

Laurent GOSSELIN*

Le traitement de la polysémie contextuelle dans le calcul sémantique

Cet article expose les problèmes théoriques et pratiques liés au traitement calculatoire de la polysémie contextuelle, et propose un modèle hypothético-déductif d'analyse des marqueurs de temps et d'aspect en français.

Mots-clés : polysémie sémantique, calcul sémantique, contexte.

On the Processing of Contextual Polysemy in Semantic Computation. This paper presents theoretical and empirical problems relative to the computation of contextual polysemy, and proposes an hypothetico-deductive model for the analysis of time and aspect markers in french.

Key Words : polysemic semantics, computational semantics, context.

PRESENTATION

Le calcul sémantique utilisé dans une perspective d'analyse prend pour entrées des *formes* linguistiques (des phrases, des textes, envisagés uniquement du point de vue morpho-syntaxique) pour leur associer des *représentations sémantiques* en sortie. Ces représentations doivent être conçues selon un format tel qu'elles puissent à leur tour servir d'entrées pour un calcul logico-pragmatique capable d'extraire des *informations* explicites ou implicites (même s'il s'avère qu'une telle séparation entre sémantique et pragmatique est trop rigide et que certaines contraintes pragmatico-référentielles très générales doivent être prises en compte dès le traitement sémantique).

* Laurent GOSSELIN, Université de Rouen, CNRS (URA 1164), Université de Rouen, IRED, 76821, Mont Saint-Aignan cedex, France.

Le calcul sémantique rencontre avec la *polysémie contextuelle* un obstacle décisif dans la mesure où, comme on va le montrer, le mode de traitement retenu pour ce phénomène est indissociable de la forme même de la théorie qui préside au calcul.

Par "polysémie contextuelle", on désigne le fait qu'un marqueur (morphème lexical, grammatical, ou construction syntaxique) puisse prendre des significations au moins partiellement différentes en fonction des contextes linguistiques dans lesquels il se trouve. Comme ces contextes sont composés de marqueurs qui sont eux-mêmes, le plus souvent, polysémiques, on parlera de "*polysémie contextuelle généralisée*" (désormais *PCG*) pour désigner le fait que la signification d'un marqueur puisse varier en fonction non seulement des formes, mais aussi des significations des autres marqueurs qui l'entourent (lesquelles varient aussi de semblable façon). Autrement dit, adopter le point de vue de la PCG, c'est refuser l'attitude courante qui consiste à isoler et à privilégier tel marqueur polysémique, pris pour objet d'étude, pour n'envisager le contexte que comme un ensemble de formes ; c'est, au contraire, essayer de penser et, si possible, de calculer l'interaction globale des marqueurs polysémiques dans l'énoncé, et au-delà, dans le texte. Prenons un exemple :

(1) (*À cette époque*) *Pierre dormait en cinq minutes*

Dans cet énoncé – qui accepterait une suite du type "*mais maintenant, il est obligé de lire pendant deux heures avant de pouvoir s'endormir*" – l'imparfait prend une valeur itérative, *dormir* équivaut à *s'endormir*, et *en cinq minutes* désigne non pas la durée du procès lui-même, mais celle de la phase préparatoire de ce procès (la durée qui sépare, par exemple, le moment où Pierre se couche de celui où il s'endort). Ces significations sont dépendantes du contexte linguistique, car elles n'apparaissent plus dans les énoncés suivants :

(2) *Pierre dormait, je l'ai réveillé* (où *dormir* ? *s'endormir* et où l'imparfait n'est pas itératif)

(3) (*À cette époque*) *Pierre mangeait en dix minutes* (l'imparfait est itératif, mais *en dix minutes* désigne la durée du procès lui-même).

De plus, il s'agit bien de PCG, car ces différentes significations contextuelles sont liées les unes aux autres et certaines d'entre elles s'impliquent réciproquement : c'est parce que le procès est interprété dans l'exemple (1) comme ponctuel et inchoatif que le circonstanciel

de durée porte sur sa phase préparatoire, mais c'est aussi à cause de ce circonstanciel que le procès est "contracté" sur sa phase initiale, etc.¹

Ce simple exemple suffit à illustrer combien les phénomènes aspectuo-temporels du français sont polysémiques (qu'on songe, entre autres, aux multiples valeurs des temps verbaux répertoriées par les grammaires) et, partant, notoirement rebelles à toute entreprise de théorisation prédictive (au point qu'un consensus minimal n'a pas même pu être obtenu sur les concepts utiles à leur description²). Aussi est-ce ce domaine que nous avons retenu pour tester un modèle de la PCG, qui calcule les significations contextuelles à partir des formes des marqueurs utilisés. Plus précisément, il s'agit d'un modèle du temps et de l'aspect en français qui permet l'assignation de représentations aspectuo-temporelles (incluant l'ensemble des relations chronologiques) aux énoncés et aux textes sur la base des marques (lexicales et syntaxiques) qu'ils présentent (ainsi que de quelques principes pragmatico-référentiels très généraux). Ce modèle a été implémenté sous la forme d'un système expert (conçu au moyen du générateur de systèmes experts HEXPERT), puis, plus récemment, dans un système programmé en C++, par un informaticien de l'université de Caen, Y. Legendre³. Ces deux implémentations ont permis de tester le modèle sur un très grand nombre d'énoncés construits ou attestés, et de corroborer ainsi un ensemble d'hypothèses nouvelles sur la nature de ces phénomènes. Les résultats obtenus au moyen de cette démarche expérimentale sont consignés dans une grammaire calculatoire du temps et de l'aspect en français (Gosselin 1996).

On évoque ici les problèmes généraux que rencontre la modélisation de la PCG, pour présenter ensuite les solutions mises en oeuvre dans notre modèle du temps et de l'aspect.

1. LA POLYSEMIE CONTEXTUELLE GENERALISEE : PROBLEMES GENERAUX

1.1. Les principes

¹ Pour d'autres exemples, et une analyse similaire, cf. B. Victorri (1996).

² Cf. par exemple l'avertissement d'O. Ducrot et J.-M. Schaeffer (1995) : "Les théories et les terminologies relatives à la temporalité linguistique sont si variées et contradictoires que nous avons préféré donner une présentation personnelle du problème".

³ Pour tous renseignements, prendre contact avec Y. Legendre, Laboratoire ELSAP, Université de Caen.

Toute sémantique grammaticale visant un certain niveau de formalisation doit prendre position vis-à-vis de deux principes, dont on attribue, à tort ou à raison, la paternité à Frege⁴ :

- a) le principe de *compositionnalité* : la signification du tout (un énoncé, par exemple) est déterminée par celles de ses parties ;
- b) le principe de *contextualité* : la signification d'une expression est au moins partiellement déterminée par le contexte dans lequel elle apparaît.

Le point de vue de la PCG conduit à reformuler le principe de contextualité comme suit :

- c) la signification d'un marqueur est déterminée par celle du tout dans lequel il apparaît.

Ce qui en fait l'exact symétrique du principe de compositionnalité.

De là, deux attitudes qui s'opposent actuellement en sémantique grammaticale :

- a) l'*atomisme*, défendu massivement par les sémantiques formelles et la psycholinguistique (sous la forme de la théorie de l'accès au lexique mental), qui consiste à retenir le principe de compositionnalité, mais à ignorer le principe de contextualité ;
- b) le *holisme*, représenté à la fois par la grammaire cognitive, qui le revendique comme tel, et, dans une certaine mesure, par la tradition sémantique française, qui reconnaît la validité simultanée des deux principes⁵.

Il faudrait encore mentionner la tendance – car il s'agit rarement d'une position explicitement défendue⁶ – illustrée par certains courants pragmatiques, à minimiser le rôle de la compositionnalité, et donc de la signification proprement linguistique, pour insister, au contraire, sur

⁴ Sur ces principes, cf. P. Engel (1994) , p. 10.

⁵ Une mise en garde terminologique s'impose. Très souvent le terme de "compositionnalité" est employé comme équivalent de "compositionnalité atomiste", d'où le fait que des systèmes compositionnels holistes soient parfois présentés comme "non-compositionnels".

⁶ Dans une tout autre perspective, le rejet explicite de la compositionnalité de la signification par J.-Cl. Milner (1989) le conduit à mettre en doute la possibilité même d'une sémantique linguistique, cf. pp. 311-314.

celui de la situation de discours et des processus inférentiels dans la détermination de la signification globale assignée à l'énoncé⁷.

Nous passons – très rapidement – en revue les problèmes rencontrés par ces différentes approches.

1.2. L'atomisme

L'atomisme ignore (ou essaie de le faire) la polysémie contextuelle. Autrement dit, dans sa forme stricte, il suppose l'invariabilité contextuelle des significations individuelles des marqueurs. Le but d'une sémantique atomiste se limite donc à proposer des règles de composition sémantique efficaces. En général, ces règles se conforment aux représentations hiérarchisées produites par la syntaxe : on "remonte" les significations à partir des noeuds terminaux, en effectuant des compositions de significations à chaque noeud branchant de l'arborescence. Si bien que l'on a pu dire que la compositionnalité était considérée par les tenants de l'atomisme comme la contrepartie sémantique de la récursivité des règles syntaxiques⁸. Adopter une démarche atomiste en sémantique suppose donc que l'on dispose 1) d'une théorie syntaxique et 2) d'un format de représentation sémantique⁹ ; les règles de composition se réduisent dès lors le plus souvent à remonter les éléments de représentation sémantique dans les arbres et à opérer les compositions selon l'ordre imposé par le type de structure syntaxique mis en oeuvre ; de sorte que comme le remarque Fr. Rastier¹⁰, la sémantique n'est, dans cette perspective, qu'un simple "décalque" de la syntaxe, tout problème sémantique devant trouver sa solution au plan de la structuration syntaxique. Mais tout ceci suppose que l'on ait réussi à "réduire" la variation contextuelle des significations, empiriquement attestée. Deux types de stratégies mises en oeuvre à cette fin peuvent être distinguées : 1) l'approche *homonymique*, qui consiste à traiter les diverses significations contextuelles d'un même marqueur comme correspondant à autant de marqueurs différents (on distingue, par exemple, un imparfait 1 : inaccompli, un imparfait 2 : itératif, un imparfait 3 : "de rupture", etc.), 2) la procédure de *conversion* (ou "recatégorisation") qui admet, tout de même, que, dans certains contextes, une substitution de traits puisse être opérée sur

⁷ Pour une critique de ce type d'approche appliquée au calcul de la référence, cf. G. Kleiber (1994), en particulier pp. 12-13.

⁸ Cf. J. Hintikka (1994), p. 5 sq.

⁹ Cf. l'exemple des grammaires d'unification : GPSG reprend la sémantique de Montague ; HPSG la sémantique des situations.

¹⁰ Cf. Fr. Rastier, M. Cavazza et A. Abeillé (1994), p. 111.

certain marqueurs ; ainsi, un procès intrinsèquement perfectif deviendrait imperfectif lorsque le verbe est conjugué à l'imparfait inaccompli (imperfectif : *il mangeait une pomme*)¹¹.

La "solution" homonymique n'est plus guère défendue aujourd'hui, en raison de son caractère totalement artificiel, mais aussi parce qu'il est, dans la pratique, très difficile d'arrêter une liste exhaustive d'acceptions distinctes (censées correspondre à autant d'homonymes) pour un même marqueur polysémique (on reviendra sur ce problème, très général). En revanche, la procédure de conversion, qui essaie de préserver les avantages de l'atomisme (voir ci-dessous), tout en tenant compte – il est vrai de façon très restrictive – des effets contextuels, est généralement admise et même tenue pour cognitivement plausible : il paraît raisonnable que la composition de deux significations porteuses de traits incompatibles soit résolue par l'effacement de l'un de ces traits. Elle est pourtant, elle aussi, contestable du point de vue empirique, car il n'est pas vrai qu'un procès perfectif devienne à proprement parler imperfectif lorsqu'il est conjugué à l'imparfait, comme l'indique le fait qu'il reste incompatible avec [*pendant* + durée], qui sert pourtant à identifier les procès imperfectifs :

(4) ?? *Il écrivait sa lettre pendant cinq minutes.*

De façon plus générale, la conversion qui consiste à *simplifier* les données (par effacement de traits) ne semble pas rendre compte de phénomènes *complexes*, qui combinent, à des niveaux différents, des caractéristiques contradictoires. Ainsi dans l'énoncé :

(5) *Il écrivait sa lettre*

le procès paraît bien intrinsèquement borné (télique ou perfectif ; il ne s'agit pas d'une activité qui pourrait se prolonger indéfiniment), mais à cause de l'imparfait, la borne finale n'est pas présentée comme atteinte ; d'où l'impossibilité d'inférer (6) à partir de (5) (ce phénomène est connu sous le nom de "paradoxe imperfectif"¹²) :

(6) *Il a écrit sa lettre.*

Perdre, au terme de la procédure de conversion, l'information selon laquelle le procès est intrinsèquement perfectif (il s'agit d'un accomplissement¹³) ne permettrait plus de rendre compte du paradoxe

¹¹ Cf. entre autres D. Dowty (1986), M. Moens et M. Steedman (1988), W.S. Bennett, T. Herlick, L. Hoyt, J. Liro et A. Santisteban (1989-90), C. Vet (1991).

¹² Cf. D. Dowty (1977) ; pour une présentation critique, cf. J. François (1981).

¹³ Cf. Z. Vendler (1967).

imperfectif, car les procès intrinsèquement imperfectifs (atéliques), référentiellement homogènes¹⁴, autorisent l'inférence au passé composé (*Il dormait* \AE *Il a dormi*).

1.3. Le holisme

Le holisme est déjà, d'une certaine façon, exprimé par Saussure à propos de la valeur syntagmatique des signes :

"Le tout vaut par ses parties, les parties valent aussi en vertu de leur place dans le tout, et voilà pourquoi le rapport syntagmatique de la partie au tout est aussi important que celui des parties entre elles"¹⁵

ou, dans une perspective proche, par Merleau-Ponty :

"À chaque moment, sous le système de la grammaire officielle, qui attribue à tel signe telle signification, on voit transparaître un autre système expressif qui porte le premier et procède autrement que lui : *l'expression*, ici, n'est pas ordonnée point par point, à l'exprimé ; chacun de ses éléments ne se précise et ne reçoit l'existence linguistique que par ce qu'il reçoit des autres et par la modulation qu'il imprime à tous les autres. C'est le tout qui a un sens, non chaque partie"¹⁶.

Mais il paraît incompatible avec le projet calculatoire, car pour opérer un calcul, il est nécessaire d'avoir des données relativement stables en entrée. Comme les significations individuelles des marqueurs sont tenues pour contextuellement variables, elles ne peuvent servir d'entrées à un système calculatoire et prédictif. On peut, au mieux, expliciter des contraintes de cohérence (i.e. de concordance des traits¹⁷ du type "si tel marqueur prend telle signification, alors tel autre marqueur prendra telle signification [porteuse d'un trait identique]", mais, comme il apparaît, ces règles prennent nécessairement pour entrées des valeurs sémantiques, qui devraient elles-mêmes avoir déjà été calculées¹⁸. On aboutit donc à une sorte de dilemme : le point de

¹⁴ Cf. B. Taylor (1977), A. Mourelatos (1978), L. Carlson (1981).

¹⁵ Saussure (1978), p. 177.

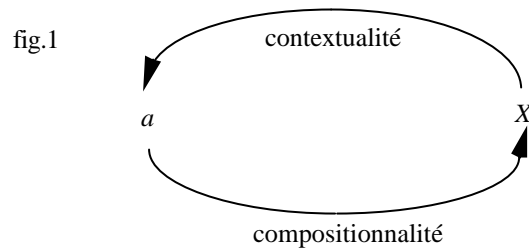
¹⁶ Merleau-Ponty (1969), pp. 40-41.

¹⁷ Voir, par exemple, les "isosémies" dans la sémantique différentielle de Fr. Rastier, dans Rastier, Cavazza et Abeillé (1994), p. 118 sq.

¹⁸ M. Cavazza le reconnaît explicitement : "Du point de vue pratique, c'est-à-dire dans le cas de simulations informatiques de la sémantique différentielle dans des modèles informatiques classiques (c'est-à-dire symboliques), les interactions

vue atomiste est compatible avec la perspective calculatoire, mais ne paraît pas adéquat à son objet (il ne peut traiter la PCG) ; le holisme prend en compte la PCG, mais semble incompatible avec le calcul. Faut-il, comme G. Lakoff, renoncer à la démarche calculatoire et prédictive¹⁹ ? Mais il resterait encore à montrer – contre l'épistémologie popperienne – qu'un tel renoncement n'altère en rien la scientificité de l'analyse²⁰, ou à défendre l'idée que la sémantique linguistique ne saurait – quelle que soit la méthode utilisée – prétendre à un même type de scientificité que les sciences de la nature.

En fait, le holisme n'est incompatible avec le calcul sémantique que pour autant qu'il inscrit la double relation, de compositionnalité et de contextualité, entre la partie (le marqueur) et le tout (la phrase, l'énoncé, le texte) dans un cercle représentable comme suit :

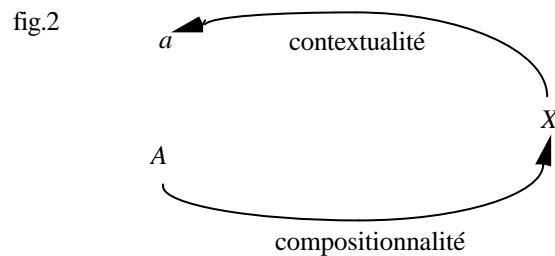


où *a* désigne une signification contextuelle particulière d'un marqueur et *X*, la signification globale du tout. *a*, en relation avec d'autres marqueurs, détermine *X* en vertu du principe de compositionnalité, tandis que, selon le principe de contextualité, *X* détermine *a*. Pour échapper à une telle circularité, la seule solution paraît être de distinguer entre une signification abstraite et invariable, associée au marqueur (notée *A*) et ses significations en contexte (*a1*, *a2* ...). On obtient alors une structure en came du type :

global/local seront entièrement du ressort de la description des contenus sémantiques par l'opérateur comme préalable à leur implémentation" (Rastier, Cavazza et Abeillé (1994), p. 87).

¹⁹ Cf. G. Lakoff (1987), p. 465 : "other theories of grammar assume some form of atomism, namely, that the meaning of a grammatical construction is a computable function of the meanings of its parts. We will argue instead that grammatical constructions in general are holistic, that is, that the meaning of the whole construction is motivated by the meanings of the parts, but is not computable from them."

²⁰ Pour une tentative en ce sens, cf. Cl. Vandeloise (1991).



où la signification *A*, parce qu'elle est invariablement associée au marqueur considéré, peut constituer une entrée au système de calcul.

De Saussure à Culioli, en passant par Guillaume, toute une tradition européenne d'étude sémantique de la polysémie repose sur cette distinction – même si elle prend des formes et des portées diverses – entre des significations ou valeurs abstraites et stables, associées aux marqueurs dans le système de la langue, et des significations empiriquement observables, et variables parce que partiellement déterminées par le contexte dans lequel elles se rencontrent.

Si cette distinction théorique rend le calcul holiste envisageable, il reste encore – et la tâche n'est pas mince – à élaborer une méthode permettant d'identifier ces valeurs stables.

1.4. L'identification des significations hors contexte

Admettre, comme il vient d'être proposé, que l'ensemble des significations contextuelles d'un marqueur provient de l'interaction d'une valeur en langue (*A*), unique et stable, avec les valeurs des autres marqueurs du contexte conduit immédiatement à une difficulté : la valeur *A*, qui – dans la perspective saussurienne – constitue l'identité différentielle du signe, n'est pas observable. Elle ne peut être appréhendée qu'indirectement, à partir des effets de sens contextuels qui, eux, sont observables au moyen, en particulier, des relations de synonymie. Prenons pour exemple les principaux effets de sens aspectuo-temporels du présent de l'indicatif, identifiés au moyen de paraphrases :

(7a) *Pierre mange* (~ est en train de manger)

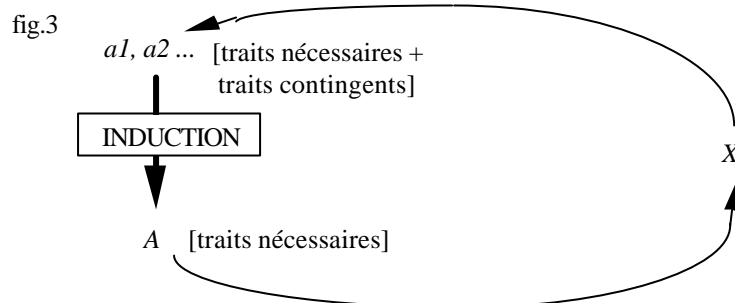
(7b) *Pierre arrive de Lyon* (~ vient d'arriver de Lyon)

(7c) *Pierre part (ce soir) pour Paris* (~ va partir pour Paris)

(7d) *Pierre fume la pipe depuis deux ans* (~ est un fumeur de pipe)

(7e) *À ce moment-là, Pierre entre (~ entra) et aperçoit (~ aperçut) Marie qui dort (~ était en train de dormir).*

La procédure de recherche qui doit donc conduire des significations contextuelles (ou effets de sens) observées ($a1, a2 \dots$) aux valeurs en langue (A) est très généralement de nature *inductive* : de la comparaison des divers effets de sens d'un même marqueur, on retient un ensemble de traits sémantiques nécessaires (i.e. qui se retrouvent dans chacun des effets de sens). On admet que le contexte "enrichit" ces significations A en leur adjoignant des traits supplémentaires, contingents, pour constituer les effets de sens, selon un schéma du type :



Cette démarche inductive rencontre certaines difficultés théoriques et empiriques qui nous conduisent à douter de sa fiabilité, et donc de son efficacité dans la perspective du calcul sémantique :

I) Un problème d'épistémologie générale tout d'abord : adopter une démarche inductive dans une pratique scientifique suppose que l'on soit en mesure d'apporter une réponse au fameux "problème de l'induction" de Hume, développé par K. Popper. Nous ne discuterons pas ce point ici.

II) Comme il n'existe pas, à proprement parler de "logique de l'induction" qui soit de quelque façon comparable à la logique déductive²¹, l'induction d'une signification abstraite (A) doit se laisser guider non par des règles formelles, mais simplement par l'intuition du chercheur. De là, la double hypothèse, généralement tacite, selon laquelle a) les phénomènes et les explications sémantiques doivent être accessibles à l'intuition, et, partant, b) qu'ils doivent être relativement *simples*. La première hypothèse sous-tend les jugements critiques couramment émis par les sémanticiens de tradition française, en particulier, sur le caractère "contre-intuitif" de telle ou telle analyse. Faudra-t-il rappeler que rien n'est plus contre-intuitif que la rotondité de la terre ? La seconde hypothèse conduit à considérer le calcul et

²¹ Cf. R. Swinburne (ed.) (1974).

l'usage de l'ordinateur comme n'étant nullement indispensables à la recherche en sémantique (il s'agirait, au mieux d'applications de la recherche linguistique fondamentale). En tout état de cause, cette double hypothèse devrait, à tout le moins, être explicitée et argumentée.

III) La procédure d'induction des valeurs *A*, qui ne retient que les traits sémantiques nécessaires (communs à l'ensemble des effets de sens qui lui correspondent) conduit, lorsque les expressions présentent un haut degré de polysémie, à des valeurs *A* extrêmement pauvres, presque vides de contenu, qui ne permettent même plus, dans certains cas, de distinguer au plan sémantique différents marqueurs (quelle pourrait être, par exemple, la valeur du présent de l'indicatif obtenue par induction à partir des effets de sens repérés ci-dessus ?). Un bon exemple de cette impasse est fourni par les recherches menées au cours des années quatre-vingt sur les valeurs du passé simple et de l'imparfait envisagées du point de vue de la chronologie relative. H. Kamp et Ch. Rohrer (1983) avaient formulé, dans le cadre de la Discourse Representation Theory et en empruntant le modèle temporel de Reichenbach, deux règles censées régir respectivement l'interprétation du passé simple et de l'imparfait dans les textes narratifs :

R1) Le passé simple introduit un événement passé (*e*) qui constitue le point de référence du texte ; chaque nouvel énoncé au passé simple fait avancer le point de référence d'un pas.

R2) L'imparfait introduit un état passé (*s*) qui contient le point de référence en cours *w* (introduit par la dernière phrase au passé simple)²².

Au cours de la décennie qui a suivi, ces deux règles ont été critiquées, sur la base de contre-exemples divers (dont les premiers avaient été fournis par Kamp et Rohrer eux-mêmes), et reformulées à plusieurs reprises²³. Au total, il est apparu que ces deux temps morphologiques pouvaient exprimer chacun toutes les relations chronologiques possibles²⁴. Puisqu'il ne paraissait donc plus envisageable de leur assigner une valeur chronologique, ni même de les distinguer de ce point de vue, il a semblé alors que seules des analyses fondées sur des principes pragmatiques et des relations discursives, négligeant même les caractéristiques aspectuelles qui leur sont traditionnellement

²² Nous traduisons.

²³ Voir, entre autres, B. Partee (1984), E.W. Hinrichs (1986), K. Eberle (1988), A. Molendijk (1990).

²⁴ On trouvera des exemples dans Gosselin (1996).

associées, seraient susceptibles de rendre compte de la chronologie relative entre procès dans les textes²⁵. Outre qu'une telle perspective, qui ne tient plus compte de la compositionnalité de la signification linguistique, ne saurait conduire à un véritable calcul sémantique, elle se heurte au fait, indéniable, que le passé simple et l'imparfait ne sont pas substituables l'un à l'autre en contexte sans conséquences pour la chronologie, ce qui veut dire qu'ils ne sont pas intrinsèquement dépourvus de contenu²⁶, mais que c'est la démarche inductive elle-même qui les en vide.

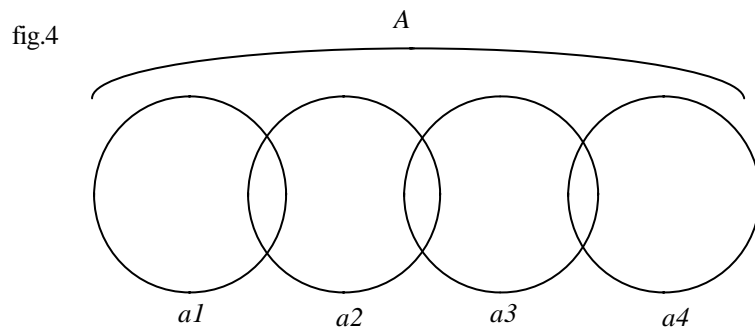
Pour répondre à cette difficulté, liée au traitement inductif de la polysémie contextuelle, la grammaire cognitive a repris à la théorie de la catégorisation en psychologie cognitive le concept de *prototype*. Il s'agissait, en fait, dans ce cas, d'assouplir les exigences liées à la démarche inductive : il n'est, dans ce nouveau cadre, plus indispensable que la valeur *A* recouvre tous et rien que les traits nécessaires des significations contextuelles (*a1*, *a2* ...) ; il faut et il suffit que chacune de ces significations partage au moins un trait sémantique avec la signification *A*²⁷. Outre qu'il paraît parfois bien arbitraire de choisir telle ou telle valeur comme prototype dans le domaine de la polysémie grammaticale²⁸, l'exigence liée au rattachement de tous les effets de sens à une même valeur prototypique s'est avérée encore trop forte pour rendre compte de la polysémie de certains marqueurs. Si bien qu'une seconde acception du prototype, fondée sur la notion de "ressemblance de famille" a dû être proposée, qui constitue, comme le souligne G. Kleiber (1990), une rupture avec la perspective antérieure, puisqu'elle revient à considérer la signification d'une expression (pour autant qu'on puisse encore dire que l'expression a une signification en langue) comme l'ensemble de ses effets de sens sémantiquement liés entre eux selon un schéma du type :

²⁵ Voir, en particulier, A. Lascarides et N. Asher (1992), N. Asher et M. Bras (1992)

²⁶ Une analyse linguistique de ces phénomènes – trop complexe pour pouvoir être présentée ici – est proposée dans Gosselin (1996), pp. 105-153.

²⁷ Cf. G. Kleiber (1990), pp. 161-162.

²⁸ Voir, par exemple, l'analyse du subjonctif en français par M. Winters (1991) qui pose, sans autres précautions, que "l'ensemble sémantique des emplois de ce mode se rassemble autour des expressions prototypiques directes et indirectes du doute" (p. 159).



Dans un tel système, le rôle du contexte s'inverse, il n'est plus d'enrichir une signification trop pauvre, mais de filtrer une signification trop riche²⁹.

IV) Quelles que soient les réponses apportées aux problèmes précédents, la démarche inductive se heurte au fait que les significations contextuelles, loin d'être simplement plus riches ou plus pauvres que les significations A, nous paraissent de nature essentiellement différente. Car ce que l'on décrit comme significations contextuelles des marqueurs recouvre, en fait, des caractéristiques de représentations sémantiques globales (X) associées aux énoncés, qui ne sauraient, en toute rigueur, être rapportées qu'à des combinaisons de marqueurs et non à l'un d'eux en particulier ; si bien que l'attribution d'un effet de sens à un marqueur peut être dénoncée comme arbitraire et illusoire : «il est bien arbitraire de déterminer, à l'intérieur d'un énoncé donné, quel est le sens qu'y possèdent, pris un par un, les mots dont il est constitué. Autrement dit, il ne nous semble pas du tout évident que le sens global de l'énoncé puisse être considéré comme la sommation des significations, mêmes contextuelles, des différents mots"³⁰. Pourquoi, par exemple, rapporter l'itération uniquement à l'imparfait dans "*Il dormait en cinq minutes*", alors qu'elle provient (comme on va l'expliquer) d'une interaction complexe entre le temps verbal et le circonstanciel de durée ? Pourquoi, en revanche, les grammaires ne considèrent-elles jamais l'itération comme une valeur possible du passé

²⁹ Il est remarquable que l'on trouve déjà ces deux types de conceptions du rôle du contexte, mises en oeuvre parallèlement, dans un cadre théorique apparemment bien différent : la psychomécanique guillaumienne telle qu'elle est appliquée par J. Picoche à l'étude du lexique. J. Picoche distingue, en effet, les polysèmes dont les "effets de sens" sont "subduits" par le "signifié de puissance", et les polysèmes sans subduction dont les effets de sens sont organisés autour d'un "noyau sémantique commun" ; Cf. J. Picoche (1986), pp. 17-46.

³⁰ O. Ducrot (1984), p. 50.

composé, alors que dans un énoncé comme *'Longtemps, je me suis couché de bonne heure'* elle provient aussi d'une interaction entre le temps verbal, l'adverbe antéposé et le procès exprimé par le verbe ? C'est précisément à cause de ce caractère illusoire des "significations contextuelles" qu'il s'est avéré, en pratique, impossible de les dénombrer (au moins en ce qui concerne les morphèmes grammaticaux) et qu'elles ne nous semblent nullement aptes à constituer le point de départ d'une procédure d'induction des valeurs *A* (quels que soient, par ailleurs, les problèmes soulevés par la démarche inductive).

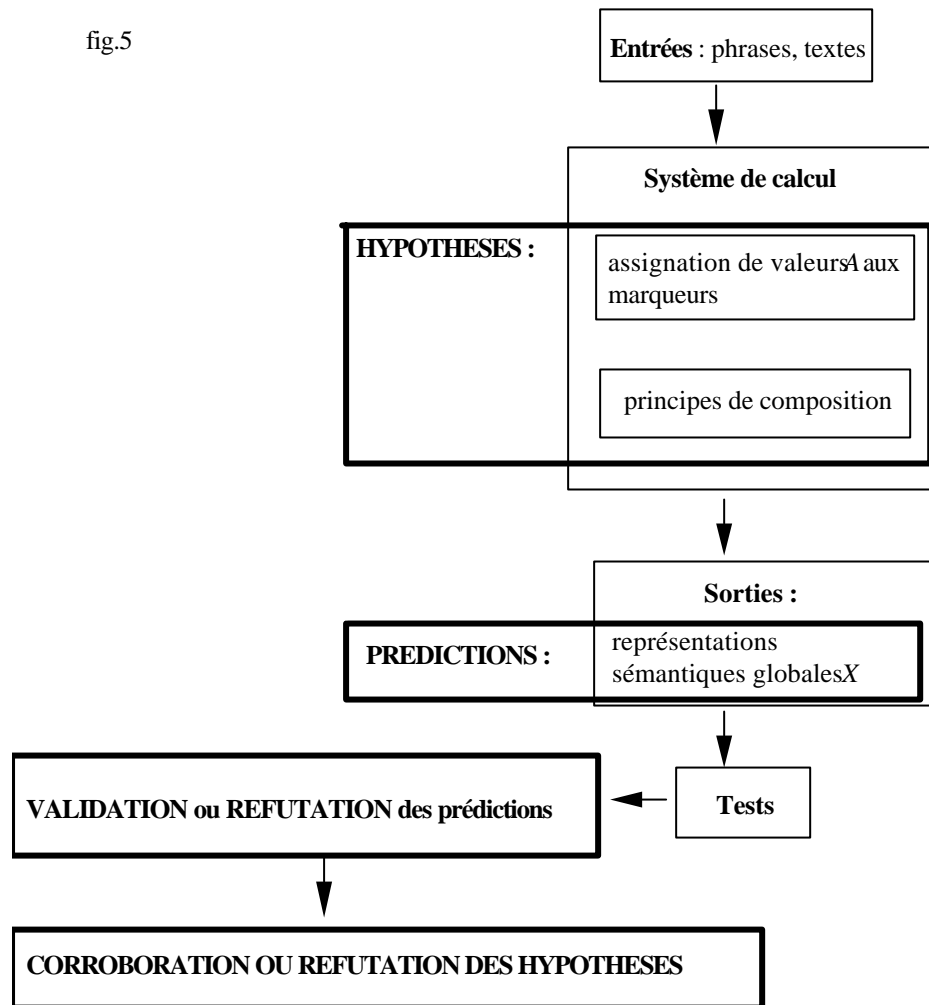
2. LE MODELE PROPOSE

2.1. Un démarche hypothético-déductive

Pour répondre à ces difficultés, nous avons opté pour l'abandon de la démarche inductive au profit d'une théorisation hypothético-déductive (conçue sur le modèle popperien³¹) telle que la description des valeurs *A* ait statut d'hypothèses à partir desquelles on puisse déduire des significations (ou plutôt des représentations sémantiques) globales associées aux énoncés, en prenant pour entrées les divers marqueurs (morphèmes et constructions syntaxiques) qu'ils présentent et en tenant compte de quelques principes généraux de composition sémantique. Ces représentations sémantiques globales, qui ont valeur de prédictions doivent alors être soumises à des tests, de façon à être validées ou réfutées. Toute réfutation d'une prédiction entraîne celle de l'une, au moins, des hypothèses de départ. Une hypothèse (nécessairement prise dans un corps d'hypothèses) est considérée comme corroborée si et seulement si elle a donné lieu à un vaste ensemble de prédictions, elles-mêmes soumises à des tests, et n'a pu être réfutée. Autrement dit, au lieu de chercher à découvrir une valeur *A* par induction à partir des "significations contextuelles" *a1, a2 ...*, de façon à pouvoir – seulement dans un second temps – calculer la représentation globale *X* et éventuellement la "signification contextuelle" *a* à partir de *A* et de son contexte, on utilise le calcul sémantique lui-même pour identifier la valeur *A* au moyen de la méthode des conjectures et réfutations ; ce que résume la figure :

³¹ Cf. K. Popper (1973) et (1985).

fig.5

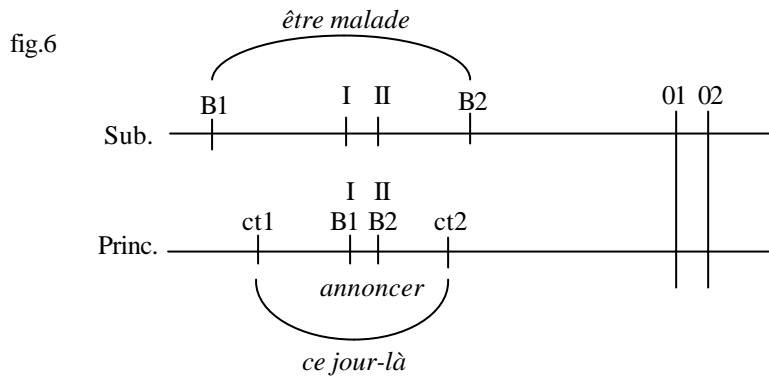


Reste qu'adopter une démarche hypothético-déductive suppose que l'on propose un modèle global d'un domaine sémantique. Car on ne peut plus s'en tenir à l'analyse de quelques marqueurs, comme c'est généralement le cas avec la méthode inductive. Seule une théorie est réfutable, ou, plus exactement, une hypothèse ne saurait être réfutée que dans le cadre d'une théorie globale, pour la simple raison qu'une hypothèse isolée ne permet pas la déduction de prédictions. C'est pourquoi nous nous sommes proposé d'analyser le domaine aspectuo-temporel dans son ensemble. Pour présenter ce modèle, nous évoquerons successivement 1) le format de représentation, 2) les principes de calcul, 3) les types de tests, et 4) l'implémentation informatique qui nous a permis de le tester.

2.2. Le format de représentation

Les représentations aspectuo-temporelles utilisées dans le modèle mettent en oeuvre quatre types d'intervalles disposés sur l'axe temporel : l'intervalle du procès [B1,B2]; l'intervalle d'énonciation [01,02], l'intervalle de référence [I,II], et l'intervalle circonstanciel [ct1,ct2]. L'intervalle de référence [I,II] correspond à ce qui est perçu/montré du procès (par ex., ce qui est asserté lorsque l'énoncé est assertif³²). À chaque énoncé est associé un et un seul intervalle d'énonciation [01,02] ; à chaque proposition (principale, subordonnée ou indépendante) sont associés au moins un intervalle de procès [B1,B2] et au moins un intervalle de référence [I,II] ; à chaque circonstanciel temporel correspond au moins un intervalle circonstanciel [ct1,ct2]. Exemple :

(8) *Ce jour-là, Pierre annonça qu'il était malade*



L'aspect grammatical se trouve défini par la relation entre l'intervalle du procès et l'intervalle de référence :

aspect aoristique : [I,II] coïncide avec [B1,B2]; ex. : *Il mangea sa soupe*

aspect inaccompli : [I,II] est inclus dans [B1, B2]; ex. : *Il mangeait sa soupe (depuis 5 min.)*

aspect accompli : [I,II] est postérieur à [B1,B2]; ex. : *Il a terminé sa soupe (depuis 5 min.)*

aspect prospectif : [I,II] est antérieur à [B1,B2]; ex. : *Il va sortir (car il est habillé).*

Le temps absolu correspond à la relation entre l'intervalle de référence et celui de l'énonciation :

passé : [I,II] est antérieur à [01,02]

³² Il s'agit approximativement de l'équivalent du *Topic Time* de W. Klein (1994).

présent : [I,II] et [01,02] coïncident³³

futur : [I, II] est postérieur à [01,02].

Le *temps relatif* est défini par la relation qui unit l'intervalle de référence de la subordonnée [I',II'] à celui de la principale [I,II] :

antérieur : [I',II'] est antérieur à [I,II]

simultané : [I',II'] et [I,II] coïncident

ultérieur : [I', II'] est postérieur à [I,II].

La situation temporelle des procès résulte de la combinaison du temps et de l'aspect. Exemples :

(9) *Il sortit son portefeuille*

[I,II] coïncide avec [B1,B2] (aspect aoristique)

[I,II] est antérieur à [01,02] (temps passé)

donc [B1,B2] est antérieur à [01,02].

(10) (*Quand j'ai été voir*) *le petit dormait*

[I,II] est inclus dans [B1,B2] (aspect inaccompli)

[I,II] est antérieur à [01,02] (temps passé)

donc B1 est antérieure à 01, mais la relation entre B2 et [01,02] n'est pas linguistiquement contrainte (le procès peut très bien se poursuivre dans le présent et même dans le futur)³⁴.

(11) *Marie croyait que Paul viendrait*

[I,II] est inclus dans [B1,B2] (aspect inaccompli dans la principale)

[I,II] est antérieur à [01,02] (temps passé)

[I',II'] coïncide avec [B'1,B'2] (aspect aoristique dans la complétive)

[I',II'] est postérieur à [I,II] (temps relatif : ultérieur)

donc la position de B2 par rapport à [01,02] reste linguistiquement indéterminée (le croit-elle encore ?) ; la

³³ Ce cas de figure est le plus fréquent, mais il y a aussi des configurations plus complexes où les deux intervalles se recouvrent partiellement

³⁴ Le même phénomène se produit avec la subordonnée de l'exemple (8) ci-dessus. La borne B2 était fixée arbitrairement dans la figure 6 (pour la commodité de lecture).

position de $[I',II']$ et donc de $[B'1,B'2]$ par rapport à $[01,02]$ l'est aussi³⁵.

Nous avons pu définir formellement des relations entre intervalles à partir d'un ensemble de relations entre bornes, dont les principales sont les suivantes :

³⁵ C'est sur la base de tels exemples qu'il nous a paru nécessaire d'abandonner la définition du temps absolu comme relation directe entre le procès et le moment de l'énonciation, ainsi que la définition du temps relatif comme relation directe entre les deux procès.

Soit i, j , deux bornes quelconques d'intervalles quelconques (éventuellement du même intervalle) ;

$i = j$ (coïncidence)

$i \infty j$ (i précède j , mais en est infiniment proche : $i = j^{-\epsilon}$)

$i \subset j$ (i précède j , mais ne se trouve pas dans son voisinage immédiat)

$i < j =_{df} (i \infty j) \vee (i \subset j)$

$i = j =_{df} (i < j) \vee (i = j)$

Ces relations entre bornes servent aussi à la définition des *types de procès* (aspect lexical). On utilise trois critères :

a) le type de bornes : soit intrinsèques ($[Bi1, Bi2]$) pour les procès téléiques (ou perfectifs), soit extrinsèques ($[Be1, Be2]$) pour les procès atéliques (ou imperfectifs)

b) les relations entre bornes : soit $B1 \infty B2$ (procès ponctuel), soit $B1 \subset B2$ (non ponctuel)

c) le fait que l'intervalle du procès subsume une série de changements, une absence de changements ou un changement atomique.

D'où la redéfinition des classes de Vendler (1967) :

état : $[Be1 \subset Be2]$ [absence de changements], ex. : *être malade, aimer la confiture*

activité : $[Be1 \subset Be2]$ [série de changements], ex. : *marcher, manger des fruits*

accomplissement : $[Bi1 \subset Bi2]$ [série de changements], ex. : *manger une pomme*

achèvement : $[Bi1 \infty Bi2]$ [changement atomique], ex. : *atteindre un sommet*.

2.3. Les principes de calcul

On admet que chacun des marqueurs aspectuo-temporels code une ou plusieurs instruction(s) pour la construction d'intervalles ou de relations entre bornes sur l'axe temporel. Ces instructions constituent la part aspectuo-temporelle de la valeur en langue (A) du marqueur. La composition sémantique réside alors dans l'assemblage de ces divers éléments de représentation (intervalles, relations entre intervalles, relations entre bornes). Le but de cet assemblage est d'abord d'obtenir une représentation globale (X) qui soit cohérente³⁶ ; or il arrive très souvent que différentes instructions codées par un même énoncé soient

³⁶ On montre dans Gosselin (96) qu'elle doit être aussi plausible et pertinente.

contradictoires (au sens où les éléments de représentation à construire sont incompatibles). Nous avons voulu montrer que ces cas de conflit sont résolus par la mise en oeuvre de modes de résolution de conflit, qui consistent à déformer – le moins possible – les représentations globales selon des procédures régulières et prédictibles, de façon à satisfaire à toutes les exigences (correspondant aux instructions codées par l'énoncé ou dépendant de principes généraux sur la bonne formation des représentations). Exemples :

a) Conflit insoluble (pour lequel il n'existe pas de mode de résolution) :

(12) * *Il marcha depuis deux heures*

Le passé simple marque l'aspect aoristique, c'est-à-dire la coïncidence de [I,II] avec [B1,B2]; [*depuis* + durée] indique au contraire que B1 précède I et mesure cet écart (comme dans '*Il marchait depuis deux heures quand je l'ai rencontré*').

b) Conflit résolu par la contraction du procès sur sa borne initiale :

(13) *Il dormit à 10h40*

dormir : activité : $Be1 \subset Be2$

passé simple : aspect aoristique : $I = B1, II = B2$

à 10h40 : circonstanciel ponctuel : $ct1 \infty ct2$

construction syntaxique du circonstanciel : intégré au SV : $ct1 = B1, B2 = ct2$

d'où $ct1 = B1 \infty B2 = ct2$, ce qui est incompatible avec l'aspect non ponctuel marqué par *dormir* ($B1 \subset B2$).

Résolution du conflit par déformation de la représentation : le procès se contracte sur sa phase initiale, ponctuelle, et *dormir* équivaut à *s'endormir*.

c) Conflit résolu par l'itération :

(14) *Il nageait pendant deux heures*

imparfait : aspect inaccompli : $B1 < I, II < B2$

pendant deux heures : $ct1 \subset ct2$ (non ponctuel)

$ct1 = B1, ct2 = B2$ (circonstanciel de durée)

Principe général sur les circonstanciels : pour qu'un procès puisse entrer en relation avec un intervalle circonstanciel, il est nécessaire que les bornes du procès soient accessibles à partir de l'intervalle de

référence³⁷ (i.e. qu'il y ait au moins coïncidence entre les deux intervalles) :

$((ct1 = B1) \& (B2 = ct2)) \emptyset ((I = B1) \& (B2 = II))$.

Un conflit se produit donc avec la valeur aspectuelle de l'imparfait, qui ne peut plus servir à présenter le procès lui-même comme inaccompli (?* *Il était en train de nager pendant deux heures*).

L'une des résolutions possibles du conflit consiste à dupliquer les intervalle de référence et de procès pour construire une série itérative (notée Bs1, Bs2) sur laquelle porte un intervalle de référence spécifique [Is, IIs]. L'aspect inaccompli, codé par l'imparfait, trouve à s'appliquer sur la série itérative dans son ensemble (d'où la possibilité d'énoncer *Il nageait pendant deux heures depuis un mois*), tandis que chacune des occurrences de procès est présentée de façon aoristique (pour satisfaire aux exigences associées au circonstanciel de durée).

C'est parce que, dans ce modèle, les conflits apparaissent très fréquents et sont résolus de façon régulière par déformation des représentations globales qu'un calcul sémantique de la PCG est possible et qu'il peut être considéré comme holiste (ainsi l'itération dans (14) n'est pas prise pour une valeur de l'imparfait, mais pour le résultat d'une résolution de conflit).

2.4. Les tests

Les représentations obtenues par le calcul sont soumises à des tests. Il s'agit généralement de *tests de compatibilité*. Par exemple, on considère habituellement que la compatibilité d'un procès avec [*en + durée*] indique son caractère télique (intrinsèquement borné), alors qu'un procès atélique (borné de façon extrinsèque) demande un complément du type [*pendant + durée*]. Mais cette conception des tests de compatibilité suppose un cadre d'interprétation atomiste, car on admet qu'un marqueur a une signification identique en présence ou en l'absence de l'élément qui sert à le tester. Autrement dit, le prédicat (qui exprime le procès) est censé avoir la même signification temporelle en l'absence du complément de durée qu'en sa présence. Or ce postulat est, comme la conception atomiste dans son ensemble, particulièrement contestable. Car il est clair, par exemple, que *dormir* a une valeur inchoative en (15) qu'il ne saurait présenter en (16) :

³⁷ Ce principe paraît cognitivement plausible : comment pourrait-on localiser ou mesurer la durée d'un procès dont les bornes excèderaient le champ de ce qui est perçu/montré ([I,II]) ?

(15) *Il a dormi en cinq minutes (♠ Il s'est endormi en cinq minutes)*

(16) *Il a dormi pendant cinq minutes (? ? Il s'est endormi pendant cinq minutes).*

Il est apparu qu'en fait très peu de combinaisons de marqueurs sont effectivement impossibles, mais que, très souvent, elles déclenchent des conflits entre instructions qui sont résolus par des procédures régulières de déformation. De sorte qu'aux tests de compatibilité (qui restent indispensables) doivent impérativement être adjoints des *tests de paraphrasabilité* destinés à identifier d'éventuelles déformations des représentations. Ainsi la possibilité d'associer dormir à [*en + durée*], comme dans l'exemple (15), n'indique nullement que dormir est intrinsèquement borné (télique), mais que sous l'effet de [*en + durée*], le procès se déforme de façon telle que seule sa phase préparatoire (l'endormissement), qui aboutit effectivement au sommeil, se trouve retenue. C'est l'association de ces deux types de tests qui permet de valider ou de réfuter les analyses et donc de corroborer ou de réfuter les hypothèses.

2.5. Le système expert

Ces principes ont été testés au moyen d'un système expert sur un très grand nombre d'énoncés, et peuvent être considérés comme (au moins provisoirement) corroborés. Ce système expert, conçu grâce au générateur de systèmes experts H-EXPERT met en oeuvre près de cinq cents règles dont la structure de base est du type : "Si tel attribut de tel objet a telle valeur, alors assigner telle valeur à tel attribut de tel objet". Ces règles sont disposées en différentes couches, qui se laissent regrouper en quatre étapes principales pour le traitement d'un énoncé. Une première étape, amenée à disparaître lorsque le système sera intégré à un analyseur plus général du français et qu'il pourra prendre pour entrées les sorties de l'analyse morpho-syntaxique, est constituée de demandes adressées à l'utilisateur concernant les caractéristiques morpho-syntaxiques de l'énoncé (la construction du verbe, le temps verbal, la présence de circonstanciels, etc.). Vient ensuite la phase de traitement proprement dit, où des éléments de représentation (des bornes, des relations entre bornes) sont calculés à partir des combinaisons de marqueurs. À côté de règles simples du type : "le passé simple implique le temps passé ($\text{II} < 01$)", on trouve aussi des règles qui aboutissent à des disjonctions (et donc à des résultats virtuellement ambigus) comme celle qui attribue à l'emploi du passé composé une double configuration aspectuo-temporelle possible (passé

aoristique ou présent accompli). En général, ces ambiguïtés virtuelles sont levées sous l'effet d'éléments du contexte qui déclenchent une procédure de désambiguïsation. Ainsi dans l'énoncé

(17) *Pierre a terminé depuis dix minutes*

le circonstanciel exclut l'aspect aoristique. Seul l'aspect accompli reste envisageable, impliquant donc la valeur temporelle de présent (d'où l'impossibilité d'énoncer "hier, il a terminé depuis dix minutes"). La troisième étape consiste à identifier les conflits et à les résoudre par l'application de règles spécifiques. Un quatrième ensemble de règles procède enfin à l'affichage des résultats.

Reprenons, en suivant très schématiquement l'ordre des étapes de traitement, l'exemple (1) que nous rappelons :

(1) (...) *Pierre dormait en cinq minutes*

1^{ère} étape : informations morpho-syntaxiques demandées à l'utilisateur et communiquées au système :

dormir : verbe de classe c (on répartit les verbes en 9 classes, en fonction de tests spécifiques³⁸)

complément du verbe : \emptyset

temps morphologique : imparfait

circonstanciel : [*en* + durée]

subordonnée : \emptyset

2^{ème} étape : instructions associées aux marqueurs :

dormir : $Be1 \subset Be2$ (procès non ponctuel, borné de façon extrinsèque)

imparfait : $B1 < I, II < B2$ (aspect inaccompli)

$II < O1$ (temps passé)

[*en* + durée] : $ct1 \subset ct2$ (circonstanciel non ponctuel)

$ct1 = Bi1, ct2 = Bi2$ (circonstanciel de durée portant sur un procès intrinsèquement borné)

À quoi il faut ajouter le principe général sur les circonstanciels (voir ci-dessus, exemple (14)) : $((ct1 = B1) \& (B2 = ct2)) \emptyset ((I = B1) \& (B2 = II))$.

3^{ème} étape : Deux conflits sont identifiés et résolus :

a) Un conflit entre l'aspect inaccompli, marqué par l'imparfait, et la présence d'un circonstanciel de durée est résolu par l'itération (comme dans l'exemple (14) ci-dessus).

b) Un conflit entre le procès borné de façon extrinsèque et le circonstanciel qui exige un bornage intrinsèque du procès conduit à déplacer le procès vers sa phase préparatoire (comme dans l'exemple (15))³⁹.

4^{ème} étape : affichage des résultats : on donne l'ensemble des relations entre bornes, en précisant – le cas échéant – de quels conflits elles proviennent.

3. DEMARCHE HYPOTHETICO-DEDUCTIVE ET UTILISATION DE L'ORDINATEUR

³⁸ Cf. Gosselin (1996), p. 186 sq.

³⁹ D'où la paraphrase : "*Pierre mettait régulièrement cinq minutes à s'endormir*".

Les critiques épistémologiques les plus vigoureuses et les mieux étayées qui ont été adressées à la théorie popperienne consistent à opposer cette *logique* de la découverte à la *pratique* scientifique réelle⁴⁰. Deux types d'arguments peuvent être distingués :

1) Les théories ne peuvent être véritablement réfutables, pour la simple raison qu'elles ne sont jamais, à elles seules, prédictives. Elles nécessitent des *propositions auxiliaires*, qui ne sont pas testées pour elles-mêmes, mais qui rendent possible la déduction de prédictions. Si bien que la réfutation d'une prédiction n'implique pas nécessairement celle d'une hypothèse théorique, car elle résulte bien souvent de propositions auxiliaires erronées ou simplement inadéquates. Comme ces propositions auxiliaires, qui n'appartiennent pas à la théorie elle-même, restent généralement implicites, il est très fréquemment impossible de dire si la théorie est ou non réfutée par une prédiction qui s'avère fausse. Aussi les théories scientifiques ne sont-elles guère plus réfutables que vérifiables.

2) Les propositions générales qui ont statut d'hypothèses "ne viennent pas du ciel" (Putnam), mais résultent généralement de procédures d'induction ou d'abduction fondées sur des observations empiriques. Il n'y a dès lors pas lieu d'opposer radicalement les deux démarches (inductive ou abductive *versus* hypothético-déductive), il ne s'agit en réalité que des deux moments de l'ordinaire va-et-vient entre observation et théorisation constitutif de la pratique scientifique réelle.

Nous n'entrerons pas ici dans le débat qui oppose les approches normatives *versus* historicistes de l'épistémologie (rappelons simplement que la logique de la découverte vise à régir la pratique et non à la décrire ou à l'expliquer – tâches dévolues à l'histoire, à la sociologie ou à la psychologie des sciences). Le but que nous nous sommes fixé a été d'essayer – pour les raisons spécifiquement linguistiques qui ont été exposées – de conformer notre pratique à la logique normative popperienne, et, dans cette perspective, l'ordinateur s'est avéré un outil indispensable dans la mesure où son utilisation permet d'éviter les principaux écueils qui guettent la démarche hypothético-déductive :

1) La formalisation requise par l'automatisation impose que *toutes* les hypothèses théoriques et propositions auxiliaires (également exprimées sous forme de règles) soient totalement explicites.

⁴⁰ Cf. P. A. Schlipp (éd.) (1974), et en particulier dans cet ouvrage, l'article de H. Putnam.

2) La possibilité de visualiser la "trace" des déductions opérées, et donc des règles effectivement mises en oeuvre lors d'une procédure d'analyse permet au chercheur de localiser la ou les hypothèse(s) responsable(s) d'une prédiction fautive. La procédure de réfutation devient donc sélective.

3) La quasi-impossibilité de distinguer clairement entre les démarches inductive ou abductive d'une part, et hypothético-déductive, de l'autre, ne se manifeste que lorsque le cheminement qui conduit des propositions générales aux particulières (les prédictions) est si simple qu'il peut être parcouru par l'intuition. L'ordinateur permet, de façon très rapide et totalement fiable, des séries de déductions extrêmement complexes mettant en oeuvre un très grand nombre de règles ; et dans ce cas, très souvent requis par l'analyse sémantique, le doute n'est plus possible : l'intuition (comme guide de l'induction et de l'abduction) s'avère impuissante.

Quant à la question du choix des hypothèses générales, soumises à l'expérimentation, il s'agit avant tout, pour le chercheur, d'une affaire privée, et en grande partie inconsciente. Mais que ces hypothèses proviennent de tâtonnements plus ou moins abductifs, d'emprunts à d'autres auteurs, d'analogies avec des lois adoptées dans d'autres domaines, ou même, comme il arrive parfois, d'erreurs de raisonnement, elles ne prennent d'intérêt qu'à partir du moment où elles sont testées par le système déductif et prédictif.

Reste bien sûr qu'une théorie hypothético-déductive, même largement corroborée (i.e. confrontée avec succès à un très grand nombre d'observations), ne saurait en aucun cas être considérée comme vérifiée, et qu'il est toujours possible que deux théories fondées sur des hypothèses générales différentes, voire contradictoires, donnent lieu à des prédictions identiques.

4. CONCLUSION

La modélisation proposée de la PCG est holiste dans la mesure où les mécanismes de composition sont ici des procédures complexes qui font intervenir de façon centrale les concepts de conflit et de résolution de conflit par déformation des représentations globales. C'est parce que le cheminement des valeurs en langue (*A*) – décrites sous forme d'instructions pour la construction d'éléments de représentation – aux représentations sémantiques globales (*X*) est souvent complexe et qu'il fait intervenir des principes généraux très abstraits – comme le principe général sur les compléments circonstanciels – que la démarche inductive fondée sur l'intuition nous paraît devoir céder la place à une

perspective hypothético-déductive. Dès lors, le recours à l'informatique, loin d'être considéré comme une simple application de la théorie linguistique, apparaît, à l'inverse, comme un outil nécessaire à sa constitution même. Que l'on considère l'extrême simplicité (apparente) des exemples proposés dans cet article, et l'on concevra combien il est nécessaire d'utiliser l'ordinateur pour opérer l'ensemble des déductions déclenchées par le traitement d'énoncés complexes (c'est ainsi que nous avons pu proposer un système de règles pour rendre compte de la "concordance des temps") et *a fortiori* pour traiter des relations chronologiques entre procès dans le texte.

Bibliographie

- Asher, N. et Bras, M. (1992) "The temporal Structure of French Texts within Segmented Discourse Representation Theory", dans Aurnague, M., Borillo, A., Borillo, M., et Bras, M. (éds) *Semantics of Time, Space, Movement and Spatio-temporal Reasoning*, Working Papers of the 4th International Workshop, Univ. Toulouse-Le Mirail, pp. 203-217.
- Bennett, W. S., Herlick, T., Hoyt, L., Liro, J. et Santisteban, A. (1989-90) "Toward a Computational Model of Aspect and Verb Semantics", *Machine Translation 4.4*, Dordrecht, Kluwer Academic Publishers.
- Carlson, L. (1981) "Aspect and Quantification", *Syntax and Semantics 14*, pp. 31-64.
- Dowty, D. (1977) "Toward a Semantic Analysis of Verb Aspect and the English "Imperfective" Progressive", *Linguistics and Philosophy 1*, pp. 45-77.
- Dowty, D. (1986) "The Effects of Aspectual Class on the Temporal Structure of Discourse : Semantics or Pragmatics", *Linguistics and Philosophy 9.1*, pp. 37-62.
- Ducrot, O. (1984) *Le dire et le dit*, Paris, Minuit.
- Ducrot, O. et Schaeffer, J.-M. (1995) *Nouveau dictionnaire encyclopédique des sciences du langage*, Paris, Seuil.
- Eberle, K. (1988) "Partial Orderings and Aktionsarten in Discourse Representation Theory", *COLING 88*, Budapest, pp. 160-165.
- Engel, P. (1994) *Davidson et la philosophie du langage*, Paris, PUF.
- François, J. (1981) "Travaux récents d'inspiration formelle sur l'aspect et l'aktionsart", *Linguisticae Investigationes V.2*, pp. 275-301.
- Gosselin, L. (1996) *Sémantique de la temporalité en français ; un modèle calculatoire et cognitif*, Louvain-la-Neuve, Duculot.
- Hinrichs, E. W. (1986) "Temporal Anaphora in Discourses of English", *Linguistics and Philosophy 9*, pp. 63-82.
- Hintikka, J. (1994) *Fondements d'une théorie du langage*, trad. N. Lavand, Paris, PUF.

- Kamp, H. et Rohrer, C. (1983) "Tense in Texts", dans Bauerle, R., Schwarze, Ch. et von Stechow, A. (éds) *Meaning, Use and Interpretation of Language*, Berlin, De Gruyter, pp. 250-269.
- Kleiber, G. (1990) *La sémantique du prototype*, Paris, PUF.
- Kleiber, G. (1994) *Anaphores et pronoms*, Louvain-la-Neuve, Duculot.
- Klein, W. (1994) *Time in Language*, Londres, Routledge.
- Lakoff, G. (1987) *Women, Fire and Dangerous Things*, Chicago U.P.
- Lascarides, A. et Asher, N. (1991) *Temporal Interpretation, Discourse Relations and Common Sense Entailment*, ms., H.C.R.C., Univ. d'Edimbourg.
- Merleau-Ponty, M. (1969) *La prose du monde*, Paris, Gallimard.
- Milner, J.-Cl. (1989) *Introduction à une science du langage*, Paris, Seuil.
- Moens, M. et Steedman, M. (1988) "Temporal Ontology and Temporal Reference", *Computational Linguistics* 14. 2, pp. 15-27.
- Molendijk, A. (1990) *Le passé simple et l'imparfait : une approche reichenbachienne*, Amsterdam, Rodopi.
- Mourelatos, A. P. D. (1978) "Events, Processes, and States", *Linguistics and Philosophy* 2, pp. 415-434.
- Partee, B. H. (1984) "Nominal and Temporal Anaphora", *Linguistics and Philosophy* 7, pp. 243-286.
- Picoche, J. (1986) *Structures sémantiques du lexique français*, Paris, Nathan.
- Popper, K. (1973) *La logique de la découverte scientifique*, trad. N. Thyssen-Rutten et Ph. Devaux, Paris, Payot.
- Popper, K. (1985) *Conjectures et réfutations*, trad. M.-I. et M. B. de Launay, Paris, Payot.
- Putnam, H. (1974) "The 'Corroboration' of Theories", dans P.A. Schlipp (éd.) *The Philosophy of Karl Popper*, La Salle, Illinois, Open Court, pp. 221-240.
- Rastier, Fr., Cavazza, M. et Abeillé, A. (1994) *Sémantique pour l'analyse*, Paris, Masson.
- Reichenbach, H. G. (1980) *Elements of symbolic Logic*, New York, Dover Publications.
- Swinburne, R. (éd.) (1974) *The Justification of Induction*, Oxford U. P.
- Taylor, B. (1977) "Tense and Continuity", *Linguistics and Philosophy* 1, pp. 199-220.
- Vandeloise, Cl. (1991) "Autonomie du langage et cognition", dans *Communications 53 : Sémantique cognitive*, pp. 69-102.
- Vendler, Z. (1967) *Linguistics in Philosophy*, Cornell U. P.
- Vet, C. (1991) "The Temporal Structure of Discourse: Setting, Change and Perspective", dans Fleishman, S. et Waugh, R. L. (éds) *Discourse Pragmatics and the Verb. The Evidence from Romance*, New York, Routledge, pp. 7-25.

- Victorri, B. (1996) "Modéliser les interactions entre une expression polysémique et son co-texte", dans les *Actes du colloque "Co-texte et calcul du sens"*, P. U. de Caen.
- Winters, M. (1991) "Subjonctif et réseau", *Communications* 53, pp. 155-170.