présenter les règles du sanskrit sous la forme la plus « économique » et la plus aisément mémorisable par les élèves *pandit*, solutions validées au moyen d'outils mathématiques pointus par W. Petersen en 2003 (cf. Petersen, 2003)?

Cet exemple illustre une caractéristique des mathématiques elles-mêmes : à savoir le fait, souvent perdu de vue, qu'elles sont là avant tout pour résoudre des problèmes issus de nos pratiques (rappelons ainsi l'exemple des paris dans le cas de Pascal et Fermat) à propos

[7] je n'entre pas ici, faute de place, dans un débat sur « l'explication » dans les sciences. Disons simplement que ces notions sont explicatives au sens où on dit en physique que la notion de groupe a un rôle explicatif en chromodynamique quantique, cf. Cohen-Tannoudji et Spiro, 90, p. 231 : « la chromodynamique quantique est la théorie de jauge non abélienne qui utilise comme invariance de jauge le groupe SU(3) des changements de couleur des quarks ».

[8] les pointilleux pourraient à juste titre s'insurger d'une telle dénomination et réclamer que l'on parle plutôt de « langue culturelle »

d'objets que nous appellerons des objets *pré-mathématiques* (en ce qu'ils semblent se prêter « naturellement » à un traitement mathématique). Il semblerait que les mathématiques traitent toujours au départ d'objets issus de la pratique humaine : *elles font la théorie des objets pré-mathématiques*. Cet exemple montre aussi qu'une activité pré (ou proto ?)- mathématique est présente dès la première grande entreprise de réflexion sur la langue.

1.5 Science humaine ou science biologique ?

Ainsi parce que l'activité langagière est inséparable d'une activité *épilinguistique* (la manière dont les locuteurs se représentent spontanément leur langue au travers d'un savoir inhérent à celle-ci) qui peut éventuellement se prolonger en une activité *métalinguistique*, elle embraye de manière quasi-automatique sur une activité mathématisante.

Mais n'est-ce pas confondre langage et représentation du langage ? Comme dit Auroux (p. 62) : « une propriété fondamentale du langage humain est que tout élément peut être utilisé comme son propre signe, être sa propre représentation ». Il peut sans doute y avoir langue sans grammairien (le sanskrit a sans doute un statut exceptionnel dans l'ensemble des langues passées et présentes), mais il n'y a pas de langue sans représentation d'elle-même, tout simplement parce qu'il n'y a pas de langue sans possibilité, toujours offerte au parlant ou à l'écoutant, *d'analyser la signification*. Or, analyser la signification se fait par un effort mental qui investit le savoir tacite que nous possédons sur notre propre langue.

De cela découlent deux conséquences importantes. La première est que si nous appelons « raisonnement grammatical » le processus réflexif par lequel nous analysons une portion de texte ou de discours afin d'en actualiser la signification de manière consciente, alors la mathématisation du linguistique concerne bien au premier chef ce « raisonnement grammatical », donnant ainsi un fondement à la recherche opérée par les spécialistes des grammaires catégorielles (Michaël Moortgat en premier lieu, cf. Moortgat, 1997) de ce qu'ils appellent « les constantes du raisonnement grammatical ». La deuxième est que, comme on peut s'y attendre depuis le début, il y a certaines différences entre l'application des mathématiques en linguistique et leur application en physique (par exemple). Si on admet généralement que les objets physiques existent indépendamment de la pensée qui les appréhende (thèse du physicalisme ordinaire), ce que nous venons de dire de l'inséparabilité en linguistique entre le signe et sa représentation semble nous contraindre en principe à abandonner un tel postulat en ce qui la concerne. C'est ce qui fait de la linguistique, spécifiquement, une « science humaine ».

Mais la linguistique n'est-elle qu'une science humaine ? Si cela était, au sens où nous l'avons entendu ci-dessus (et que l'on retrouve chez maints auteurs philosophes, tels que Habermas par exemple, cf. Habermas, 1987), il s'ensuivrait une différence majeure avec la physique. Dans cette dernière en effet il y a de nombreux cas où la mise en place d'un modèle mathématique a *précédé* la découverte effective d'entités (par exemple les *bosons*). Dans le cas de la linguistique, il semble que ceci serait quasi impossible : on ne risquerait pas de voir apparaître de nouvelles réalités à partir de modèles mathématiques postulés pour la simple raison qu'étant immergés dans le langage… nous serions les premiers informés de leur existence, bien avant le modèle mathématique qui pourrait en rendre compte !

Mais n'est-ce pas avoir une vision trop herméneutique de la linguistique ? La théorie « Principes et Paramètres »[9], en donnant un modèle (même s'il est rudimentaire) des paramètres permettant d'expliquer la diversité langagière du point de vue de l'ordre entre Sujet, Verbe et Objet a, en un certain sens, « prévu » l'existence de langues possédant l'ordre

[9] C'est-à-dire la théorie défendue par l'école chomskyenne à partir des années quatre-vingt selon laquelle une langue se développe à partir d'un ensemble de principes commun à toutes les langues, qu'on appelle « Grammaire Universelle » et de paramètres ajustés différemment pour chaque langue (comme « tête à droite ou à gauche »).

« Objet-Verbe-Sujet » même si ces langues, comme il est prévu aussi, sont extrêmement rares (le Hixkaryana[10]) (cf. Baker, 2001). Cela montre qu'il est également possible de concevoir le langage en tant que système que l'on peut mettre à distance et étudier de manière objective. En ce cas, il est perçu comme un système biologique et la recherche sur les langues ne se distingue pas d'une recherche sur les systèmes naturels dans leur ensemble. Les mathématiques pourront alors s'y appliquer ni plus ni moins qu'elles ne le font en biologie (nous reviendrons en fin de chapitre sur ce rapport à la biologie).

2 Les mathématiques et le pouvoir d'expression des langues

2.1 Des questions linguistiques et des questions mathématiques

Nous commencerons cette section par la citation suivante d'Edward Keenan et Dag Westerstahl (1997, Handbook of Logic and Language, p. 839) :

> *In the past 15 years the study of generalized quantifiers has deepened, considerably, our understanding of the expressive power of natural language. It has provided answers to questions which have arisen independently in language study, and it has raised, and often answered, new questions, ones that were largely inconceivable without the basic concepts of generalized quantifier theory. In turn, our new linguistic understanding has prompted some novel mathematical questions (and an occasional theorem) whose interest derives, in part at least, from their natural interpretation in a non-mathematical domain*

Par cette citation, nous sommes confrontés à deux questions: la première est bien sûr celle de l'applicabilité de concepts forgés dans le domaine des mathématiques et de la logique à un problème d'essence linguistique et la seconde, parfois sous-estimée, est celle de la pertinence de la recherche d'application des mathématiques (ou de la logique) à la linguistique, du point de vue du développement des disciplines formelles. Formulées de manière plus directe ces questions reviennent à demander :

1) est-ce que, oui ou non, l'utilisation de nouveaux outils mathématiques permet de faire des avancées dans la connaissance des phénomènes linguistiques ?

2) est-ce que, oui ou non, le développement d'outils logico-mathématiques dans le champ linguistique permet de faire surgir de nouvelles questions mathématiques, voire de nouveaux théorèmes, ce qui aurait un intérêt du point de vue de ces disciplines prises indépendamment de leur champ d'application ?

Prenons comme exemple ce que nous propose justement Keenan, avant de voir également ce que propose R. Bernardi sur le sujet des items de polarité et la notion de « constante grammaticale » chez Moortgat, puis de conclure avec un autre problème et un autre angle de vision, concernant les phénomènes de portée des quantificateurs, où nous verrons qu'il existe des concepts opérants particuliers dans les sciences du langage actuelles, ceux de *programme* et d'*exécution* de programme. Par le biais de ces deux notions, nous serons alors capables d'établir un lien entre la linguistique et non pas la physique (décidément « trop loin ») mais la biologie.

---

[10] Une langue parlée dans une région de l'Amazonie

2.2 Le problème des quantificateurs généralisés

L'origine des *quantificateurs généralisés* est à situer chez Montague (1974), lorsque le philosophe américain décide d'interpréter les syntagmes nominaux non plus comme de « simples » SN, auxquels la structure prédicative du verbe peut s'appliquer, mais comme des objets d'un « type » plus élevé qui sont susceptibles de s'appliquer eux-mêmes à la partie verbale de la phrase. Le concept en vient à être autonomisé et étudié pour lui-même par Barwise et Cooper en 1981 (Barwise et Cooper, 1981). Après quoi, il sera considérablement développé par van Benthem (1986), Westerstahl (1985), et Keenan (voir aussi Corblin, 2002). De quoi s'agit-il ? Pour faire bref, et afin de ne pas abreuver le lecteur de trop de détails, disons que, partant de la logique, il est possible de donner une lecture des quantificateurs qui sorte du domaine de la logique proprement dite (où ils sont restreints à seulement deux instances :$\forall$, l'universelle et $\exists$, l'existentielle). Pour cela, on remarque qu'une formule de logique comme (1) ci-dessous :

(1) $\forall x\ homme(x) \Rightarrow mortel(x)$

peut aussi s'écrire comme (2) :

(2) $\forall(\lambda x.homme(x) \Rightarrow mortel(x))$

faisant apparaître le quantificateur ($\forall$) comme une propriété de second ordre, c'est-à-dire un prédicat s'appliquant lui-même à un prédicat (mais qui se trouve dans une hiérarchie des types à la Russell juste en dessous), ou dit encore autrement : comme *une famille d'ensembles*. « Pour tout » est ainsi la propriété d'être une propriété que tous les individus de l'univers possèdent, ou bien : est représentable par une famille d'ensembles qui, tous, contiennent l'univers, or de tels ensembles il ne peut en exister qu'un, c'est l'univers lui-même, d'où :

(3) $[[\forall]] = \{D\}$ (où D est l'univers)

De la même manière,

(4) $[[\exists]] = \{X \subseteq D\ ;\ X \neq \varnothing\}$

A cette formulation simple, on en préfère cependant une autre, car dans le langage, la quantification n'est pas « absolue » : elle est toujours relative, non pas à « l'univers » de la logique abstraite, mais à un « univers » particulier. Par exemple (1) n'est pas pris comme référant aux entités du monde en général, mais comme étant à propos des hommes. Ce n'est donc pas la dénotation de $\forall$ que nous cherchons, mais celle de « $\forall x \in A$ » où A peut être par exemple « la collection des hommes ». (2) peut alors s'écrire plutôt :

(5) $\forall_A(\lambda x.B(x))$ (A : hommes et B : mortels)

ou, mieux encore :

(6) $(\forall(\lambda x.A(x)))\ (\lambda x.B(x))$

faisant apparaître cette fois $\forall$ comme une relation de prédicats, ou, extensionnellement, comme une *relation entre ensembles* (une famille de couples d'ensembles au lieu d'une famille d'ensembles). On peut également le décrire comme une fonction $[[\forall]]$ qui, à tout couple de prédicats (P, Q) associe « vrai » si $P \subseteq Q$ et « faux » sinon.

A partir de là naît une idée importante : celle selon laquelle, après tout, n'importe quelle fonction de ce genre (définie dans l'ensemble des couples de prédicats et à valeurs dans {vrai, faux}) pourrait servir à définir un quantificateur, lequel en ce cas, serait qualifié de « quantificateur généralisé ». De fait, on peut facilement montrer qu'un tel schéma convient à plusieurs expressions quantifiées de la langue, comme le montre le tableau suivant (voir aussi Corblin, 2002) :

Groupe nominal	interprétation ensembliste sur un univers D
tout homme, chaque homme	$\{A \subseteq D\ ;\ HOMME \subseteq A\}$
(au moins) un homme	$\{A \subseteq D\ ;\ HOMME \cap A \neq \varnothing\}$

aucun homme	$\{A \subseteq D ; \text{HOMME} \cap A = \emptyset\}$
au moins trois hommes	$\{A \subseteq D ; \text{Card}(\text{HOMME} \cap A) \geq 3\}$
trois hommes	$\{A \subseteq D ; \text{Card}(\text{HOMME} \cap A) = 3\}$

A partir duquel on tire que, par exemple, « au moins trois » est la fonction qui à deux prédicats P et Q d'extensions respectives [[P]] et [[Q]] associe « vrai » si $\text{Card}([[P]] \cap [[Q]]) \geq 3$ et « faux » sinon.

La question vraiment intéressante alors est celle de savoir quels sont, parmi tous les quantificateurs généralisés possibles, ceux qui sont effectivement réalisés dans une langue. *A priori, il n'y a pas de raison pour qu'une telle question reçoive une réponse formelle*. Or, les travaux des chercheurs précédemment cités, et en particulier de Keenan, ont justement prouvé qu'il existait une telle réponse à cette question. Les quantificateurs réalisés dans les langues humaines possèdent les propriétés d'*extension*, de *conservativité* et d'*isomorphie*, dont les définitions sont rappelées en annexe.

2.3 La question des items de polarité négative

Cette caractérisation des quantificateurs de la langue à partir de propriétés formelles resterait toutefois peu de choses si la théorie des quantificateurs généralisés ne parvenait pas à faire plus en nous donnant *la clé* de problèmes comme celui posé par les *items de polarité*. Depuis Klima (1964), on appelle ainsi (NPIs dans l'abréviation anglo-saxonne) des expressions qui sont assignées (*licensed*) par un déclencheur qui se trouve être canoniquement une expression négative comme *not* ou *ne…pas*. Par exemple, en anglais, on dira :

(7) *John hasn't **ever** been to Moscow*
(8) *John didn't see **any** birds on the walk*

et non:

(9) **John has ever been to Moscow*
(10) **John saw any birds on the walk*

de même qu'en français, on dira:

(11) *il n'a pas mangé **le moindre** vermisseau*
(12) *je ne crois pas qu'il ait **jamais** vu d'éléphant*

et non :

(13) **il a mangé le (un) moindre vermisseau*
(14) **je crois qu'il a jamais vu d'éléphant*[11]

Comme l'ont fait remarquer divers auteurs, il ne suffit pas de dire qu'un NPI est déclenché par *not* ou *ne…pas*, car on peut trouver des exemples d'introduction de ces items dans d'autres contextes. Citons par exemple :

(15) *aucun sportif n'a ramené **la moindre** médaille*
(16) *si **jamais** tu passes par Grenoble, n'oublie pas de me rendre visite*

[11] A distinguer bien sûr de : *je crois qu'il n'a jamais vu d'éléphant* qui, en tout état de cause, n'est pas la négation de (12) laquelle serait : *je crois qu'il a déjà vu un éléphant*

(17) *le seul espoir qui lui reste de* ***jamais*** *revoir son pays*

(15) et (17) montrent l'usage d'un item de polarité négative dans le contexte d'une expression quantifiante (***aucun*** *sportif, le* ***seul*** *espoir qui lui reste*), (16) montre un exemple d'un tel emploi dans l'antécédent d'une conditionnelle (mais pas dans le conséquent, on opposera ainsi : *si* ***jamais*** *tu passes à Grenoble, tu viendras me voir* à *si tu passes à Grenoble, tu viendras* ***jamais*** *me voir* !).

Ce que l'on a découvert grâce à la théorie des quantificateurs généralisés, c'est tout simplement (Ladusaw, 1979, Fauconnier, 1975) que *les items de polarité négative apparaissent dans les arguments des fonctions monotones décroissantes*. Si, d'une façon générale, nous représentons, comme nous l'avons fait jusqu'à présent, par Q la fonction associée à un déterminant quantificateur, laquelle a deux arguments : A et B, nous pouvons définir quatre classes particulières de déterminants : ↑MON↑, ↑MON↓, ↓MON↓ et ↓MON↑, selon que la fonction est croissante ou décroissante en son premier argument et croissante ou décroissante selon son second. A titre d'exemple, *aucun* est ↓MON↓, d'où (15), mais tout aussi bien : *aucun sportif qui ramène la moindre médaille ne laisse le public indifférent*, *tout* est ↓MON↑, ce qui donne :

(18) *tout pêcheur qui ramène le moindre poisson est acclamé*

mais interdit :

(19) **tout pêcheur ramène le moindre poisson*

Une autre question résolue par cette théorie est celle de l'ensemble des syntagmes nominaux qui peuvent apparaître après le *of* anglais dans des partitives telles que : *more than ten of John's cats* ou bien *all but two of his ten children*. Le problème peut être transposé au français en se demandant quels syntagmes nominaux apparaissent après *de* dans des partitives comme : *plus de la moitié de tes copains* ou bien *trois des élèves qui l'accompagnaient*. La solution fait intervenir la notion mathématique de *filtre*.
Ceci permet de répondre en partie à notre question 1 : oui, l'utilisation d'outils mathématiques permet de faire des avancées dans la connaissance des phénomènes linguistiques.

2.4 De quelques théorèmes de Keenan

Pour ce qui est de la question 2 maintenant, il faut nous référer aux travaux de Keenan, qui ont effectivement enrichi les mathématiques de quelques théorèmes. Donnons-en un exemple. La classe CONS est celle des quantificateurs conservatifs (cf. définition ci-dessus). A l'intérieur de cette classe, on peut distinguer deux sous-classes : INT (pour « intersectifs ») et CO-INT (pour « co-intersectifs »). Q est dit *intersectif* si et seulement si pour tous A, B, A', B' inclus dans le domaine D, si $A \cap B = A' \cap B'$, alors $Q(A)(B) = Q(A')(B')$, autrement dit si et seulement si Q(A)(B) ne dépend que de l'intersection de A et de B. Q est dit *co-intersectif* si et seulement si Q(A)(B) ne dépend que de la différence A – B. Les cardinaux ainsi que les déterminants *quelques*, *au moins n*, *au plus n* etc. sont intersectifs alors que *tous*, *tous sauf n*, *presque tous* sont co-intersectifs. Il est alors intéressant de se demander comment se situe la classe CONS par rapport à ses deux sous-classes. Peut-on en particulier sortir de CONS par composition booléenne ? Existe-t-il des quantificateurs conservatifs qui ne soient pas engendrés par composition booléenne à partir d'éléments de INT et de CO-INT ? C'est alors que vient un théorème de Keenan (cf. Keenan & Westerstahl, 97) :

Théorème : *pour tout domaine D, $CONS_D$ est la fermeture booléenne complète de $INT_D \cup$ $CO\text{-}INT_D$*

Autrement dit, les opérateurs de $\mathrm{INT_D} \cup \mathrm{CO\text{-}INT_D}$ engendrent *tous* les quantificateurs de CONS par composition booléenne. Ce théorème nous permet ainsi d'aller plus loin dans la connaissance de la structure des quantificateurs conservatifs.
Dans le même ordre d'idée, on est amené à différencier les déterminants selon qu'ils effectuent ou non des « tris » sur ce à quoi ils s'appliquent. Par exemple, un *quelques* portant sur une classe A donnée peut toujours être remplacé par un *quelques* portant sur tout l'univers. On peut passer de (20) à (21) ci-dessous :

(20) *quelques étudiants sont végétariens*

(21) *quelques individus sont étudiants et végétariens*

cette propriété ne s'applique pas au cas de *la plupart*. (22) ci-dessous n'a pas de traduction analogue à (21) :

(22) *la plupart des étudiants sont végétariens*

On dit que *la plupart* est un trieur inhérent (*inherently sortal*) et que *quelques* est un trieur réductible. Un autre théorème de Keenan permet de caractériser ces notions.

Théorème : *un quantificateur conservatif Q est un trieur réductible si et seulement s'il existe une fonction booléenne h de deux variables telle que pour tous A et B inclus dans le domaine D, Q(A)(B) = Q(D)(h(A, B))*

2.5 L'intégration des propriétés sémantiques dans une dérivation syntaxique

La conception sur laquelle nous avons basée notre aperçu des quantificateurs généralisés est, comme on sait, en harmonie avec les grammaires catégorielles (dont la source se situe dans l'article de Bar-Hillel cité en introduction) pour la raison simple que dans les deux cas, *la notion de fonction mathématique joue un rôle central*. Dans les grammaires catégorielles standard, les catégories complexes (comme `(sn\s)/sn`) sont interprétées comme des fonctions : `(sn\s)/sn` est interprété comme le type d'un objet syntaxique qui prend deux arguments de type `sn` avant de retourner une phrase (type `s`), autrement dit le type d'un verbe transitif. Dans la conception dérivée de Montague, et qui inclut Lambek (1958), les syntagmes nominaux, pour pouvoir jouer leur rôle de quantificateur généralisé, se voient dotés d'un type d'ordre supérieur, du genre `s/(sn\s)` ou `(s/sn)\s` (deux types qu'on peut obtenir à partir de `sn` dans le calcul de Lambek). Ces types sont homomorphes aux types qu'ont les fonctions qui les traduisent au plan sémantique (c'est-à-dire les fonctions utilisées pour la construction des représentations sémantiques) si on admet que les objets syntaxiques de type `sn` se traduisent par des entités (**e**), ceux de type `s` par des propositions susceptibles d'être vraies ou fausses (**t**) et si on admet qu'un type syntaxique A/B ou B\A se traduit par un type sémantique <t(B), t(A)> où t(E) est la traduction du type syntaxique E en un type sémantique[12]. Si on omet la médiation par le type sémantique (qui serait ici : <<**e**, **t**>, **t**>), on obtient une connexion directe entre le type syntaxique `s/(sn\s)` et la fonction du paragraphe précédent (que nous pouvons écrire : $\lambda Q.\forall x\, personne(x) \Rightarrow Q(x)$ dans le cas de *toute personne*, ou bien $\lambda Q.\forall x\, personne(x) \Rightarrow \neg Q(x)$ dans le cas de *personne*, ou de *nobody*). La fonction associée à *personne* étant *monotone décroissante*, marquons cela par une « modalité » (nous reviendrons sur cette idée plus loin) affectant le type de l'argument : `s/◊(sn\s)`. Nous sommes alors (Moortgat, 1997, Bernardi, 2002) dans un système dit « multimodal » régi par des règles qui sont en harmonie avec celles qui régissent les constructeurs de type binaire / et \.

[12] ceci est une première approximation, l'association systématique d'un type e aux entités sn pouvant être contestée, comme l'a fait Claire Beyssade (cf. Beyssade et Dobrovie-Sorin, 2005)

La paire {/, \} étant obtenue par résiduation à partir d'un produit •, on peut construire aussi par résiduation un couple de modalités dites duales tel que l'on ait :

(23) $\Diamond A \to B \Leftrightarrow A \to [\,]\, B$

Les règles d'introduction et d'élimination de ces opérateurs sont les suivantes :

[[] I]	**[[] E]**	**[◊ E]**	**[◊ I]**
$\dfrac{\langle\Delta\rangle^{-} \to A}{\Delta \to [\,]\, A}$	$\dfrac{\Delta \to [\,]A}{\langle\Delta\rangle^{-} \to A}$	$\dfrac{\Delta \to \Diamond A \qquad \Gamma[\langle A\rangle^{-}] \to B}{\Gamma[\Delta] \to B}$	$\dfrac{\Gamma \to A}{\langle\Gamma\rangle^{-} \to \Diamond A}$

On peut prouver que *nobody left yet* est « déductible », à la différence de **everybody left yet,* avec une catégorie pour l'adverbe *yet* définie comme : `[]◊(sn\s)\ []◊(sn\s)`. Le calcul consiste à trouver une bonne parenthétisation pour la phrase à analyser (*nobody <left yet>*$^-$) et à raisonner comme suit :

- si *left* est de type `sn\s` (celui d'un verbe intransitif), alors par application de [◊ I], <*left*>$^-$ est de type `◊(sn\s)`,
- par application de [[] I], *left* est de type `[]◊(sn\s)`,
- *left* peut donc être composé avec *yet*, ce qui donne : *left yet* de type `[]◊(sn\s)`, et donc :
- par [[] E] : <*left yet*>$^-$ est de type `◊(sn\s)`, et est donc composable avec *nobody*, donnant : *nobody <left yet>*$^-$ de type `s`.

Si un tel raisonnement n'existe pas pour **everybody left yet*, c'est parce qu'il n'existe pas de parenthétisation acceptable permettant de dériver cette phrase.

Une autre manière de présenter cette dérivation est la suivante :

Nobody *<left* *yet>*$^-$

`s/◊(sn\s)` `<sn\s` `[]◊(sn\s)\ []◊(sn\s)>`$^-$

`<[]◊sn\s`

`<[]◊(sn\s)>`$^-$

`◊(sn\s)`

`s`

L'idée est que les catégories peuvent se combiner en tenant compte d'un jeu de transformations unaires possibles sur elles. De cette manière, les propriétés sémantiques étudiées (ici la sensibilité aux contextes croissants et décroissants pour ce qui est de l'introduction d'items de polarité négative) deviennent intégrées à une construction (dérivation) syntaxique.

Que sont alors les construits formels pour rendre compte de cette « explication » ? Peut-on d'abord parler « d'explication » des phénomènes de la quantification et de la polarité ? Les phénomènes de polarité dans les langues semblent être largement répandus : s'il doit y avoir des *universaux*, n'est ce pas à ce genre de phénomène qu'il faut en donner le statut ? le fait que les items de polarité négative apparaissent dans certains contextes seulement, comme s'ils étaient déclenchés par l'action à distance d'un élément déterminé semble faire partie de leur définition elle-même. A partir de là, relier leur occurrence à celle d'éléments associés à des fonctions monotones décroissantes apparaît bien explicatif dans la mesure où le phénomène empirique (constaté dans un grand nombre de langues au moins) est relié à un dispositif abstrait, formel, qui existe indépendamment de ce qu'on étudie (la théorie des fonctions monotones).

En ce qui concerne l'intégration syntaxe-sémantique maintenant, on devra noter que les constantes de type utilisées (`s, sn, n`…) sont tout à fait secondaires. Il ne faut pas s'obnubiler sur le caractère statique temporaire du modèle : ce qui compte avant tout, *c'est la dynamique du système*, autrement dit, *plus que les symboles* (y compris ceux de constructeur : /, \, •, ◊, []), *les règles* de leur introduction et de leur élimination et la *dynamique* à l'œuvre dans ces règles, dont nous venons d'avoir un rapide aperçu avec la dérivation précédente, où on utilise le fait qu'un environnement représenté par < … >$^{-}$ est utilisé pour neutraliser (annihiler) une modalité [] au moment opportun.

On aura d'ailleurs remarqué la symétrie des règles utilisées : [[] I] et [[] E] (resp. [◊ I] et [◊E]) sont clairement symétriques : l'utilisation de l'une à la suite de l'autre ne ferait que nous reconduire à la situation de départ, mais ce n'est bien sûr pas de cette façon absurde que les symétries sont utilisées. Dans la dérivation précédente, on a une *séquence* d'applications de règles : ([◊ I], **[[] I]** , [\ E], **[[] E]** , [/ E]) où [/ E] est intercalé entre [[] I] et [[] E] pour passer de <left>$^{-}$ à <left yet>$^{-}$ autrement dit pour continuer la composition des expressions en mode *modalisé,* puisque la fonction associée à l'expression *nobody* exige ce mode sur son argument. Ainsi dans cette séquence, [[] I] et [[] E] fonctionnent comme des *marqueurs* qui délimitent une partie de texte qui se trouve analysée dans un mode spécifique.
La déduction peut encore se linéariser sous la forme d'une séquence « concrète » (par opposition à la séquence abstraite des applications de règles) :

(24) $\backslash^{-1}$ (nobody **[]**$^{-1}$ $\backslash^{-1}$ (**[]** ◊left yet))

où les symboles $\backslash^{-1}$, **[]**$^{-1}$ **, [], ◊,** ◊$^{-1}$ sont des formes abrégées des règles [\ E], [[]] E], [[] I], [◊I], [◊E] qu'on peut interpréter comme des opérateurs jouant sur des structures :

- ◊ : met un environnement < … >$^{-}$ et surcharge le type d'une modalité ◊,
- **[]** : enlève un environnement < … >$^{-}$ et surcharge le type d'une modalité [],
- ◊$^{-1}$: enlève un environnement < … >$^{-}$ et décharge le type d'une modalité ◊,
- **[]**$^{-1}$: met un environnement < … >$^{-}$ et décharge le type d'une modalité []
- $\backslash^{-1}$ et $/^{-1}$ concatènent les expressions,
- \ et / extraient des expressions de leur contexte

(24) contient alors l'expression avec en plus l'indication d'une suite d'opérations nécessaires pour la « décoder » correctement :

- surcharger *left* des modalités ◊ et [],
- concaténer avec *yet*
- décharger [], en conférant une propriété structurelle marquée par < … >$^{-}$
- concaténer avec *nobody*

R. Bernardi (2002) donne d'autres exemples de dérivations réussies comme *John doubts anybody came at all*, où <(anybody, came), at all>$^{-}$ est composé dans le mode approprié.

Nous en venons donc aux véritables « atomes » du raisonnement grammatical, qui ne sont ni les catégories, ni même les constructeurs, mais *les règles*. Et on peut de ce fait s'inscrire en faux contre l'idée que « les linguistes « formalistes », au contraire des physiciens, pren[draient] pour termes de leurs calculs les éléments de la langue naturelle eux-mêmes » (Auroux, p. 56), car bien évidemment, [[] I], [[] E], [◊ I] et [◊E] *ne sont pas* des éléments de la langue naturelle. On peut les voir comme des marqueurs d'opérations que le linguiste *ajoute* par rapport au réel pour en produire l'intelligibilité, comme un ensemble de concepts est ajouté en biologie moléculaire ou en physique des particules pour mieux comprendre la réalité. Bien entendu ces ajouts ne sont pas arbitraires : il ne s'agit dans aucun des cas de se contenter de « sauver les phénomènes ». La physique théorique sait se doter d'instruments de mesure pour vérifier que les entités ajoutées ne sont pas de pures vues de l'esprit (elles se manifestent par des effets que l'on a prévus initialement, ce qui conduit au constat qu'elles «existent » bel et bien), comme le linguiste formaliste sait se doter d'un calcul pour vérifier que les dérivations ont bien lieu par le jeu des règles ajoutées, lorsqu'on a équipé les expressions linguistiques des types appropriés. On notera alors que *les déductions*, sur le statut duquel nous nous interrogions (s'agissait-il des mécanismes eux-mêmes à l'œuvre dans la production ou dans la réception des expressions ?), trouvent dans cette conception leur fonction véritable : celle de *montages théoriques* visant à valider des hypothèses.

3 Linguistique et informatique théorique : une question d'évaluation

3.1 Le « calcul » : fantasme d'automate ou processus naturel

Comment « calcule-t-on » la portée des quantificateurs dans une phrase comme (25) ci-dessus ? Voilà bien une question susceptible de faire hurler un linguiste empiriste avant tout attaché au recueil d'exemples et contre-exemples. Une telle question est hérétique à ses yeux d'abord parce qu'elle postule une idée de « calcul », en général attachée à une conception computo-représentationnelle du langage, et ensuite parce qu'elle fait mention de la notion de « portée », qui évoque irrésistiblement la logique et son travers immédiat : le logicisme.
Il nous faut donc préciser le contenu d'une telle question, dans la ligne de ce que nous avons avancé jusqu'à présent.

En premier lieu, il y a des notions qui nous semblent indispensables parce qu'avant même d'avoir pris le temps d'examiner un très grand nombre de langues, nous savons que dans au moins un nombre non négligeable d'entre elles, elles répondent à des observations réelles. La présence de particules servant à la quantification (déterminants ou adverbes) semble attestée dans la plupart des langues connues (pour un autre horizon linguistique que les langues indo-européennes, on pourra s'intéresser à la particule *dou* en chinois mandarin). A partir de là se pose la question de la *portée* de ces éléments, c'est-à-dire de la zone de phrase affectée par eux.

En second lieu, la notion de « calcul » ne doit pas être nécessairement vue comme référant à un calculateur empirique (éventuellement réalisé par un cerveau humain) qui effectuerait méthodiquement chaque pas l'un après l'autre. L'hypothèse reste neutre quant au dispositif éventuel chargé de s'assurer que les choses se passent bien selon ce que prévoit le calcul. Comme nous le verrons plus loin, la dimension « calcul » renvoie à une dimension de la *conversion* (ou de *l'évaluation*). Pour reprendre les comparaisons que nous avons déjà évoquées, le passage de la recette à la confection effective du plat, comme celui de la structure en hélice de l'ADN à la synthèse effective des protéines, sont autant de « calculs » au sens où nous l'entendons ici. Autrement dit, *le calcul transcende les calculateurs*. Il est défini avant même que l'on conçoive les dispositifs concrets susceptibles de le simuler. Cela rappelle la situation de Turing avant l'informatique : les machines ont existé avant l'informatique, avec cette différence toutefois que les ordinateurs ont permis de *réaliser concrètement* les machines de Turing, (au moment de leur création, ces dernières étaient des objets théoriques en attente de leur réalisation matérielle, ce qu'on appelle d'ailleurs aujourd'hui des *machines virtuelles*), alors que dans ce qui nous occupe ici, dans le domaine du langage (et comme on le verra plus loin, en biologie), le calcul n'attend pas sa réalisation concrète. Tout juste pouvons-nous dire qu'il attend (peut-être) une *simulation*. Que celle-ci intervienne ou non, cela peut sembler de moindre importance : ce n'est pas cela qui va valider la théorie, tout au plus pourra-t-elle servir à quelques fins pratiques, au statut d'ailleurs pas toujours assuré (les fameuses « industries de la langue »).

De toutes manières, le dispositif formel (un automate) qui simulerait ce calcul ne serait peut-être pas très utilisable, car de deux choses l'une : ou bien il serait récursif et il est presque sûr qu'il ne serait pas une simulation, ou bien il ne le serait pas (certains calculs seraient « interminables ») et alors, il serait peu utilisable.

Nous touchons là à un aspect central, au cœur de la mathématisation des sciences du langage et sur lequel il va falloir revenir en détails : la langue, comme les autres processus vivants, comme aussi les processus qui ont lieu à l'échelle de l'univers, est le siège *de processus qui ne sont pas tous « convergents »*.

La logique sait bien traiter les processus convergents, mais les autres ? En physique, nous pouvons évoquer dans cet ordre d'idées les nécessités de la « renormalisation » : dans certaine situations où on devrait calculer des quantités finies, les intégrales qui entrent dans le calcul divergent. Le physicien intervient alors pour modifier la mesure d'intégration. Dans la langue, le moindre soupçon quant à l'acceptabilité d'un énoncé nous fait osciller entre plusieurs alternatives de sens… Nous avons alors des procédures rapides de réparation qui nous font passer en général à autre chose soit que nous ayons *choisi le sens* à accorder à l'énoncé problématique, soit que nous ayons décidé de le laisser *en suspens*. On notera ainsi que la poésie (principalement à partir du début du XXème siècle) semble souvent fondée sur cette possibilité de mise en suspens du sens, parfois à chaque mot. *L'ordinateur, lui, ne laisse pas en suspens le sens*.

Ces considérations n'empêchent toutefois pas d'examiner ce qui se passe lorsque les processus convergent, ou en tout cas sont maîtrisables : la phrase (25) certes nous conduit à une oscillation de significations, mais celles-ci sont au nombre de deux et sont clairement identifiables. De plus, un contexte ou un dialogue élémentaire devrait suffire à faire opérer un choix. Nous allons découvrir néanmoins que son interprétation découle déjà d'un processus où se montre une *bifurcation*.

3.2 « Solutions » au calcul

3.2.1 Solution 1 : des choix pour le moins « ad hoc »

Essayons de calculer la signification de :

(25) *tout habitant rencontre un employé de mairie*

La section précédente, qui concernait les items de polarité, s'appuyait sur le modèle catégoriel. Nous pourrions ici aussi poursuivre la discussion dans le même cadre. Toutefois, cela nous obligerait à traiter de détails techniques que nous préférons éviter ici. Nous allons donc plutôt nous appuyer sur l'autre modèle de référence, celui des grammaires syntagmatiques. Par là, nous serons plus proche de la démarche d'origine concernant le traitement de telles questions, qui fut celle de Richard Montague (1930 - 1971). On sait que Montague (1974) introduit l'idée que les SNs sont de type sémantique complexes <<**e**, **t**>, **t**> afin justement de traiter la quantification, d'où le titre d'ailleurs de son fameux article : « The Proper Treatment of Quantification in Ordinary English ». <<**e**, **t**>, **t**> est, rappelons-le, le type d'un foncteur qui prend pour argument un prédicat et retourne comme valeur une valeur de vérité. Nous donnerons toutefois une forme plus simple de la correspondance syntaxe-sémantique que celle que donnait Montague dans son texte des années soixante-dix, une forme plus proche de celle qu'en donnait Gerald Gazdar dans les années quatre-vingt, puis finalement qu'en donneront Irene Heim et Angelika Kratzer dans leur ouvrage « Semantics in Generative Grammar » (1998).

Dans la perspective Montague – Gazdar, dite « règle à règle », à chaque règle syntaxique, correspond une instruction de composition des représentations sémantiques. Ces derniers sont des λ-termes simplement typés (au moyen des types sémantiques construits au moyen des types primitifs **e** et **t**). Voici une grammaire élémentaire (« grammaire-jouet ») qui peut donner une interprétation sémantique d'une phrase telle que (25) ou (26). Elle est divisée en deux parties : une partie « syntaxe » et une partie « sémantique » (σ dénote alors la fonction qui, à un constituant X donné, associe son interprétation sémantique). Nous ajoutons une colonne de commentaires.

syntaxe	sémantique	commentaires
$S \rightarrow SN\ SV$	$\sigma(S) = (\sigma(SN))(\sigma(SV))$	La sémantique du syntagme nominal est appliquée à celle du syntagme verbal.
$SN \rightarrow Det\ N$	$\sigma(SN) = (\sigma(Det))(\sigma(N))$	La sémantique du syntagme déterminant est appliquée à celle du nom.
$SV \rightarrow V\ SN$	$\sigma(SV) = \sigma(SN) \circ \sigma(V)$	La sémantique du syntagme nominal **est composée** avec celle du verbe.

Cette grammaire fonctionne au moyen d'un petit « lexique » qui comprend :

V → /*rencontre*/, sémantique : $\lambda x.\ \lambda y.\mathbf{rencontre}(x, y)$

N → /*habitant*/, sémantique : $\lambda u.\ \mathbf{habitant}(u)$

N → /*employé de mairie*/, sémantique : $\lambda v.\ \mathbf{employé}(v)$

Det → /*tout*/, sémantique : $\lambda P.\ \lambda Q.\ \forall x\ P(x) \Rightarrow Q(x)$

Det → /*un*/, sémantique : $\lambda P.\ \lambda Q.\ \exists x\ P(x) \wedge Q(x)$

Si nous partons des mots de la phrase, selon une perspective bottom-up (dans ce qui suit représentée tête en bas !), nous avons la construction suivante :

Det : /*un*/ — N : /*employé de mairie*/
$\lambda P.\ \lambda Q.\ \exists x\ P(x) \wedge Q(x)$ — $\lambda v.\ \mathbf{employé}(v)$
résultat de l'application : SN : /*un employé de mairie*/
de sémantique : $\lambda Q.\ \exists x\ (\lambda v.\ \mathbf{employé}(v))(x) \wedge Q(x) > \lambda Q.\ \exists x\ \mathbf{employé}(x) \wedge Q(x)$

V: /rencontre/ — SN : /un employé de mairie/
$\lambda x.\ \lambda y.\mathbf{rencontre}(x, y)$ — $\lambda Q.\ \exists x\ \mathbf{employé}(x) \wedge Q(x)$
résultat de l'application : SV : /*rencontre un employé de mairie*/
composition des fonctions : $\lambda z.(\ \lambda Q.\ \exists x\ \mathbf{employé}(x) \wedge Q(x))((\lambda x.\ \lambda y.\mathbf{rencontre}(x, y))(z))$ ce qui donne par réduction : $\lambda z.(\ \lambda Q.\ \exists x\ \mathbf{employé}(x) \wedge Q(x))(\lambda y.\mathbf{rencontre}(z, y))$, d'où :
$\lambda z.(\exists x\ \mathbf{employé}(x) \wedge (\lambda y.\mathbf{rencontre}(z, y))(x)) >$
$\lambda z.(\exists x\ \mathbf{employé}(x) \wedge \mathbf{rencontre}(z, x))$

SN : /*tout habitant*/ — SV : /*rencontre un employé de mairie*/
$\lambda Q.\ \forall x\ \text{habitant}(x) \Rightarrow Q(x)$ — $\lambda z.(\exists e\ \mathbf{employé}(e) \wedge \mathbf{rencontre}(z, e))$
par application :
$(\lambda Q.\ \forall x\ \text{habitant}(x) \Rightarrow Q(x))\quad (\lambda z.(\exists e\ \mathbf{employé}(e) \wedge \mathbf{rencontre}(z, e))) >$
$\forall x\ \text{habitant}(x) \Rightarrow (\lambda z.(\exists e\ \mathbf{employé}(e) \wedge \mathbf{rencontre}(z, e)))(x) >$
$\forall x\ \text{habitant}(x) \Rightarrow \exists e\ \mathbf{employé}(e) \wedge \mathbf{rencontre}(x, e)$

mais ceci ne représente *qu'une seule lecture* ! De plus, on aura noté « l'anomalie » que constitue la contrepartie sémantique de la règle SV $\rightarrow$ V SN. Là où partout ailleurs, il faut appliquer un terme à un autre, dans l'application de cette règle, il faut *composer* les deux termes.

Une autre construction est proposée par Heim et Kratzer, dans l'esprit des grammaires génératives modernes, qui incluent la notion de déplacement : elle consiste à utiliser les opérations dites de « quantifier raising », c'est-à-dire de montée des quantificateurs vers le haut de la structure. Dans une phrase telle que (25), l'ambiguïté provient du fait que le SN quantifié objet peut monter (« virtuellement ») par-dessus le SN quantifié sujet ou au contraire rester dominé par lui. Chez Heim et Kratzer en réalité on n'utilise pas le modèle syntagmatique ci-dessus, on utilise les deux opérations fondamentales des grammaires de « l'ère minimaliste », à savoir la fusion (merge) et le déplacement (move).

3.2.2 Solution 2 : où c'est le cadre d'évaluation qui permet des lectures différentes

La solution 2 que nous allons proposer (empruntée à C. Barker et à P. de Groote) est en un sens une généralisation de la première. En effet, un trait remarquable du traitement qu'offre Montague de la quantification est son passage, pour les syntagmes nominaux, à un type sémantique complexe, qui n'est plus **e** mais <<**e, t**>, **t**>. Même les noms propres reçoivent un tel type. Dans ce cas, le nom propre n'est plus le prototype de la constante logique (le genre d'entité auquel on attribuerait sans difficulté le type **e**, celui des constantes individuelles du modèle), c'est sémantiquement un être bien plus complexe : un prédicat de prédicat. Autrement dit, se trouve associé à *Pierre* la propriété d'être une propriété possédée par Pierre, ce qui se note : λP. P(pierre). Ceci a été remarqué évidemment dès les années soixante dix. Mais une chose n'avait pas été suffisamment remarquée, c'est qu'entre la constante **pierre** et le prédicat de prédicat λP. P(**pierre**), il y a *un renversement de perspective*. **pierre**, en tant que constante, dénote l'argument auquel n'importe quel prédicat réalisé dans le contexte pourra s'appliquer.

Argument :	**Fonction** (**contexte**) :
pierre	dort
	écoute la radio
	regarde la pluie tomber par les carreaux sales de sa fenêtre

λP. P(**pierre**), au contraire, dénote une fonction qui va pouvoir s'appliquer à n'importe quel contexte.

Fonction :	**Argument** (**contexte**) :
λP. P(**pierre**)	dort
	écoute la radio
	regarde la pluie tomber par les carreaux sales de sa fenêtre

Ainsi la complexification du type sémantique du syntagme nominal s'accompagne-t-elle d'un mouvement au terme duquel le syntagme nominal est sémantiquement une fonction prête à s'appliquer à un contexte.

Cette vision des choses existe depuis longtemps en informatique. C'est elle qui est utilisée notamment dans la gestion du contrôle. Si une interruption intervient au cours d'une suite de calculs, il faut immédiatement capturer la suite des calculs encore à faire, cela suppose qu'à chaque instant, un argument de la fonction calculée soit cette suite de calculs elle-même.

Cette situation est très semblable à celle de l'interprétation des SN : le SN quantifié « tout habitant » doit être lui aussi capable de prendre pour argument n'importe quelle suite de phrases. Mais si nous appliquons au langage une telle « sémantique des continuations », alors ce ne sont pas seulement les SNs qui doivent avoir un type complexe, mais également les autres constituants.

Ainsi avions-nous auparavant la situation suivante :

syntaxe	sémantique	commentaires
S → SN SV	σ(S) = (σ(SV))(σ(SN))	La sémantique du syntagme **verbal** est appliquée à celle du syntagme **nominal**.
SV → V SN	σ(SV) = (σ(V))(σ(SN))	La sémantique du verbe **est appliquée** à celle du sn.

Ajoutons-y :

SN → /*Paul*/, SN → /*Marie*/ etc. SV → /*chante*/, V → /*rencontre*/

avec les sémantiques :

/Paul/ : **p**, /Marie/ : **m**, */chante/* : λx. **chante**(x) et */rencontre/* : λx. λy. **rencontre**(y, x), les types sémantiques étant : σ(S) : **t** ; σ(SN) : **e** ; σ(SV) : <**e, t**> ; σ(V): <**e**, <**e, t**>>. Avec ce que nous avons vu plus haut, nous pouvons attribuer un type aux *continuations* de ces diverses expressions :

- pour une (sous) phrase : <**t, t**>
- pour un syntagme nominal : <**e, t**>
- pour un syntagme verbal : <<**e, t**>, **t**>
- pour un verbe : <<**e**, <**e, t**>>, **t**>

d'où on déduit les types des « sémantiques CPS » des expressions des diverses catégories syntaxiques :

- σ(S)* : <<**t, t**>, **t** >
- σ(SN)* : <<**e, t**>, **t**>
- σ(SV)* : <<<**e, t**>, **t**>, **t**>
- σ(V)* : <<<**e**, <**e, t**>>, **t**>, **t**>

ce n'est bien sûr pas une surprise de retrouver le type sémantique <<**e, t**>, **t**> pour un syntagme nominal : c'est exactement celui que Montague lui avait attribué. On notera d'autre part qu'un type <<α, **t**>, **t**> revient à une double négation (en « interprétant » **t** comme ⊥) et qu'il n'est équivalent à α qu'en vertu de la *loi logique de double négation*, valide seulement en logique classique.

Les sémantiques directes et « continuées » des mots inclus dans le lexique ci-dessus sont données dans le tableau ci-après :

Paul	**p**	λu.u(**p**)
Marie	**m**	λu.u(**m**)
chante	λx. **chante**(x)	λP.P(λx. **chante**(x))

rencontre	$\lambda x.\ \lambda y.$ **rencontre**(y, x)	$\lambda R.R(\lambda x.\ \lambda y.$ **rencontre**$(y, x))$

qu'on peut maintenant compléter avec des noms et des déterminants quantificateurs :

chanson	$\lambda x.$ **chanson**(x)	$\lambda P.P(\lambda x.$ **chanson**$(x))$
vedette	$\lambda x.$ **vedette**(x)	$\lambda P.P(\lambda x.$ **vedette**$(x))$
toute	$\lambda P.\ \lambda Q.\ \forall x\ P(x) \Rightarrow Q(x)$	$\lambda U.U(\lambda P.\lambda Q.\forall x\ P(x) \Rightarrow Q(x))$
une	$\lambda P.\ \lambda Q.\ \exists x\ P(x) \wedge Q(x)$	$\lambda U.U(\lambda P.\lambda Q.\exists x\ P(x) \wedge Q(x))$

Il reste alors à transformer les contreparties sémantiques des règles syntaxiques puisqu'on ne peut plus désormais obtenir par exemple la représentation d'une phrase (S) en appliquant la représentation du SV à la représentation du SN sujet (les représentations en question sont respectivement de types : $\langle\langle \mathbf{t}, \mathbf{t}\rangle, \mathbf{t}\rangle$, $\langle\langle\langle \mathbf{e}, \mathbf{t}\rangle, \mathbf{t}\rangle, \mathbf{t}\rangle$, $\langle\langle \mathbf{e}, \mathbf{t}\rangle, \mathbf{t}\rangle$ et elles ne sont plus composables de manière directe, ou plus exactement, l'application d'une fonction de type $\langle\langle\langle \mathbf{e}, \mathbf{t}\rangle, \mathbf{t}\rangle, \mathbf{t}\rangle$ à une fonction de type $\langle\langle \mathbf{e}, \mathbf{t}\rangle, \mathbf{t}\rangle$ donne un objet de type **t** et non une fonction de type $\langle\langle \mathbf{t}, \mathbf{t}\rangle, \mathbf{t}\rangle$, comme il est demandé). Or, on peut démontrer le résultat suivant (adapté de C. Barker, 2002) :

Lemme : étant donnée une règle $C \rightarrow AB$, avec $c = m(a, b)$ (où nous avons noté pour plus de simplicité a, b et c les représentations sémantiques respectives de A, B et C et où m est l'opération qui relie a, b et c, ici l'application), il existe deux règles : $C \rightarrow AB$, avec $c^* = m_1(a^*,b^*)$ et $C \rightarrow AB$, avec $c^* = m_2(a^*,b^*)$ qui sont telles que pour i=1, 2, $m_i(a^*,b^*)(\lambda x.x) = m(a, b)$.

Il suffit que m_1 et m_2 vérifient la contrainte suivante, appelée *schème de continuation* :

$m_1(a^*,b^*) = \lambda \underline{u}.(a^*(\lambda x.\ (b^*\ (\lambda y\ \underline{u}(m(x, y)))))))$

$m_2(a^*,b^*) = \lambda \underline{u}.\ (b^*(\lambda y.\ (a^*\ (\lambda x\ \underline{u}(m(x, y)))))))$

Démonstration : on doit d'abord démontrer que pour toute fonction g et toute sémantique continuée x^*, $x^*(\lambda u.g(u)) = g(x)$. Noter qu'en ce qui concerne notre lexique ci-dessus, c'est bien le cas, d'une manière triviale. On généralise ensuite par induction sur la structure de l'expression. Ceci étant on voit alors que :

$$[\lambda \underline{u}.(a^*(\lambda x.\ (b^*\ (\lambda y\ \underline{u}(m(x, y)))))))]\ (\lambda x.x) > a^*(\lambda x.\ (b^*\ (\lambda y\ (\lambda x.x)\ (m(x, y))))))$$
$$> a^*(\lambda x.\ (b^*\ (\lambda y\ (m(x, y))))$$
$$> a^*(\lambda x.\ (m(x, b))))$$
$$> m(a, b)$$

idem pour $m_2(a^*,b^*)$.

Ce lemme fait surgir la question suivante : *d'où vient qu'il y ait deux modes de composition possibles associés à une même règle binaire ?* (et, plus généralement n! pour une règle n-aire $B \rightarrow A_1\ A_2\ \ldots\ A_n$). On pourrait penser qu'il s'agit d'une décision arbitraire. Or, il n'en est rien. Des informaticiens théoriciens (Griffin, 1990, Parigot, 1992, de Groote 1994) ont montré deux liens importants :

1- le lien, établi par P. de Groote (1994) entre la programmation par continuation (CPS) et une extension du λ-calcul, dont une variante a été définie par Parigot sous la dénomination de *λμ-calcul*,

2- le lien, établi par M. Parigot (1990) entre ce λμ-calcul et la logique classique, sous la forme d'une extension de l'isomorphisme de Curry-Howard.

Il en résulte d'une part que toutes les transformations CPS que nous effectuons (cf. celles que nous avons effectuées ci-dessus) peuvent s'exprimer en λμ-calcul et que d'autre part, les λμ-termes que l'on obtient sont les décalques de preuves qui sont faites *non plus en logique intuitionniste mais en logique classique* (cf. annexe). Cela peut paraître contradictoire avec une affirmation souvent entendue selon laquelle la logique classique ne serait pas « constructive » (c'est-à-dire n'aurait pas de contenu algorithmique). De fait elle est constructive et possède un contenu algorithmique mais, contrairement à la logique intuitionniste, elle n'est pas *confluente* : cela signifie que la normalisation des $\lambda\mu$-termes ne possède pas la propriété dite « de Church-Rosser », c'est-à-dire la propriété selon laquelle tous les chemins possibles suivis pour arriver à la normalisation d'un terme conduisent au même résultat. *La logique classique offre ainsi des possibilités de bifurcation*, et c'est justement ce qui se traduit dans la possibilité d'associer deux modes de composition m_1 et m_2 à une règle binaire. Evidemment, dans le cas général, ces deux modes ne vont pas donner le même résultat, et c'est ce dont nous avons besoin justement dans le cas de la sémantique des langues naturelles.

3.3 Richesse de l'approche logico-mathématique

Le résultat ci-dessus est étonnant. Il montre toute la fécondité d'une approche mathématisée du langage qui, seule, peut aller assez loin dans les propriétés formelles de certains phénomènes pour qu'on puisse dégager des isomorphies inattendues entre par exemple, d'un côté des propriétés structurelles, dynamiques d'un système logique (en l'occurrence le plus usité : la logique classique) et de l'autre des propriétés non moins structurelles, mais d'objets empiriques comme les phrases d'une langue. On remarquera également au passage qu'ici les mathématiques ne sont pas un simple moyen de notation, un langage servant à exprimer de manière « plus précise », « plus rigoureuse » telle ou telle propriété que l'on connaissait déjà : leur évolution propre, autonome, apporte des solutions inédites à des problèmes soulevés par des phénomènes que l'on constatait mais qu'on ne pouvait résoudre que par des moyens *ad hoc* (*quantifier raising* etc.).

3.3.1 Sur la compositionnalité

Les conséquences en retour sur la théorie linguistique sont nombreuses. C. Barker (2001) en note une qui concerne la théorie de la compositionnalité. Depuis Frege, on sait que la sémantique s'est basée sur le principe selon lequel « la signification d'une expression linguistique complexe est une fonction des significations de ses sous-expressions immédiates », voire sous une forme plus élaborée : « une fonction des significations de ses sous-expressions immédiates *et de la manière dont elles sont combinées* », c'est bien sûr la voie qui a été prise par toute la tradition montagovienne. Or, cela ne suffit pas puisqu'une phrase obtenue par une seule suite de règles syntaxiques bien définie peut avoir deux lectures ! Il faut donc encore modifier le principe : « la signification d'une expression linguistique complexe est une fonction des significations de ses sous-expressions immédiates, de la manière dont elles sont combinées, *et de leur mode sémantique de composition* ».

3.3.2 La place des stratégies d'évaluation

De fait, un tel mode de composition est plutôt un mode *d'évaluation*. Les informaticiens relient en effet la normalisation d'une preuve à *l'évaluation* d'une forme symbolique (l'exécution d'un programme). Les résultats concernant les transformations CPS sont classiquement interprétés dans la littérature informatique (cf. Griffin, 1990 : « A formulae-as-types notion of control ») comme portant sur le *contrôle* des calculs ou sur les divers modes d'évaluation possibles (cf. la distinction en programmation fonctionnelle entre « call by name » et « call by value », sur laquelle nous revenons en annexe). L'ambiguïté d'une phrase comme (26) résulte *donc d'une pluralité de stratégies d'évaluation*. Tout se passe comme si les expressions linguistiques étaient des formes symboliques à évaluer au sein d'un système non déterministe, où des choix de stratégie de contrôle restent à faire. Cela ouvre notamment la voie à *une théorie de l'intonation prosodique en tant que manière de forcer les choix à effectuer*.

Avec les exemples donnés jusqu'ici, il semble qu'une position totalement empiriste à l'égard du langage soit mise en difficulté : elle ne parviendrait à dégager ni les régularités des phénomènes de polarité ni les isomorphies que nous venons d'évoquer avec le fonctionnement de systèmes artificiels. Le partisan d'une telle position peut sans doute tenter d'invalider la démarche au prétexte qu'il n'y a rien là de « prouvé » empiriquement… et que pour qu'il y ait preuve, il faudrait accumuler un nombre « suffisant » de faits empiriques, mais outre qu'une accumulation de faits empiriques n'a jamais fait une preuve à proprement parler, c'est ignorer que dans les autres sciences également, on échoue toujours à apporter une « vérification expérimentale » complète et on se contente de modèles qui ne sont que partiellement vérifiés, ce qui n'empêche pas de croire en leur vérité, même partielle. Le fameux « hypotheses non fingo » de Newton ne doit pas être mal interprété : le grand physicien anglais n'a jamais voulu dire qu'il rejetait la notion d'hypothèse (cf. Koyré, 1965), mais seulement que les hypothèses ne devaient jamais être de simples manières de « sauver les phénomènes », autrement dit de simples « vues de l'esprit ». Pour lui, les hypothèses avaient prétention à saisir quelque chose du réel.

4 Retour à une question : le point de vue computo-représentationaliste est-il nécessaire ?

4.1 Une hypothèse d'espace

Revenons à la question posée à la section … qui avait été motivée par une observation de S. Auroux au début de son ouvrage d'épistémologie des sciences du langage, où il s'interrogeait sur la capacité d'un modèle computationnel à expliquer le comportement linguistique humain, autrement dit sur le fait de savoir si, comme il le disait : « [la] compétence linguistique [de chacun d'entre nous] s'explique parce que nous avons des algorithmes implémentés dans notre tête ? ». On aura remarqué que dans les sections précédentes, nous n'avons jamais eu recours explicitement à cette hypothèse. Y aurions-nous eu alors recours *implicitement* ?

Je ne le pense pas. Encore une fois pas plus que le physicien mathématicien n'a recours à l'hypothèse d'un Grand Calculateur lorsqu'il formule des lois de la physique. Que nous utilisions des items de polarité négative spontanément dans le contexte de quantificateurs monotones décroissants ou que les phrases avec expressions quantifiées soient interprétables de diverses manières ne résulte pas nécessairement d'un algorithme programmé pour cela dans nos têtes. *Je ferai plutôt l'hypothèse qu'il s'agit des propriétés générales d'un espace de l'interprétation que tous les locuteur-interprètes partagent*, à la façon dont ils partagent un espace géométrique pour leurs déplacements. Cet espace est extérieur aux individus : ceux-ci viennent « l'habiter » en vertu même de leurs *aptitudes biologiques* acquises concernant le langage. Cet espace externe est évidemment fait de phrases déjà existantes, de récits mille fois entendus et transmis par l'écriture puis l'imprimerie, de dictons rabâchés, d'actes de langage qui se répondent en vertu de multitudes de dialogues qui ont déjà eu lieu. L'apport du dialogue est d'ailleurs probablement fondamental dans l'instauration d'un système de polarités, en raison de l'asymétrie constitutive entre les rôles de proposant et de répondant (ou de questionneur et questionné).

4.2 Des structures « transindividuelles »

Je reprendrai donc ici la proposition de S. Auroux, qui est de considérer que « les structures cognitives [sont] externes à l'individu [,] ont connu leur développement grâce à la technologie intellectuelle de l'écriture [et] dépendent également d'instruments externes (livres bibliothèques, instruments de calcul et d'observation, etc.) ainsi que des structures sociales de production et de cumulation des connaissances », tout en faisant remarquer que l'on trouve un peu la même idée chez le Lacan des *Ecrits* lorsqu'il écrivait, à propos de « la vérité de l'Inconscient », dans « Fonction et champ de la parole et du langage » (1953) que « le plus souvent déjà *elle est écrite ailleurs*[13]. A savoir : dans les monuments […], dans les documents d'archives aussi […], dans l'évolution sémantique […], dans les traditions aussi, voire dans les légendes qui sous une forme héroïsée véhiculent mon histoire ; dans les traces, enfin, qu'en conservent inévitablement les distorsions […] etc. » (p. 136-7). La différence tient évidemment à ce que le Lacan des années cinquante se concentre sur l'Inconscient dont il croit possible de mettre à jour la vérité à la manière dont on rétablit un chapitre censuré, alors qu'il n'est pas question de cela ici, mais seulement de l'idée d'envisager les processus à la source de la signification en tant que *transindividuels.*

Comme tout espace, cet espace de l'interprétation est soumis à des contraintes et à des lois concernant les mouvements des objets qui s'y meuvent. En même temps qu'il est produit et reproduit par les locuteurs au sein de structures sociales et d'institutions, les systèmes cérébraux de ces mêmes locuteurs *se moulent dans ses formes*, ce qui peut donner l'illusion qu'il a son siège dans chaque cerveau individuel. En ce sens, il n'y a sans doute pas *d'algorithme implémenté* qui engendrerait ou reconnaîtrait des phrases correctes du point de vue des polarités. En entendant une phrase ou en en produisant une, le système de compréhension du locuteur *s'attend à* trouver telle ou telle forme en vertu de la manière même dont il a été modelé dès l'origine. Nous savons bien que lorsque nous essayons de juger de la grammaticalité d'une phrase, c'est à une sorte de « sentiment de la langue » (selon une expression de l'écrivain Jean Paulhan) que nous faisons appel, qui tient plus à ce que les musiciens appellent « avoir de l'oreille » qu'à la mise en œuvre d'un algorithme infaillible. En faisant appel à ce sentiment, nous explorons en nous des souvenirs enfouis parfois très loin au sujet d'une structure de langue. Ce sont des souvenirs qui ont été déposés dans notre mémoire par les multiples institutions, lectures et conversations auxquelles nous avons participé dans le passé. Si « norme » il y a, c'est moins d'une norme édictée que d'une *norme de conformité à un espace préétabli* qu'il s'agit, semblable à une *notion de moindre effort* pour une dynamique se déroulant dans un espace physique.

Si la tradition grammairienne (englobant la grammaire générative) nous a habitué depuis des décennies à voir les énoncés sous l'angle d'une seule dichotomie (*acceptable* / *inacceptable*), cela ne doit pas nous faire oublier que le domaine linguistique est plutôt celui de la *nuance* et de la *gradation* : nombreux sont les items qui servent avant tout à établir des degrés sur une échelle. On notera par exemple la gradation suivante :

> *Pierre n'a pas pris le moindre poisson ≅ Pierre n'a pas pris le plus petit poisson < Pierre n'a pas pris un seul poisson < Pierre n'a pas pris de poisson < Pierre a tout juste pris un poisson < Pierre a pris un seul poisson < Pierre a pris un poisson < Pierre a pris quelques poissons < … etc… < pas un poisson n'a échappé à Pierre*

où les échelons extrêmes sont des échelons marqués du point de vue des polarités. Quand il étudie ce qu'il appelle « l'usage informationnel du langage », Dessalles (2000) insiste fort justement à mes yeux sur l'alternance au sein d'une conversation entre *saillance* et *banalisation* : un premier locuteur attire l'attention d'un autre sur une situation saillante, le second répond par une tentative de banalisation, ce à quoi le premier réplique en tentant de rehausser la saillance et ainsi de suite jusqu'à l'arrêt de la conversation. Les procédures de

[13] c'est moi qui souligne

captation de l'attention et de banalisation passent par des marquages particuliers d'énoncés au moyen de polarités ou de tout autre moyen syntaxique semblable (structures clivées par exemple).

D'une manière générale, le linguistique apparaît comme le domaine du contraste et de la saillance, le domaine où se pose la question : *comment détacher en tant que* **pertinent** *un élément particulier d'un flux informationnel ininterrompu* ? En cela, il se distingue du biologique qui n'est concerné *que* par le flux informationnel.

5 Le biologique et le linguistique

5.1 La conversion comme mécanisme général

Tous les livres élémentaires de biologie le disent (cf. le récent « Comment évoluent nos gènes » de Bernard Dujon, 2005), « la propriété fondamentale de la vie, celle qui différencie les organismes vivants du monde inanimé, c'est *la reproduction* ». Au centre des mécanismes moléculaires de l'hérédité figurent les notions de *transcription* et de *réplication*. La vie reproduit du similaire, mais cette notion de similarité elle-même semble ne pas pouvoir être définie en dehors d'une référence à la notion d'information. Pour les biologistes contemporains, la réplication, comme d'ailleurs la mutation, sont subordonnées à *l'information*, laquelle est le fondement de la notion *d'individuation*, qui joue un rôle si important dans le monde du vivant. Dans le monde physique, en effet, deux morceaux de matière peuvent être rigoureusement identiques, avec leurs atomes et molécules interchangeables alors qu'on sait qu'il n'en est pas de même pour le vivant, où chaque organisme est *unique*. Le cycle biochimique n'est ainsi pas suffisant pour l'explication de la vie : se superpose à lui un « cycle informationnel ». L'existence de ce dernier permettrait de comprendre la raison d'être de la duplication elle-même, et donc de la vie : l'information est en effet un capital bien fragile (non stockable dans l'espace de manière tangible) qui, pour perdurer, n'a guère d'autre solution que la reduplication permanente, d'où l'ADN et sa structure en hélice comme solution unique.

Des logiciens se lancent aujourd'hui[14] dans la collaboration avec des biologistes pour modéliser les mécanismes cellulaires, en trouvant en eux des similarités avec le fonctionnement de machines de Turing (cf. la notion de transcription), comme avec la notion d'interaction intervenue récemment en logique et mise à profit dans l'étude des interactions entre processus (cf. parallélisme). Parmi eux figure Antoine Danchin, qui intitule une de ses conférences : « The cell as a living computer » (consultable à l'adresse : http://www.pasteur.fr/recherche/unites/REG/lectures/Stockholm1_04.pdf).
Dans cette conférence, le biologiste définit différentes disciplines par leurs objets respectifs, et en particulier :

« - Physique : matière, énergie, temps…
- Biologie : Physique + information, codage, contrôle… »

A cela nous pourrions ajouter la linguistique en tant que science qui se définirait comme : **Biologie + pertinence, argumentation, dialogue**.

Les analogies avec la biologie existent en effet. Dans les deux cas, on trouve une idée de *dynamique de la conversion*, du genre de celle qui fait passer du texte du programme à son exécution, du texte de l'ADN à la production de protéines et de la recette de la crêpe à la crêpe…. Il y a *conversion* en effet chaque fois qu'une séquence (éventuellement infinie) d'objets élémentaires (qu'il s'agisse de gènes, de mots ou de symboles divers) donne lieu à des transformations en des objets finis d'un autre ordre (des protéines pour les séquences

[14] cf. une journée récente à la Sorbonne (22 avril 2005) sur « Le Logique et le Biologique », organisée par Jean-Baptiste Joinet.

d'ADN, des « significations localisées » pour des suites de mots, des calculs concrets et effectifs pour des schémas d'induction généraux). Le λμ-calcul, présenté en annexe, est un exemple de mécanisme de conversion.

En ce qui concerne notre compréhension du langage, comme nous l'avons déjà évoqué plus haut, mais sans nous étendre, des linguistes et psycho-linguistes (Thomas Bever et David Townsend, 2002) ont mis en évidence l'existence de deux phases dans la compréhension d'un discours ou d'un texte (je distingue les deux, le discours étant plutôt oral et le texte écrit) : une phase d'analyse initiale d'un sens vraisemblable et une autre de reconstitution de la structure syntaxique dérivationnelle complète cohérente avec la forme et le sens. Ils relient cette théorie de la compréhension à une théorie semblable en ce qui concerne la perception : lorsque nous cherchons nos lunettes dans l'appartement, disent-ils, et que finalement nous les trouvons, nous les avons en réalité « vues » deux fois, la première est passagère, « mais elle laisse le clair souvenir d'un évènement qui paraît pré-perceptif », la seconde est bien sûr au moment où nous les trouvons, et elle nous paraît être la première parce que la vraie première a été immédiatement enfouie dans notre inconscient. Il en irait semblablement avec le langage : le tout-venant du discours et du texte est écouté ou lu sans nécessairement une attention particulière, comme un balayage furtif exercé par notre ouïe ou notre regard de lecteur, et ce n'est qu'en certains moments bien localisés (lorsque nous ne sommes pas sûrs de bien comprendre quelque chose ou bien lorsque nous cherchons une information précise et que nous voulons confronter ce qu'on lit ou ce qu'on entend avec la requête que nous avons en tête) que nous procédons à l'analyse détaillée (« dérivationnelle » pour reprendre le qualificatif très chomskyen utilisé par les auteurs) d'un fragment du discours proféré ou du texte lu. Autrement dit, nous aurions une différenciation entre deux niveaux tout aussi différents que le sont le texte de l'ADN et une suite d'actions biologiques déterminées, entre un niveau des instructions générales évoquées par les mots et les phrases mais pas nécessairement évaluées et celui de la mise en activation de ces instructions afin de produire un « texte concret », c'est-à-dire une évaluation « mentale » authentique. Celle-ci peut se traduire alors par la reconnaissance d'un accord avec une requête ou par une séquence d'actes : cas d'un ordre par exemple, où là, l'opération est très similaire avec ce qui se passe dans la cellule, ou bien encore par la production par soi-même d'autres phrases et d'autres mots, dans une situation de dialogue, et dans le but de poursuivre le flux langagier, selon un processus qui, après tout, n'est pas sans rappeler… la réplication de l'ADN.

5.2 Programmes et pertinence

Ce que nous avons vu concernant l'interprétation de phrases avec expressions quantifiées montre à mes yeux particulièrement bien le rapprochement qu'il est possible d'établir entre programme, interprétation sémantique et mécanismes cellulaires. La notion de programme génétique n'est plus une simple métaphore à partir du moment où on peut décrire les mécanismes cellulaires comme des « ordinateurs vivants ». De même, les mécanismes de l'interprétation sémantique, nous l'avons vu, sont similaires à ceux que l'on trouve en matière d'évaluation de programmes, ce qui rend l'interprétation de la phrase analogue à une évaluation de forme symbolique. Cependant, il ne s'agit jamais de « tout » évaluer, mais seulement, dans un flux permanent d'informations, les fragments *pertinents*.

Si la vie trouve son origine dans la nécessité de maintenir un cycle informationnel, on peut suggérer que le langage (et tout ce qu'il draine avec lui en matière de symboles et donc de culture et d'organisation de la société) trouve la sienne dans la nécessité de maintenir *un flux de pertinence dans l'information*, tâche qui s'effectue principalement par le dialogue et plus généralement par ce que Wittgenstein appelait des *jeux*. La logique et les mathématiques s'intéressent alors aux structures formelles que cela met en œuvre. De même que la vie trouve une solution dans une structure formelle : celle de l'hélice de Crick et Watson, l'hominisation en trouverait une, voire plusieurs, dans la structure formelle liée à l'activité langagière (dialogique, conversationnelle etc.). Dans les deux cas, ce que nous entendons par « structure formelle » réside dans l'identification de quelques invariants qui assurent la reproduction[15], la reduplication et la diffusion.

Dans cette problématique, la question de savoir si l'on traite de « faits » ou de « normes » apparaît secondaire, ou du moins elle tient la même place que dans les autres sciences humaines, où « fait » et « norme » apparaissent souvent interchangeables. Il ne s'agit évidemment pas de norme « prescriptive » (au sens de l'Académie, voire du *Grévisse*, dictant les règles du *bien parler*) mais d'une norme créée par l'usage. En biologie, après une mutation génétique, la reproduction continue, emmenant avec elle les résultats de mutations qui se sont avérées viables. Dirons-nous en ce cas que l'espèce mutante *suit une nouvelle norme* ? On le pourrait sans doute, et en ce cas, il n'y aurait pas de différence fondamentale avec le cas du langage lorsqu'une nouvelle « mode » langagière surgit. Quant à savoir ce qu'il en est de la situation de chaque individu par rapport à cette « norme », on peut simplement concevoir qu'il en est de chacun comme d'une trajectoire dans un espace. Les trajectoires sont individuelles et l'espace peut être vu aussi comme leur fusion. Sont-elles alors dessinées avant ou après l'espace qui se révèle être leur support ? C'est une question de point de vue, question irrésoluble puisque la possibilité même du point de vue est incluse dans la solution…

6. Conclusion

Au début de ce texte, nous nous interrogions sur la viabilité d'une approche mathématique en sciences du langage. L'objection essentielle reposait dans l'accusation formulée par certains (notamment des socio-linguistes) selon laquelle on en arrivait à faire de la langue « une invention des linguistes », voire pire encore (une telle pensée étant tellement hors de l'horizon des auteurs en question qu'ils ne l'ont encore jamais formulé publiquement ainsi) : *une invention des mathématiciens*. De fait, il ne semble pas que la distorsion entre le réel et la théorie soit plus forte ici qu'elle ne l'est dans les autres sciences, qui depuis au moins Galilée, ont été forcées d'admettre l'idéalisation et la perte du contact charnel avec le monde. Schrödinger a théorisé ce mouvement de la science sous le concept (repris et développé par M. Bitbol dans Bitbol, 1990) « d'élision du sujet ». Husserl a identifié ce moment comme celui d'une crise. Rien de nouveau donc. Allons-nous pour autant contester la démarche scientifique dans ce qu'elle nous apporte de plus profond dans notre connaissance du monde ?

J'ai donné un certain nombre d'exemples (les *pratyaharas* de Panini, les items de polarité négative, la quantification) de l'intervention des mathématiques en tant que produisant des effets explicatifs. Dans le cas de la grammaire de Panini, il ne s'agissait pas à vrai dire « d'expliquer » un phénomène linguistique « objectif », mais plutôt d'expliquer *a posteriori* un geste de classification linguistique. C'était déjà indiquer que la mathématisation dans les sciences (notamment humaines) est toujours ambiguë car elle englobe en une seule symbolisation l'objet à appréhender et le geste qui l'appréhende. Cela est vrai surtout pour les sciences du langage où l'on peut difficilement séparer le signe de son interprétation (et à partir de là de son interprétation à tous les niveaux : épilinguistique comme métalinguistique), à moins de réduire le signe à la pure marque matérielle[16], or une telle réduction est impossible dans les systèmes naturels (elle l'est peut-être dans les systèmes artificiels, et encore ce n'est pas évident[17]), sauf à produire un pur fantasme technologique (celui qu'on aperçoit dans les soi-disant « industries de la langue »). Plutôt donc que de « masquer » cette ambiguïté derrière

[15] On notera que c'est le même mot, « reproduction » qui est utilisé aussi bien par les biologistes en ce qui concerne la vie, que par les sociologues (cf. Bourdieu) en ce qui concerne la société.

[16] Conception à distinguer radicalement de celle de la « matérialité du signifiant » : toujours se rappeler que, en conformité avec la doctrine de Saussure, le signe a bien deux faces, dont une est le signifiant et l'autre le signifié.

[17] Les symboles du mathématicien sont loin d'être de pures marques matérielles…

une affirmation positiviste de réduction du signe à sa marque, on ferait mieux de la montrer pour ce qu'elle est, et « d'avouer » en quelque sorte qu'on formalise *à la fois la matière et le geste* du linguiste.

Dans les autres exemples, nous avons vu les mathématiques œuvrer selon leur régime normal. On identifie certaines formes mathématiques sous des apparences linguistiques : des fonctions, des relations d'ordre, des filtres, des idéaux… et on les met en relation avec d'autres formes qui elles-mêmes revêtent d'autres apparences linguistiques : des items de polarité par exemple. La théorie mathématique est utilisée pour comprendre ces relations systématiques. Ou bien, dans le cas de l'évaluation d'expressions quantifiées, on rapporte le phénomène à un modèle général de l'évaluation. Les propriétés globales d'un tel modèle sont utilisées à leur tour pour rendre compte de comportements dans le domaine d'application qui, sans cette démarche, devraient faire l'objet de stipulations *ad hoc* pour qu'il en soit rendu compte (à la manière dont s'y prend la grammaire générative classique avec ses transformations de *quantifier-raising*).

La question difficile est de savoir si, dans ce dernier cas notamment, le modèle est une simple *métaphore* ou si, au contraire, thèse beaucoup plus forte, il est « l'authentique » cadre général à l'intérieur duquel se situe le cas particulier du langage standard. Autrement dit, s'agit-il seulement, comme le disait Osiander à propos de Copernic, de « sauver les phénomènes » ou bien s'agit-il de plus, comme le prétendait Galilée (qui souhaitait qu'on le considérât davantage comme « philosophe » que comme « astronome ») ?

Si on adopte la seconde perspective, autrement dit si l'on admet que le modèle n'est pas pure métaphore, il faut reconnaître qu'une irrésistible pente nous entraîne vers une théorie computo-représentationnelle de l'esprit, dont le langage n'est qu'une région, ou, en termes plus brefs et abrupts, vers la théorie selon laquelle nous parlons ainsi parce que nous avons une grammaire dans la tête et que celle-ci fonctionne comme un grand algorithme… Or, ce n'est pas la voie que j'ai choisie. Il me semble au contraire que pour sortir de ses apories présentes, la science du langage doit se porter vers la conception, soutenue par S. Auroux, *de l'extériorité des structures cognitives*. A tout le moins remarquera-t-on que « l'hypothèse du sujet » n'est jamais intervenue dans les explications que j'ai données des exemples. C'est à une « hypothèse d'espace » que j'ai fait appel (ou *de temporalité*… ceci n'est pas tranché). La proposition d'Auroux se résume dans le concept d'*hyperlangue*, (S. Auroux, op. cit. p.115) :

> *[…] l'espace-temps, par rapport à l'intercommunication humaine, n'est pas vide, il dispose d'une certaine structure que lui confèrent les objets et les sujets qui l'occupent. Appelons* hyperlangue*, cet espace-temps ainsi structuré. Introduire un nouvel objet (par ex. un sujet doté de capacités linguistiques déviantes, un dictionnaire, ou encore le moyen de communiquer à distance) change la structure de l'hyperlangue. Les* évènements *dans l'hyperlangue (ce que nous appelons les « discours ») en changent également (plus ou moins) la structure […]*

On peut en rapprocher la conception qui était celle de l'analyse du discours en France dans les années quatre-vingt, telle que défendue à l'époque par Michel Pêcheux[18], pour laquelle tout discours s'inscrit dans un *interdiscours* qui en détermine les lois de fonctionnement, les préconstruits, les formulations possibles et impossibles etc. Pêcheux parlait de *matérialité discursive* comme « [un] ressassement de paroles entendues, rapportées ou transcrites, [un] fourmillement d'écrits citant des paroles, et d'autres écrits » (cité par F. Mazière, p. 53).

En résumé, il peut y avoir mathématisation hors d'une conception de la langue qui serait régie par le seul concept de *grammaire*. Il s'agit alors de l'étude des propriétés logico-

[18] cette position est redevenue accessible aujourd'hui grâce à la publication de l'excellent petit ouvrage « l'analyse du discours », par Francine Mazière, dans la collection « Que sais-je ? » (n°3735)

mathématiques d'un type d'espace, indépendamment de toute hypothèse sur l'instance qui « les met en œuvre » (idée qui d'ailleurs apparaît incongrue dès que formulée). Il semble peu vraisemblable d'atteindre un jour un état de connaissance du langage caractérisé par une grammaire telle qu'à partir d'elle et d'elle seule un « sujet » (de préférence artificiel !) puisse produire et interpréter « toutes les phrases ». Il semble encore plus invraisemblable qu'un tel système puisse être « décidable »… voir « décidable en temps polynomial » ! Cette question ne fait pas seulement intervenir un point de vue de *limitation technique* (est-il possible d'envisager un système formel F engendrant toutes, et rien que, les phrases d'une langue L), mais également un point de vue théorique sur la langue. Une telle conception supposerait en effet que le point de vue *inductiviste* (la langue est susceptible d'une définition au moyen d'un schéma de récursion) et *compositionnaliste* (la signification d'une expression comme fonction des significations de ses parties) soit adopté une fois pour toutes. Or, rien n'est sûr en ce domaine. Nous avons déjà dû modifier le principe de compositionnalité plus haut pour tenir compte des ambiguïtés de portée des quantificateurs. Hintikka (1994) a, de son côté, donné des arguments contre la compositionnalité[19].

Finalement, la perspective développée ici concernant la mathématisation dans les sciences du langage repose sur des *montages* (des déductions, des réductions, des « preuves ») chargés de valider des hypothèses et dont la réussite n'est jamais établie à l'avance (comme ce serait le cas pour une approche transcendantale du langage, guidée par le concept de « grammaire pure » par exemple), en ce sens elle peut être qualifiée « d'expérimentale ». Ajoutons que ces montages mêlent des observables et des inobservables, comme dans un montage physique (voir l'exemple des items de polarité) et qu'ils sont souvent liés indissolublement aux gestes mêmes du linguiste visant à poser un terme ou une relation entre termes. De ce point de vue, ils n'ont certainement pas toute « l'objectivité » requise par un modèle physique, mais apparaissent comme très consubstantiels à l'activité mathématique qui elle-même est toujours inscrite dans ce léger décalage entre entité posée et geste ou acte consistant à la poser (cf. Cavaillès, p. 32, à propos des mathématiques : « tout sens posant est en même temps sens posé d'un autre acte »), même si, dans les sciences du langage, le mouvement d'abstraction (« le principe que toute idée a une réalité formelle ») n'est pas allé aussi loin que dans les mathématiques en général (même si donc, pour reprendre les termes de Cavaillès, « l'idée de l'idée » n'a pas encore « manifesté [toute] sa puissance génératrice »).
D'un autre point de vue, il est utile de comparer les modèles en linguistique à ce qu'ils sont ailleurs. Cet ailleurs est l'union de deux domaines : la physique et la biologie. Les modèles en linguistique se distinguent des modèles en physique pour la raison que nous avons donnée plus haut : en physique, l'impossibilité d'appréhender le réel par nos sens et notre langage ordinaire est compensée par l'apport des mathématiques en tant que *constituantes* de ce réel, alors que dans le langage, les mathématiques ont toujours tendance à se surajouter à un discours déjà théorique qui *va avec* l'activité langagière. Si par ailleurs, les modèles en biologie se distinguent des modèles en physique par l'intégration nécessaire du paramètre temporel (les modèles en physique se ramènent à une dynamique dans un espace de phases, alors qu'en biologie, ils contiennent à la fois une dynamique dans l'espace des phases et une dynamique temporelle qui modifie constamment cet espace (cf. Danos, Journée « le logique et le biologique » 22 avril 2005, Sorbonne)), en sciences du langage, le paramètre *supplémentaire* par rapport au temps « biologique » est *le temps de l'histoire*, autrement dit une temporalité scandée par les *alternances* : de tour de parole, entre question et réponse, entre ordre et exécution, entre discours et action, toutes alternances qui s'inscrivent dans la langue de manière durable sous forme de *polarités*. C'est dire qu'aucune mathématisation n'y

[19] ces arguments ne sont pas si éloignés de ceux que donne Auroux (op. cit. pp. 70-71) à propos de phrases insérées dans des contextes de telle manière que leur valeur apparaisse immédiatement déterminée à partir du contexte, autrement dit à partir d'une forme globale plutôt qu'à partir des éléments.

est achevée, ou alors il faut intégrer au modèle suffisamment de paramètres pour qu'il puisse s'adapter aux nombreux changements auxquels la langue aura toujours à faire face. La notion de paramètre est ici plus vaste que dans la théorie dite «Principes et Paramètres » de N. Chomsky : un paramètre ne saurait être simplement un « switch » dans un système par ailleurs immobile et achevé.

Bibliographie

Ajdukiewicz, K., [1935], 1967 : 'Die syntaktische Konnexität', *Studia Philosophica*, vol. 1, 1-27; English translation: 'Syntactic Connexion', in S. McCall (ed.), *Polish Logic 1920-1939*, Oxford, Oxford University Press, pp 207-231..

Auroux S., 1998 : *Le langage, la raison et les normes*, Paris, PUF.

Bar-Hillel Y., [1953], 1968, « Une notation quasi-arithmétique destinée aux descriptions syntaxiques », *Langages* n° 9, 9-22.

Baker, M., 2001: *The Atoms of Language, The Mind's Hidden Rules of Grammar*, Basic Books

Barker, C., 2002: 'Continuations and the nature of quantification', *Natural Language Semantics*, 10 (3),

Barwise, J. et Cooper, R., 1981: 'Generalized Quantifiers and Natural Language', *Linguistics and Philosophy* 4,

Bever, T. & Townsend, T., 2002 : 'Quelques phrases sur notre conscience perceptive des phrases', in E. Dupoux, *Les langages du cerveau*, Paris, Odile Jacob, pp. 147-160

van Benthem, J., 1986: *Essays in Logical Semantics*, D. Reidel Pub. Co, Dordrecht,

Bernardi, R, 2002: *Reasoning with Polarity in Categorial Type Logic*, PhD thesis, Utrecht University

Beyssade C. & Dobrovie-Sorin, C., 2005 : *Définir les indéfinis*, CNRS Editions, Paris,

Bitbol, M, 1990 : *L'élision, essai sur la philosophie d'Erwin Shrödinger*, in Schrödinger, *L'esprit et la matière*, ed. du Seuil, Paris

Bitbol, M, 2000 : *Physique et Philosophie de l'esprit*, Nouvelle Bibliothèque Scientifique, Flammarion, Paris

Carey, S.,1998: 'Knowledge of number: Its evolution and ontogenesis'. *Science, 242,* 641-642.

Carey, S.,2002 : 'Evidence for numerical abilities in young infants: A fatal flaw?' *Developmental Science, 5(2),* 202-205.

Carey, S., 2004 : 'Bootstrapping and the origins of concepts'. *Daedalus*, 59-68.

Cavaillès, J., 1960 : *Sur la logique et la théorie de la science*, Presses Universitaires de France, Paris

Chomsky N., 1956: "Three models for the description of language", *Transactions on Information Theory* IT-2, Institute of Radio Engineers, 113-124 (trad. fr. *Langages* n° 9, 1968, 51-76).

Chomsky, N., 1996 : *The Minimalist Program*, MIT Press, Cambridge Mass

Cohen-Tannoudji, G. & Spiro, M., 1990: *La matière-espace-temps, la logique des particules élémentaires*, coll. essais, Folio n°138

Corblin, F., 2002 : *Représentation du discours et sémantique formelle*, coll. linguistique nouvelle, PUF, Paris

Culioli A., 1968 : « La formalisation en linguistique », *Cahiers pour l'Analyse* 9, 106-11.

Culioli A.,1990 : *Pour une linguistique de l'énonciation*. Tome I : *Opérations et représentations*, Paris, Ophrys.

Curry, H., 1961 : 'Some logical aspects of grammatical structure", in Jakobson, R. (ed.) *Studies of language and its Mathematical Aspects*, Providence, Proc. of the 12th Symp. Appl. Math., pp. 56-68.

Desclés, J-P., 1990 : *Langages applicatifs, langues naturelles et cognition*, Hermès, Paris.

Dessalles, J. L., 2001 : *Aux origines du langage, une histoire naturelle de la parole*, ed. Hermès,

Dowty, D., 1988 'Type raising, functional composition and non-constituent conjunction' in Oehrle, Bach and Wheeler (eds) *Categorial Grammars and Natural Language Structures*, Kluwer

Dujon, B. 2005 : *Comment évoluent nos gènes?* Le Pommier et La cité des Sciences et des techniques, Paris

Fauconnier, G, 1975 : 'Polarity and the scale principle', *Papers from the 11th Regional Meeting of the Chicago Linguistic Society*, Univ. of Chicago.

Gardies J-L., 1975, *Esquisse d'une grammaire pure*, Paris, Vrin.

Gazdar, Klein, Pullum, Sag, 1985: *Generalized Phrase Structure Grammar*, Blackwell, Oxford

Griffin, J. 1990: 'A formulae-as-types notion of control'. In *conference record of the seventeenth annual ACM symposium on Principles of Programming Languages*, pp 47-58

de Groote, P. : 1994, 'On the Relation between the λμ-Calculus and the Syntactic Theory of Sequential Control', in F. Pfenning (ed) *Proceedings of the 5th International Conference on Logic Programmaing and Automated Reasoning (LPAR'94),* LNCS, 822, Springer-Verlag, pp 31-43.

de Groote, P.2001a: 'Type raising, continuations and classical logic', *Proceedings of the 13rd Amsterdam Colloquium*, ILLC, Amsterdam

de Groote, P. 2001b : 'Towards Abstract Categorial Grammars'. In *proceedings of* ACL 2001, Toulouse, France.

de Groote, P. & Pogodalla, S., 2004 : 'On the expressive power of Abstract Categorial Grammars: Representing context-free formalisms'. In *Journal of Logic and Information*, volume 13, number 4.

Habermas, J. 1987 : *Logique des sciences sociales et autres essais*, coll. Quadrige, PUF, Paris

Heim, I. & Kratzer, A., 1998 : *Semantics in Generative Grammar*, Blackwell

Hintikka, J., 1994: *Fondements d'une théorie du langage*, PUF, Paris

Jacobson, P., 1996 : *The Syntax-Semantics Interface in Categorial Grammar*, in The Handbook of Contemporary Semantic Theory (S. Lappin, ed), Blackwell,

Kanazawa, M., 2006: "Abstract families of abstract categorial languages" in *proceedings of* WoLLIC 2006, CSLI, Stanford University.

Keenan E. & Westerstahl, D. 1997 : 'Generalized quantifiers', in *the Handbook of Logic and Language* (van Benthem et ter Meulen eds), pp 837-894, North-Holland

Klima, A.1964 : 'Negation in English', *The Structure of Language*, J. A. Fodor and J. J. Katz (eds) Prentice-Hall, Englewood Cliffs, NJ

Koyré, A., 1965 : *Études newtoniennes*. Paris : Gallimard, 1991. (Bibliothèque des idées)

Lacan, J. [1953], 1964 : 'Fonction et champ de la parole et du langage', *Les Ecrits*, ed. du Seuil, Paris

Ladusaw, W., 1979: *Polarity sensitivity as inherent scope relations*, PhD Dissertation, University of Texas at Austin

Lambek J., 1958, "The mathematics of sentence structure", *American Mathematical Monthly* 65, 154-165.

Lambek J., 1961, "On the calculus syntactic types" *Proceeding of symposia in Applied Mathematics*, vol. XII, America Mathematical Society, Providence, Rhode Island, 166-178.

Lecomte & Retoré, 2001: ‘Extending Lambek grammars: a logical account of minimalist grammars’, *Proceedings of the 39th meeting of the Association for Computational Linguistics*, Toulouse, July 2001, pp. 354--361.

Mazière, F., 2005 : *L'analyse du discours*, coll Que sais-je ? n°3735, PUF

Montague R., 1974: *Formal Philosophy. Selected Papers of Richard Montague*, edited and with an Introduction by R. H. Thomason, New Haven & Londres, Yale University Press.

Moortgat, M., 1997: *Categorial Type Logics* in (van Benthem et ter Meulen eds) *The Handbook of Logics and Language*, North-Holland,

Moortgat, M. 1999: ‘Constants of grammatical reasoning’. In Bouma, Hinrichs, Kruijff and Oehrle (eds.) *Constraints and Resources in Natural Language Syntax and Semantics*, CSLI, Stanford,

Parigot, M., 1992: ‘λμ-calculus: an algorithmic interpretation of classical natural deduction’, in Voronkov, ed. *Proceedings of the International Conference on Logic Programming and Automated reasoning*, LNAI 624, Springer

Pêcheux, M., 1975 : *Les Vérités de La Palice*, Maspero

Petersen, W., 2003 : ‘A set-theoretical investigation of Panini's Sivasutras’. In: *Proceedings of the Eighth Meeting on Mathematics of Language* (MOL-8). Bloomington, Indiana, July 20-22, 2003.

Pollard, C, 1988 : ‘Categorial grammar and phrase structure grammar : an excursion on the syntax-semantics frontier’ in Oehrle, Bach and Wheeler (eds) *Categorial Grammars and Natural Language Structures*, Kluwer

Pollard, C. & Sag, I., 1987 *Information-Based syntax and Semantics*, vol. 1 Fundamentals, CSLI, Stanford

Pollard, C. & Sag, I., 1994 *Head-Driven Phrase Structure Grammar*, University of Chicago Press, Chicago

Pollard, C., 2001 : ‘Higher-order grammar, a categorical foundation for type-logical constraint based grammar’, *Proceedings of Formal Grammar*, Vienne,

Retoré, C. ‘Les mathématiques de la linguistique computationnelle. Premier volet: la théorie des langages’. *La Gazette des mathématiciens* (Société Mathématique de France). Volume 115 janvier 2008 pp. 35 - 62

Retoré, C. ‘Les mathématiques de la linguistique computationnelle. Second volet: Logique’. *La Gazette des mathématiciens* (Société Mathématique de France). Volume 116 avril 2008 pp. 29-63

Retoré, C.& Stabler, E.,2004 : ‘Resource logics and minimalist grammars’. *Research on Language and Computation* 2(1), pages 3-25.

Spelke, E. 2000, ‘Core Knowledge’, *American Psychologist*, 55, 1233-1243

Spelke, E. & Tsivkin, S. 2001: ‘Language and number: a bilingual training study’, *Cognition*, 78, 45-88

Stabler, E., 1997: ‘Derivational minimalism’. in *Logical Aspects of Computational Linguistics* (Retoré, ed).. Springer, 1997, pages 68-95.

Stabler, E., 2001: ‘Recognizing head movement’.in *Logical Aspects of Computational Linguistics*, de Groote, Morrill, Retoré (eds). Springer, 2001, pages 245-260.

Steedman, M., 1985: ‘Dependency and coordination in the grammar of Dutch and English’, *Language* n°61, 523-68

Steedman, M., 1987: ‘Combinatory grammars and human language processing’, in *Modularity in Knowledge Representation and Natural Language Understanding*, (Garfield, ed.) 187-205, MIT Press, Cambridge,

Szabolcsi, A, 1997, *Ways of Scope Taking*, Kluwer

Vermaat, W. *The Logic of Variation*, PhD thesis de l’Université d’Utrecht, janvier 2006

Westerståhl, D. : ‘Quantifiers in formal and natural languages’, in Gabbay, Guenthner, *Handbook of Philosophical Logic*, vol. IV, Reidel, Dordrecht, 1989

Annexes

Annexe 1 : Les propriétés d'extension, de conservativité et d'isomorphie des QGs

1) Un déterminant δ représenté par une relation binaire Q sur les parties d'un univers D est dit satisfaire la propriété d'*extension* si pour tous A, B $\subseteq$ E $\subseteq$ E', $Q_E(A, B) \Leftrightarrow Q_{E'}(A, B)$, où Q_E désigne la restriction de Q aux intersections des parties de D avec E.

Cela signifie tout simplement que la vérité d'un énoncé se traduisant par Q(A, B) ne se modifie pas quand l'univers croît ou décroît (à condition évidemment d'inclure A et B). Comme l'ensemble le plus petit contenant à la fois A et B est leur union $A \cup B$, on en déduit que l'énoncé en question est évaluable en ne s'intéressant qu'aux individus de $A \cup B$. Par exemple le déterminant **tout** (**chaque**) possède évidemment cette propriété puisque TOUT(A, B) se définit par l'inclusion de A dans B (peu importent les éléments extérieurs à A ou à B !), idem pour **la plupart** qui ne fait que comparer le nombre Card($A \cap B$) au nombre Card(A – B).

2) Un déterminant δ représenté par une relation binaire Q sur les parties d'un univers D est dit satisfaire la propriété de *conservativité* si pour tous A, B $\subseteq$ E, $Q_E(A, B) \Leftrightarrow Q_E(A, A \cap B)$

3) Un déterminant δ représenté par une relation binaire Q sur les parties d'un univers D est dit satisfaire la propriété d'*isomorphie* si pour toute permutation π de E et pour tous A, B $\subseteq$ E, $Q_E(A, B) \Leftrightarrow Q_E(\pi(A), \pi(B))$, autrement dit la relation ne dépend pas des identités des individus dénotés.

Annexe 2 : Les tratégies d'évaluation "call by name / call by value"

Ces deux expressions désignent différents ordres d'évaluation des sous-expressions dans une forme symbolique. Dans les langages de programmation fonctionnelle (LISP, Scheme…), les « programmes » sont des expressions symboliques qui représentent généralement des applications de certaines expressions (jouant le rôle de fonctions) à d'autres expressions (jouant le rôle d'arguments). On peut avoir par exemple une expression définie au moyen d'une expression particulière `define` (cette dernière agissant bien comme une fonction, mais avec un effet de bord : celui consistant à mettre en mémoire de la machine l'association entre un nom de fonction, ici `try`, et une forme symbolique, qu'on appelle le corps de la fonction), comme :

```
(define (try a b)
        (if (=0 a)
             1
            b))
```

la function `try` ainsi définie retourne la valeur 1 si a est égal à 0 et la valeur de b sinon. L'expression symbolique suivante est alors bien formée :

```
(try 0 (/ 1 0))
```

elle représente l'application de la fonction `try` aux arguments : 0 et 1/0 (qui bien sûr en principe provoque une erreur). On voit que le compilateur d'un tel langage peut réagir de deux manières différentes selon la façon dont il évalue les arguments :

- s'il évalue les arguments *avant* d'appliquer `try` aux valeurs trouvées, bien sûr, il va renvoyer un message d'erreur : 1 n'est évidemment pas divisible par 0,
- si en revanche il évalue les arguments *le plus tard possible*, il va renvoyer 1 puisqu'on est dans la situation où a = 0 : inutile d'aller voir la valeur du second argument.

La première stratégie suit *l'ordre applicatif* et est appelée stratégie « call by value », la seconde est dite suivre *l'ordre normal* et est appelée stratégie « call by name ».

En ce qui concerne la phrase (26), le syntagme verbal peut être évalué de deux manières : ou bien il est évalué après que les deux constituants l'aient été (« call by value ») ou bien le verbe est d'abord évalué, puis le complément (« call by name »). Dans un cas, le quantificateur de l'objet pourra avoir une portée plus large que celui du sujet, dans l'autre, il ne le pourra pas.

Annexe 3 : Le $\lambda\mu$-calcul

Quand on présente les logiques intuitionniste et classique sous la forme de calculs des séquents, ce qui les différencie au premier coup d'œil, c'est que les séquents intuitionnistes ont une seule formule en partie droite alors que les séquents classiques ont un nombre arbitraire de formules à droite. Les séquents intuitionnistes s'interprètent ainsi facilement en termes fonctionnels : les éléments de la partie gauche sont les *inputs* et l'élément de la partie droite est l'*output*, ce qui est cohérent avec la définition même d'une fonction (une seule valeur est associée à un choix donné de valeurs pour les arguments), d'où le *déterminisme* du calcul intuitionniste (le fait qu'il soit *confluent*). En logique classique, comme il y a plusieurs formules en partie droite, il faut à chaque pas d'un calcul pouvoir sélectionner celle sur laquelle on travaille et qui devient de ce fait la formule courante à droite. Pour cela, il faut d'une part pouvoir nommer les formules et d'autre part pouvoir sauter directement à une formule nommée en utilisant simplement son nom. Cette manière de faire a alors son corollaire du point de vue d'une extension du λ-calcul. Des noms peuvent être attribués à des sous-termes et au cours de l'évaluation (donc au cours de l'exécution du programme) on doit pouvoir sauter[20] directement à certains sous-termes uniquement à partir de leurs noms

[20] que l'on pense à l'instruction `goto`, très traditionnelle en programmation.

(on dit aussi « étiquettes »). Les termes du λμ-calcul contiennent donc des termes nommés et des termes non nommés, qui sont définis récursivement de la manière suivante (Parigot, 1990) :

- x est un terme non nommé, si x est une λ-variable,
- λx.u est un terme non nommé, si x est une λ-variable et u un terme non nommé,
- (t u) est un terme non nommé, si t et u sont des termes non nommés,
- μα.e est un terme non nommé, si e est un terme nommé et α une μ-variable,
- [α]t est un terme nommé, si t est un terme non nommé et α une μ-variable.

Les règles de réduction de base sont :

- **Réduction logique** : (λx.u v) > u[v/x]
- **Réduction structurelle** : (μβ.u v) > μβ.u[[β](w v)/[β]w]
- **Renommage** : [α]μβ.u > u[α/β]

où u[[β](w v)/[β]w] est obtenu à partir de u en remplaçant inductivement chaque sous-terme de la forme [β]w par [β](w v).
commentaire : la règle de réduction logique est standard, c'est la β-réduction du λ-calcul,
la règle de réduction structurelle dit qu'appliquer un terme μβ.u à v, c'est remplacer dans u tous les sous-termes nommés β, de la forme donc : [β]w, par [β](w v), autrement dit aller chercher chaque sous-terme distingué et lui donner v comme argument supplémentaire.
Le système logique servant au typage des λμ-termes contient les règles classiques pour les parties lambda, étendues au second ordre (le système AF_2 de Krivine), mais rajoute des règles spécifiques pour le nommage :

$$\frac{t : \Gamma \rightarrow A, \Delta}{[\alpha]t : \Gamma \rightarrow A^{\alpha}, \Delta} \qquad \frac{e : \Gamma \rightarrow A^{\alpha}, \Delta}{\mu\alpha.e : \Gamma \rightarrow A, \Delta}$$

où on voit bien que l'étiquette donnée à un terme t correspond à la distinction d'une formule dans la partie droite et la μ-abstraction au fait de transformer une formule étiquetée en une formule courante (active). Comme on est en logique classique (plusieurs formules en partie droite), il faut régler ce qui concerne la négation afin d'avoir un système complet. ¬A est définie comme A ⇒ ⊥, où ⊥ est le type « absurde » (ou bien celui des entités observables dans certaines interprétations comme celle proposée par de Groote 2001). On a les quatre règles :

$$\frac{u : \quad \Gamma, A^{x} \rightarrow \bot, \Delta}{\lambda x.u : \Gamma \rightarrow \neg A\ , \Delta} \qquad \frac{t : \Gamma \rightarrow \neg A\ , \Delta \qquad u : \Gamma' \rightarrow\ A\ , \Delta'}{(t\ \ u) : \Gamma, \Gamma' \rightarrow \bot, \Delta\ , \Delta'}$$

$$\frac{t : \quad \Gamma \rightarrow \bot, \Delta}{[\alpha]t : \quad \Gamma \rightarrow \Delta} \qquad \frac{e : \Gamma \rightarrow \Delta}{\mu\alpha.e : \Gamma \rightarrow \bot, \Delta}$$

Dans la dernière règle, e n'a pas de sous-terme étiqueté α.

Noter que les λ-variables apparaissent uniquement à gauche et les μ-variables uniquement à droite. La règle d'introduction à droite de la négation permet de changer la place de la formule A et de la faire passer à droite, au prix de sa négation. La règle « de contradiction » est en haut à droite. Le nommage de la formule « absurde » permet de la faire disparaître, alors que la μ-abstraction au moyen de la variable α sur une expression e qui ne possède pas de sous-terme étiqueté α revient à faire réapparaître la formule « absurde ». Comme le fait remarquer Parigot, grâce aux règles qui permettent de changer les formules de côté, (chaque formule nommée dans les conclusions peut être remplacée par sa négation parmi les hypothèses), on peut se ramener à un système avec au plus une conclusion en partie droite. En ce cas l'opérateur μ n'est plus utilisé que pour gérer l'absurdité. D'où la conclusion que « l'opérateur μ est, dans un certain sens, le contenu algorithmique de la règle de l'absurdité ». De Groote utilise cet aspect pour formaliser la sémantique des langues naturelles, en faisant de **t** (le seul type « observable ») l'équivalent de ⊥. Dans la formulation qu'il donne du problème de l'ambiguïté de portée des quantificateurs, le λ-terme associé à *toute personne* : λP.λQ.∀x **personne**(x) ⇒ Q(x) est remplacé par le λμ-terme : μα.λP.∀x **personne**(x) ⇒ α(x), à cause de la preuve suivante.

Admettons que **personne** soit une constante de type <**e**, **t**> (donc ¬**e** dans la traduction adoptée) et que ⇒ soit de type <**t**, <**t, t**>>. Considérons les axiomes : α : ¬**e** → ¬**e** et x : **e** → **e**, on obtient par la règle de contradiction : [α]x : ¬**e, e** → **t**. Avec l'axiome non logique correspondant à **personne**, on obtient par ailleurs :

x **: e** → **personne**(x) : **t,**

et avec celui de ⇒, on obtient :

α : ¬**e,** x : **e** → **personne**(x) ⇒ [α]x : **t**

d'où par la règle d'introduction de ∀ :

α : ¬**e** → ∀x **personne**(x) ⇒ [α]x : **t**.

et d'après la règle de μ-abstraction :

$\rightarrow \mu\alpha.\ \forall x\ \mathbf{personne}(x) \Rightarrow [\alpha]x : \mathbf{e}$

(Bien voir ici la différence entre la représentation sémantique classique par un λ-terme de type **<<e, t>, t>** et la « nouvelle » représentation au moyen d'un $\lambda\mu$-terme qui est de type **e**, autrement dit, pour passer de l'une à l'autre, on a appliqué l'équivalent de la règle de double négation).