

SÉM
iotiques

DÉC.
1994
n°6-7



LINGUISTIQUE COGNITIVE et MODÈLES DYNAMIQUES
n° dirigé par Jean Petitot



DIDIER - ERUDITION

Linguistique cognitive et Modèles dynamiques



Avant-propos	5
Yves Marie VISETTI	15
Les Modèles connexionnistes entre perception et sémantique	
Tim van GELDER	49
Compositionality and Dynamics in Neural Network Representations	
Ronald W. LANGACKER	69
Structural Syntax : The View from Cognitive Grammar	
Franson MANJALI	85
Dynamic Semiotics or the Case for Actantial Case	
Peter GÄRDENFORS	99
Frameworks for Properties : Possible Worlds vs. Conceptual Spaces	
Per Aage BRANDT	121
Pour une phrastique intégrale	
Pierre OUELLET	137
La Sémiotique cognitive : les sciences de l'esprit entre la nature et la culture	
Wolfgang WILDGEN	161
Les Fondements sémiotiques et philosophiques d'une grammaire morphodynamique	
David PIOTROWSKI	175
Approche morphodynamique de deux régimes fondamentaux du langage	
Jean PETITOT	187
Dynamical Constituency : An Epistemological Analysis	
Résumés	227
Abstracts	231

RÉSUMÉS

*Linguistique cognitive
et Modèles dynamiques***Per Aage BRANDT***Pour une phrastique intégrale*

121-136

L'article propose une nouvelle conception de la phrase, selon laquelle l'énonciation (pragmatique), l'articulation propositionnelle, l'organisation casuelle et finalement la rection verbale et adverbiale, dans cet ordre, *précèdent* la stemmatique (la syntaxe en arbre) qui détermine la linéarisation grammaticale. Cette interprétation intégrale de la phrase est rendue possible par une topologie catastrophiste, le *papillon*.

Peter GÄRDENFORS*Cadres pour l'étude des propriétés : mondes possibles
vs espaces conceptuels*

99-120

L'article montre que le type traditionnel de théorie des modèles utilisé en sémantique intensionnelle ne fournit pas un cadre approprié à un tel propos. La discussion est centrée sur la notion de propriété. On présente d'abord la définition intensionnelle standard d'une propriété en termes de mondes possibles. Certains des problèmes philosophiques qu'elle pose sont soulignés. On introduit alors l'alternative des espaces conceptuels comme cadre approprié pour la sémantique. Les espaces conceptuels consistent en un certain nombre de "dimensions qualitatives" qui sont souvent dérivées de mécanismes perceptuels. On montre comment de tels espaces peuvent servir de base à un nouveau critère de définition d'une propriété, critère échappant aux problèmes de l'approche traditionnelle. On montre qui plus est que ce dernier est utile pour comprendre les effets de prototype des propriétés et le rôle des propriétés dans le raisonnement non monotone.

Ronald W. LANGACKER*La Syntaxe structurale vue par la grammaire cognitive*

69-84

Cet article propose une comparaison entre la syntaxe structurale de Tesnière et la théorie actuelle de la grammaire cognitive. Les notions de base de cette dernière sont présentées et appliquées à l'analyse d'exemples significatifs. Il apparaît que les affinités entre les deux perspectives sont profondes, bien qu'il subsiste des différences importantes. En particulier, la notion fondamentale chez Tesnière de "translation" ne semble pas pertinente et la conception des "actants" est trop

simplifiée. Cependant, l'œuvre de Tesnière demeure remarquable et peut être reconnue comme une entreprise intellectuelle comparable à celle de la grammaire cognitive.

Franson MANJALI

Sémiotique dynamique : les actants et les cas

85-97

Depuis Saussure, la tendance de considérer les plans de l'expression et du contenu du langage d'un point de vue discret a oblitéré le caractère essentiellement dynamique et continu du signifié des énoncés. Avec les notions *d'actant* et de *connexion*, Tesnière a introduit la perspective d'une continuité dynamique dans la sémantique des cas. Les 'morphologies archétypes' de René Thom ont fourni une compréhension topologique de la manière dont une liste finie de cas peut servir de système classificatoire au dynamisme infini du monde physique. L'article décrit également la théorie des *karaka* telle que l'a élaborée le philosophe indien du langage Bharthari (au 7^e siècle avant J. -C.) ; pour celle-ci, la définition d'un nombre limité de catégories casuelles est un moyen d'appréhender un dynamisme actantiel potentiellement infini.

Pierre OUELLET

*La Sémiotique cognitive : les sciences de l'esprit
entre la nature et la culture*

137-159

La sémiotique cognitive permet de faire le lien entre la perception sensorielle et la perception sémantique (ou discursive) *via* les processus d'imagerie mentale, dont la structure et la dynamique sont isomorphes à l'expérience sensori-motrice et dont le rôle, dans le langage, est de donner un contenu intuitif aux représentations conceptuelles. Deux processus, l'exploration et la rotation mentales, caractéristiques de l'imagerie, conditionnent l'énonciation descriptive, notamment celle qui a pour objet le monde naturel. Un texte de Claude Dourguin nous sert d'exemple pour montrer comment ces processus jouent dans la perception discursive et fondent la catégorie ontologique de "chemin" (Jackendoff), qui est la figure du "trajet" qu'épouse le double mouvement du sujet et de l'objet de la perception dans le "champ de vision" ou le "champ de présence" représenté. L'étude de ces représentations littéraires de l'activité sensori-motrice (réelle ou imaginaire) permet de dégager des "esthésies" ou des configurations de la connaissance sensible, qui montrent comment l'histoire propre aux systèmes symboliques (comme la littérature) modulent et modalisent notre expérience perceptive.

Jean PETITOT

La Constituance dynamique : une analyse épistémologique 187-225

Le développement des modèles dynamiques en sciences cognitives soulève des problèmes épistémologiques majeurs. L'un des plus délicat concerne la possibilité d'édifier une théorie dynamique-connexionniste de la constituance des représentations mentales. Il a fait l'objet il y a quelques années d'un vif débat ayant opposé Jerry Fodor et Zenon Pylyshyn à Paul Smolensky. Or il existe une solution dynamique au problème de la constituance, fondée sur les modèles

morphodynamiques de René Thom. L'article explique comment elle permet de reprendre le débat.

David PIOTROWSKI

*Approche morphodynamique de deux régimes
fondamentaux du langage*

175-186

Parmi les principaux régimes de structure qui fondent l'analyse linguistique et, corrélativement, déterminent l'objectivité en langue, figurent la dualité du signe et le *différentiel* d'acceptabilité (couple *possible/impossible*). Après un rappel sommaire de ces deux régimes et des difficultés qu'ils soulèvent, nous en proposons une schématisation dans le cadre de la théorie morphodynamique. Le différentiel d'acceptabilité est analysé comme un principe de production de grandeurs oppositives, et l'intrication entre le signifiant et le signifié apparaît alors comme une conséquence de ces modalités de construction.

Tim van GELDER

*Compositionnalité et dynamique des représentations
implémentées dans des réseaux neuronaux*

49-67

Est-il possible de comprendre, à partir d'une perspective entièrement dynamique, la façon dont les réseaux connexionnistes peuvent intégrer les représentations compositionnelles ? Cet article explore le concept même de *compositionnalité* en analysant six de ses aspects et en révélant ainsi un espace étonnamment riche de formes possibles pour les représentations compositionnelles. Une de ces formes, la compositionnalité dynamique, peut être implantée dans des systèmes dynamiques au moyen de "chaînages d'attracteurs" dans lesquels la trajectoire d'un système dans son espace d'états est déterminée par un paysage d'attracteurs qui se modifie lorsque les paramètres de contrôle changent. Des versions de cette idée existent dans l'œuvre de René Thom, en particulier dans les développements que lui a donnés Jean Petitot, et, chez Robert Port, dans ses travaux connexionnistes consacrés à la reconnaissance de *patterns* temporels.

Yves Marie VISETTI

Les Modèles connexionnistes entre perception et sémantique 15-48

L'essor de la modélisation connexionniste mobilise, au moins à titre instrumental, une part de la conceptualité mathématique ayant trait au Continu et à l'espace ; les modèles qui en résultent favorisent le développement des conceptions topologiques et dynamiques dans les sciences cognitives ou les sciences du langage. Qu'il s'agisse ainsi de la nature des représentations ou de la compréhension des processus, la place du continu et du spatial ne cesse de croître, au sein d'une grande diversité d'acceptions et de méthodes. L'article analyse plusieurs travaux marquants dans cette perspective : modélisation psycholinguistique dans la ligne du groupe connexionniste PDP, grammaires harmoniques et modèles de polysémie, modèles d'ancrage perceptif du symbolique, problématique du courant de la Vie Artificielle. Bien que faisant appel à des techniques fortement apparentées, ces travaux n'en sont pas moins les indices d'*attitudes représentationnelles* divergentes, dont on s'efforcera de distinguer les formes, les valeurs, et les méthodes caractéristiques.

Wolfgang WILDGEN

*Les Fondements sémiotiques et philosophiques
d'une grammaire morphodynamique*

161-174

Le paradigme morphodynamique initié par René Thom permet la construction d'une sémiotique et d'une grammaire morphodynamiques. Les fondements de ces constructions sont précisés par la formulation d'une série de postulats : le postulat réaliste, le postulat de Peirce, le postulat génétique, le postulat de Thom, le postulat de Tesnière-Thom. Pour finir, une version faible du lien entre ontologie, représentation cognitive et structure linguistique est formulée.

ABSTRACTS

Cognitive Linguistics and Dynamical Models

Per Aage BRANDT

Towards an Integrated Sentence Analysis

121-136

This article proposes a new conception of the sentence, according to which (pragmatic) enunciation, propositional articulation, case structure, and finally verbal and adverbial binding *precede* — in this order — the “stemmatic” level (that of syntactic trees) determining grammatical linearization. This integrated interpretation of sentence forms is made possible by a catastrophe topology, the so-called *butterfly catastrophe*.

Peter GÄRDENFORS

*Frameworks for Properties : Possible Worlds
vs. Conceptual Spaces*

99-120

It is argued that the traditional kind of model theory used in intensional semantics is not the right kind of framework for its purpose. The discussion in this paper is focussed on the notion of property. First, the standard intensional definition of a property, which is formulated in terms of possible worlds, is presented. Some of the philosophical problems this definition leads to are outlined. As a remedy, conceptual spaces are suggested as an appropriate framework for semantics. Conceptual spaces consist of a number of ‘quality dimensions’ which are often derived from perceptual mechanisms. It is demonstrated how such spaces can be used as a basis for a new criterion of what a property is. This criterion is shown to elude the problems of the traditional approach. Furthermore, it is argued that the criterion is useful for understanding prototype effects of properties and the role of properties in non-monotonic reasoning.

Ronald W. LANGACKER

Structural Syntax : The View from Cognitive Grammar

69-84

This paper offers a comparison between Tesnière’s structural syntax and the modern theory of cognitive grammar. Basic notions of the latter are introduced and applied to the analysis of representative examples. The affinities between the two frameworks are found to be extensive, yet there are also important points of difference. In particular, Tesnière’s basic notion of “translation” is found

AVANT-PROPOS

1. Le tournant cognitif et dynamique post-structural des disciplines sémio-linguistiques

“If connectionism was the most dramatic theoretical revolution of the 1980s, it appears that dynamics is the connectionism of the 1990s”. Cette affirmation ici même de Tim van Gelder exprime bien la spectaculaire montée en puissance des approches dynamiques du langage qui s’est opérée ces dernières années.

Le tournant dynamique des disciplines sémio-linguistiques est inséparable de leur tournant cognitif, qui les a ouvertes à nouveau sur des problématiques plus “mentalistes” comme la phénoménologie ou la psychologie cognitive. Le retour amont des descriptions eidétiques et formelles de compétences vers des explications naturalistes de performances concrètes a conduit à une remise en perspective complète et à un profond renouvellement des traditions structuralistes.

La naturalisation des descriptions formelles a évidemment projeté au premier plan la question de leur implémentation dans des substrats physiques et biologiques. Pendant les années 1980, les modèles connexionnistes neuromimétiques ont considérablement fait progresser la compréhension de phénomènes cognitifs de base comme la catégorisation, l’apprentissage ou l’inférence inductive. Le transfert de modèles sophistiqués venant de la physique statistique a conduit à les formuler dans un univers mathématique où, tout naturellement, le point de vue dynamique s’est imposé. Comme Daniel Amit l’a montré, l’introduction d’une hypothèse de rétroaction complète (les réseaux en couches étant essentiellement “feed-forward” (pro-actifs) et donc peu rétroactifs) permet de réinterpréter les phénomènes bien connus de “réverbération” à la Hebb comme une stabilisation de la dynamique de l’assemblée neuronale considérée dans l’un de ses attracteurs sur une durée d’ordre “psychologique” (de l’ordre de quelques centaines de ms, l’échelle des activités neuronales étant, elle, de l’ordre de quelques ms)¹.

L’implémentation connexionniste des structures et processus cognitifs a ainsi convergé avec un point de vue dynamique. Cela a permis tout naturellement de faire le lien avec les modèles morphodynamiques précurseurs qui avaient été développés en Europe pour le structuralisme, à la suite des propositions de René Thom, en particulier par Per Aage Brandt, Wolfgang Wildgen et nous-même. Ce structuralisme dynamique

¹Pour une discussion de ces modèles dynamiques neuromimétiques, cf. le prochain numéro de “The Behavioral and Brain Sciences” où un “target article” de D. Amit est commenté par de nombreux spécialistes.

avait déjà conduit en sémiotique à des remises en perspective notables du structuralisme “statique” hjelmslevo-greimassien. Il avait fait l’objet de discussions approfondies de la part de certains des meilleurs spécialistes. Pensons par exemple à Jean-Claude Coquet dans le cadre de l’Ecole sémiotique de Paris, à Pierre Ouellet à Montréal ou à Umberto Eco et Patrizia Violi à Bologne et à San Marino. Ces dernières années, il a fusionné avec le cognitivisme dynamique issu des nouvelles orientations (anti-formalistes) de la linguistique américaine (Ron Langacker, Len Talmy, Ray Jackendoff, Georges Lakoff, etc.).

Il est particulièrement satisfaisant, tout autant sur le plan scientifique que sur celui, plus sociologique, des traditions intellectuelles, de voir ainsi s’unifier, à propos d’un programme de recherche fondamental et lourd d’avenir, d’une part les sciences naturelles (physico-mathématiques, biologiques et informatiques) et les sciences humaines authentiques ainsi que, d’autre part, les traditions européennes “continentales” et les traditions anglo-saxonnes. La naturalisation de l’esprit et du sens ouvre l’horizon de l’unification méthodologique des programmes de recherche concernant les *Geisteswissenschaften*. Il s’agit d’un *factum rationis* de première grandeur : sans emphase et sans risque d’être démenti, on peut affirmer que les sciences humaines sont en train de devenir l’un des fers-de-lance des technosciences “dures”. Il y a là une convergence remarquable — un tournant cognitif et dynamique — dont ce numéro spécial de *Sémiotiques* se voudrait être le reflet.

2. Quelques points caractéristiques

Avant de présenter brièvement les contributions à ce dossier, rappelons quelques caractéristiques essentielles de ce tournant cognitif et dynamique qui représente l’alternative post-structurale majeure aux eidétiques descriptives formalistes.

1. La critique du formalisme, du générativisme et des axiomes de “clôture” affirmant l’autonomie du langage et de ses structures formelles. C’est de façon profonde que les langues naturelles sont “naturelles”. Leur “naturalité” est incompatible avec les dogmes formalistes et fonctionnalistes sur la centralité de la syntaxe. En effet ceux-ci reposent tous d’une façon ou d’une autre sur la thèse biologiquement falsifiable que l’esprit naturalisé implémente, comme un ordinateur, des programmes écrits dans un langage formel (cf. le “mentalais” de Fodor). Bref, les sciences des langues naturelles sont des sciences naturelles et non pas des sciences formelles.

2. L’insistance corrélatrice sur le primat du sémantique, sur l’inséparabilité du sens et de la grammaire. Comme l’affirme ici même Ron Langacker : “A pivotal theoretical issue is the relation between

meaning and grammar. (...) The central claim of cognitive grammar [is] that meaning and grammar are indissociable". On connaît aussi les admirables travaux de Len Talmy sur la façon dont la grammaire spécifie des contenus sémantiques.

3. L'interprétation cognitive (donc mentale et représentationnelle) du sémantique en termes de "structure conceptuelle" (cf. par exemple les travaux de Ray Jackendoff). Cela ne ramène pas pour autant les eidétiques structurales à un "psychologisme" (serait-il celui de la psychologie cognitive). Cela signifie que le sens sémio-linguistique renvoie à des contenus de *pensée*, contenus qui, eux-mêmes, peuvent (ou non) référer au monde extérieur. La conséquence en est la critique radicale des points de vue logicistes sur la sémantique qui auront dominé ce siècle (les conceptions purement dénotationnelles et vériconditionnelles, seraient-elles intensionnelles, du sens).

4. L'ouverture de la structure conceptuelle sur le monde et sur le corps. L'esprit est "incarné" (*embodied*). Les structures sémio-linguistiques et leurs universaux sont fondamentalement contraints par les structures qualitatives du monde et par la compatibilité entre le langage, la perception et l'action. D'où un renouveau spectaculaire des problématiques *phénoménologiques* (celles du second Husserl et de Merleau-Ponty).

5. En particulier, l'importance des liens entre la perception et le langage conduit à la thèse de *l'ancrage* de celui-ci dans celle-là. D'où la problématique de *l'iconicité*. L'iconicité des structures, en particulier celle des structures syntaxiques, ne signifie évidemment pas que les structures sont des figures concrètes. Il ne s'agit pas d'une figurativité au sens classique, mais d'une *iconicité abstraite*, de nature *schématique*. Les représentations mentales sont conçues de façon schématique, comme des Gestalts généralisées, comme une *imagerie mentale* qui, ainsi que Kant l'avait déjà profondément pensé à propos du schématisme des concepts empiriques, est un système de règles pour la construction de référents. C'est dire que les images-schémas structurant les représentations concernent des *types* (et non pas des occurrences, des *tokens*).

La conception gestaltiste des structures du langage s'est véritablement imposée ces dernières années puisque, tout récemment, Herbert Simon lui-même, dans un "target article" d'un numéro spécial de la *Stanford Humanities Review* (Suppl. vol. 4, n°1, 1994) a pu défendre la thèse que les contenus sont visualisés comme des images mentales (au sens de Sheppard et Kosslyn) et même, plus précisément, que : "a mental picture formed by retrieving some information from memory or by visualizing the meaning of a spoken or written paragraph is stored in the same brain tissue and acted on by the same mental processes as the picture recorded by the eyes".

6. L'iconicité schématique des structures est en fait une thèse sur le *formatage* des représentations mentales. Elle met en cause le dogme *propositionnaliste* (qui est une pierre angulaire de toutes les conceptions logicistes) que les représentations ont *a priori* un format propositionnel. La thèse est que le format est *topologico-dynamique*. On pourrait la faire remonter à Peirce et à sa théorie des graphes existentiels. L'idée est que *l'a priori spatio-temporel* est plus profond que *l'a priori* symbolique (il s'agit là aussi d'une thèse phylogénétique sur l'homínisation : le système perceptif visuel est hérité d'une très longue évolution et est une composante fondamentale du psychisme animal alors que l'idéographie et l'écriture sont des épiphénomènes culturels extrêmement récents). Philosophiquement, on peut dire qu'il s'agit de l'actualisation du problème des intuitions pures de l'Esthétique transcendantale : celle-ci est spatio-temporelle (comme chez Kant) et non pas symbolique (comme chez Hilbert et tous les formalistes).

7. Les espaces et les dynamiques qui sous-tendent le schématisme sémio-linguistique sont par conséquent *abstraites*. Ils ne se réduisent en rien à l'espace extérieur. La schématicité des contenus mentaux *n'est pas* une thèse réductionniste sur leur réductibilité à des états de chose externes (on retrouverait alors les conceptions dénotationnelles du sens). Elle concerne leur formatage.

8. La schématicité fonde les structures sémio-linguistiques dans une activité cognitive de base, celle de la *catégorisation* au moyen de *prototypes*. Dans les grammaires cognitives, même les structures syntaxiques les plus abstraites sont conçues comme typage de prototypes (catégorisation d'événements par exemple).

Ces points de vue, centraux en sémiotique morphodynamique depuis Thom, sont bien devenus ceux de la linguistique et de la sémantique cognitives contemporaines comme l'atteste une analyse récente de Peter Gärdenfors² :

(i) le sens se définit comme une conceptualisation dans un modèle cognitif (et non pas comme des conditions de vérité dans des mondes possibles) ;

(ii) les modèles cognitifs sont essentiellement déterminés perceptivement : "A central hypothesis of cognitive semantics is that the way we store perceptions in our memories has the *same form* as the meanings of words" ;

(iii) "semantic elements are based on *spatial* or *topological* objects (not symbols)" ;

(iv) "cognitive models are primarily *image-schematic* (not propositional)" ;

(v) la sémantique est première par rapport à la syntaxe, celle-ci n'est pas un calcul formel ;

(vi) "concepts show *prototype* effects."

²Cf. "Conceptual Spaces as a Basis for Cognitive Semantics" (à paraître).

3. La question de la constituance dans les modèles dynamiques

Les modèles dynamiques (implémentables dans des réseaux connexionnistes) utilisés pour modéliser les schèmes cognitifs identifient un contenu à un attracteur. Une question centrale est alors celle de l'interprétation, dans ce contexte modélisateur, de la *constituance* et de la *compositionnalité* des structures, c'est-à-dire celle de la modélisation dynamique des structures en constituants. Ce problème est en quelque sorte "la revanche" du point de vue symbolique formel classique (où constituance et compositionnalité sont des *a priori* constitutifs et ne font donc pas question). La difficulté est double. D'abord, il faut comprendre ce que peut être une "syntaxe" d'attracteurs. Ensuite, il faut tenir compte du fait que, dans les modèles dynamiques, les schèmes sont des *types* "incarnés" (*embodied*) et que les relations syntaxiques entre types *émergent* des *interactions* dynamiques entre leurs *tokens*. Il faut donc élaborer une théorie des *interactions entre attracteurs* et montrer qu'il existe des *interactions archétypes*. Comme René Thom l'a montré dès la fin des années 1960, la théorie des *bifurcations* permet de développer une telle théorie des interactions. Celle-ci sert de base aux modèles morphodynamiques de la syntaxe dans les théories *actantielles* — qu'elles soient casuelles (Hjelmslev, Fillmore, Anderson), grammaticales (Tesnière) ou narratives (Greimas).

4. Présentation du dossier

Pour des raisons de cohérence interne du dossier, nous avons adopté un ordre de présentation allant d'articles largement synthétiques faisant le bilan des approches dynamiques à des articles plus spécifiques proposant soit des résultats particuliers, soit des discussions de points techniques, soit des suggestions pour la recherche future.

Dans un premier article, Yves Marie Visetti présente une large synthèse du rôle croissant des modèles topologiques et dynamiques en sciences cognitives et esquisse les différentes "attitudes représentationnelles" qui structurent ce champ.

Après avoir rappelé les modèles connexionnistes classiques de type PDP ("Parallel Distributed Processing", l'ouvrage de référence édité par D. Rumelhart et J. McClelland), il présente les travaux de J. Elman qui ont montré que, pour apprendre (de façon supervisée par des jugements d'acceptabilité) les régularités syntaxiques d'un corpus de phrases, un réseau connexionniste doit construire sur ses couches cachées (dans son espace d'états internes) des "*paradigmes*", c'est-à-dire regrouper les représentations de façon que leurs *rapports fonctionnels* se trouvent identifiés à des *rapports de position*. Ces résultats sont fondamentaux car ils valident la thèse structuraliste de la projection de l'axe paradigmatique sur l'axe syntagmatique. Visetti analyse ensuite les modèles de

³Signalons à ce propos la parution des actes du remarquable Colloque «Le Continu en sémantique linguistique» organisé en 1992 par B. Victorri et C. Fuchs («Continuity in Linguistic Semantics», John Benjamins, 1994).

P. Smolensky et de B. Victorri concernant l'organisation topologique et dynamique des espaces de traits de l'analyse componentielle ainsi que la polysémie³. Il explique comment s'élaborent ainsi des modèles de schèmes. Il en vient ensuite aux travaux concernant la modélisation des images-schèmes des grammaires cognitives, qui fonctionnent comme des "intermédiaires" entre la diversité des formes perceptives et les catégorisations linguistiques. Un tel schématisme permet d'actualiser une conception pictorielle-déictive du langage (isologie entre les structures qualitatives du monde et les structures linguistiques qui les symbolisent) en échappant aux cercles vicieux que l'on trouve habituellement dans ce genre de point de vue (comme dans le *Tractatus* de Wittgenstein). Visetti conclut alors en analysant les approches *non représentationnalistes* de la genèse adaptative des structures et de l'adaptation organisme-environnement qui se développent rapidement en Vie artificielle (travaux de G. Edelman, F. Varela, P. Bourguin).

Tim van Gelder analyse quant à lui le problème de la constituance et de la compositionnalité. Ayant organisé avec son collègue Bob Port de l'Université d'Indiana un important Colloque ayant abouti à un ouvrage de référence sur l'usage des modèles dynamiques en sciences cognitives (*Mind as Motion*, MIT Press, à paraître, 1995), il centre sa réflexion sur cette question : "the problem, in short, is to reconcile compositionality, dynamics and neural networks". La difficulté est que le concept même de syntaxe ne fait pas partie de la théorie standard des systèmes dynamiques. De même que les concepts de catégorisation et d'apprentissage ne faisaient, évidemment, pas partie de la physique statistique des verres de spins mais sont devenus des concepts physico-mathématiques à partir du transfert de ces modèles aux sciences cognitives, de même l'usage des modèles dynamiques en sémio-linguistique rétroagit sur la théorie des systèmes dynamiques en posant le problème d'une "syntaxe" d'attracteurs. Van Gelder analyse alors six "key issues" permettant de comprendre la différence entre les systèmes symboliques statiques-discrets et les systèmes dynamiques où la constituance et la compositionnalité émergent de processus d'interaction. Il montre que les conceptions connexionnistes classiques ne sont pas suffisantes et qu'il faut penser la syntaxe à partir de dynamiques "externes" qui contrôlent les attracteurs de dynamiques internes. Il expose pour conclure la façon dont Bob Port a pu ainsi élaborer des modèles de reconnaissance phonétique : "the input [auditory] pattern (...) shapes the way the state of the recurrent network evolves by modeling the dynamical landscape".

⁴Cet article de Langacker reprend en partie un article de fond présenté au remarquable colloque sur Lucien Tesnière organisé par Madame Madray-Lesigne à Saint-Aignan (Rouen) en 1992.

De façon particulièrement sympathique et intéressante pour la linguistique européenne, Ron Langacker, l'auteur de la "bible" *Foundations of Cognitive Grammar*, établit un parallèle entre la "syntaxe structurale" de Lucien Tesnière et sa propre grammaire cognitive⁴. Après avoir rappelé que "the basic tenet of cognitive grammar is that grammatical structure reduces to conventional patterns for the construal of

conceptual content and its phonological symbolization", il résume les principaux caractères de la schématicité grammaticale, en particulier en ce qui concerne les opérations de catégorisation et de "profiling". Développant alors de façon détaillée un exemple, il montre quelles sont les ressemblances et les différences avec le point de vue de Tesnière. La différence principale tient à la conception "sémantique" (conceptuelle) de la syntaxe elle-même. Mais les affinités sont néanmoins très profondes et Langacker peut donc conclure qu'avec la syntaxe structurale et la grammaire cognitive il s'agit bien d'une entreprise commune, malgré les différences d'époques et de contextes scientifiques.

C'est également à une confrontation historique que se livre Franson Manjali, l'un des meilleurs jeunes représentants indiens des grammaires cognitives et des modèles dynamiques⁵. Après avoir remis en perspective la façon dont les réinterprétations cognitives et morphodynamiques des syntaxes actantielles (Tesnière, Greimas, Fillmore) et du sémantisme verbal ont modifié la conception structurale (saussurienne) du langage, il présente le rôle essentiel joué par les relations actantielles (les rôles casuels : *karaka*) chez les grammairiens indiens (Panini, Patanjali et surtout Bhartrhari) : selon Sharfe, "they are the prime moving factors of the whole grammar". Cela est particulièrement intéressant car l'on trouve déjà clairement formulée l'idée réaliste que les cas (ici du Sanskrit) sont des *intermédiaires* entre la réalité et les catégories grammaticales et qu'il existe donc "a common substratum of ontology, cognition and grammar".

Peter Gärdenfors reprend dans le contexte de la sémantique cognitive la question des *propriétés*. Il montre d'abord que les conceptions référentielles et vériconditionnelles de la sémantique, même lorsque l'on dépasse le point de vue extensionnel tarskien pour adopter des points de vue intensionnels à la Kripke-Hintikka-Montague, aboutissent à des théories non naturelles et non viables des propriétés (elles seront par exemple identifiées à des applications de mondes possibles dans des ensembles d'individus). Il devient par exemple impossible de comprendre la perception, la signification ou l'induction des propriétés. En fait, ces théories *sous-déterminent* massivement la référence et ne permettent pas de distinguer des propriétés "naturelles" de propriétés artificielles ("goodmaniennes"). Bref, "there is something rotten in the kingdom of semantics". Pour résoudre ce problème, Gärdenfors, après avoir sévèrement critiqué les conceptions logicistes et propositionnalistes, s'appuie sur la notion d'*espace conceptuel* (espace de qualités muni d'un minimum de structures géométriques) et explique comment des langages formels peuvent être interprétés dans des espaces conceptuels à travers des "location functions" à la Stalnaker. C'est donc le concept de *localisation* dans des espaces conceptuels qui prend en sémantique cognitive la place des concepts d'individu et de monde possible constitutifs des sémantiques intensionnelles. Gärdenfors explique alors certains des bénéfices théoriques d'un tel point de vue. D'abord, la possibilité de caractériser la

⁵Signalons l'intéressant recueil «*Language and Cognition*» édité par F. Manjali (New Delhi, Bahri Publications, 1993).

“projectibilité” des propriétés naturelles par un critère mésologique de *convexité* des régions correspondantes. Cela permet de résoudre en partie le problème de l’inférence inductive. Ensuite, la possibilité de développer une théorie satisfaisante des effets de *prototypicalité*. Enfin, une réinterprétation des *raisonnements non monotones* comme raisonnements sur des situations génériques (typiques).

Per Aage Brandt, qui avait déjà montré dans *La Charpente modale du sens* comment le système modal des langues pouvait être interprété en termes de dynamiques “externes” (dynamiques de contrôle) dans des modèles morphodynamiques, présente ici une sorte de parcours génératif “inversé” des structures phrastiques. Il part de la situation communicationnelle générale et “descend”, par spécifications successives, jusqu’aux structures phrastiques. L’idée de base est de concevoir le schéma pragmatique et énonciatif de la scène communicationnelle comme une interaction dialogique entre trois méta-actants Ego/Alter/Quid et de la particulariser par paliers. Une importance particulière semble devoir être accordée à la genèse des universaux casuels-actantiels-narratifs (instrumental, agentif, objet, datif) comme “poids casuels”, c’est-à-dire comme polarisations ou tropismes dans l’espace “externe” de l’interaction dialogique E/A/Q. Leur lien avec le sémantisme verbal tient au fait que, dans une schématisation en termes de catastrophes élémentaires où les actants occupent les *minima* d’une fonction “énergie” (une fonction “harmonie” au sens de Smolensky), les verbes spécifient les *maxima* (les seuils) séparant les *minima*.

Pierre Ouellet a été l’un des tout premiers (dès 1977) à saisir l’importance du point de vue morphodynamique en sémiotique. Il propose ici une réflexion sur les liens existant entre “la dynamique psychophysique ou la phénoménologie de notre expérience sensori-motrice” et “la structure anthropologique et historique de notre imaginaire”. Il commence par commenter certaines caractéristiques du tournant cognitif-dynamique post-structural : l’ouverture du structuralisme statique vers la psychologie cognitive et l’expérience phénoménologique, le rôle décisif de la perception et de l’imagerie mentale, la “gestaltisation” du sémiotique. Mais ce qui l’intéresse ici est de relier ce qui se joue ainsi à “l’amont du sens” à ce qui se joue à “l’aval”. Pour cela, il élabore une théorie des différentes *esthésies*, c’est-à-dire des “différentes modalités historico-culturelles de la connaissance sensible”. Il commente alors un exemple de perception *discursive*, de “donner à voir littéraire” emprunté à la littérature descriptive contemporaine, où se noue une noétique de la perception à une noématique des mouvements de l’apparaître.

Depuis la parution en 1982 de *Catastrophe Theoretic Semantics*, Wolfgang Wildgen s’est attaché à fonder une grammaire morphodynamique sur des bases cognitives et sémiotiques. Il propose ici des éléments de base pour une sémantique *réaliste*, c’est-à-dire fondée sur le principe que le signe renvoie à une réalité externe qui est accessible à la

fois corporellement, phénoménologiquement et scientifiquement. Ce principe ("postulat réaliste") joint à une conception génétique du signe peircien impose des contraintes considérables à toute théorie sémiotique. Celle-ci doit être en effet une sémio-physique (une physique du sens) qui articule le sens à la fois sur les prégnances biologiques et sur les saillances perceptives à travers lesquelles le monde se structure et se donne phénoménologiquement.

Dans une approche dynamique, l'acceptabilité est pensée de façon continuiste (et non plus discrète : oui/non) en termes de *degrés*. C'est plus une optimisation de contraintes qu'un tri effectué par un automate accepteur de langages. David Piotrowski propose de relier cette question de l'acceptabilité à la dualité saussurienne *Sé/Sa* du signe linguistique. Après une analyse épistémologique — en grande partie inspirée des travaux de J.-C. Milner — de la nature particulière, prescriptive et "législative", de l'objectivité linguistique, il reprend la dualité du signe. Comment la repenser dans un cadre où ce ne sont plus les règles formelles de la langue (de la compétence) qui viennent au premier plan mais les structures dynamiques de la performance ? L'idée (originale) de Piotrowski est d'adopter un point de vue "holistique" sur les significations. Par le signifié restreint qu'il évoque, un signifiant activerait des éléments de contenu (des grandeurs de substance de contenu) qui, par propagation d'activations, fourniraient des accès à un espace de sens. On pourra remarquer que, si l'on remplace dans cette perspective "signifié" par "symbole mental", on débouche sur une conception assez proche de celle proposée aux narratologues par Herbert Simon : les *Sa* activent des symboles (i.e. les "chargent" dans une mémoire vive) et ceux-ci, en fonction de ce qui est déjà présent en mémoire, d'une architecture implémentant un ensemble de prescriptions et des connaissances stockées dans une mémoire à long terme, activent contextuellement d'autres symboles par une réaction en chaîne. L'idée est donc que les signifiés (au sens classique, saussurien, du terme) fonctionnent *comme un espace de contrôle* pour des dynamiques "internes" (correspondant à une "perception" du sens)⁶. Cela permet alors d'utiliser des modèles morphodynamiques puisque ceux-ci expliquent précisément comment des dynamiques internes engendrent des écarts différentiels dans des espaces externes au moyen de leurs bifurcations. Les oppositions sémiques et les traits distinctifs apparaissent ainsi comme des conséquences d'activités mentales. Ce renversement d'optique sur la catégorisation de la "nébuleuse" saussurienne du signifié permet de penser le différentiel d'acceptabilité comme la source dynamique des catégories sémiques (et non l'inverse). On rejoint ainsi ce qui a été démontré par Elman, à savoir que c'est l'apprentissage (supervisé par des jugements d'acceptabilité) des régularités syntagmatiques qui force le système cognitif à construire des paradigmes.

⁶Cette idée retrouve l'opposition fondamentale entre "encyclopédie" et "dictionnaire".

Enfin, pour conclure ce dossier, nous revenons de façon détaillée sur le débat ayant opposé J. Fodor, Z. Pylyshyn et B. McLaughlin à P. Smolensky à propos de la constituance dans les modèles connexionnistes. Nous avons montré techniquement ailleurs comment les modèles morphodynamiques permettaient de dépasser les limites du connexionnisme standard et de développer le concept d' "Attractor Syntax". Reprenant les termes d'une étude de 1989, nous nous focalisons ici sur le contenu épistémologique du débat. Il est en effet crucial pour la perspective dynamique.

Jean Petitot

Remerciements

Je remercie Jean-Claude Coquet de m'avoir proposé d'organiser ce numéro spécial de *Sémiotiques*. Je le prie, ainsi que les lecteurs, de m'excuser pour le retard considérable que j'ai fait prendre à la publication. Je remercie également Florence de Chalonge pour son impeccable travail éditorial. Enfin, j'exprime ma profonde reconnaissance envers tous les collègues qui ont accepté de collaborer à ce numéro.

Les Modèles connexionnistes entre perception et sémantique

Notes sur la nature et les fonctions des représentations

Yves Marie Visetti

1. Le Continu et l'espace dans les sciences cognitives

L'essor de la modélisation connexionniste mobilise, au moins à titre instrumental, une part de la conceptualité *mathématique* ayant trait au Continu et à l'espace. Les modèles qui en résultent suivent ainsi de près, et parfois favorisent, le développement des conceptions topologiques et dynamiques dans les sciences cognitives. Qu'il s'agisse ainsi de la nature des représentations ou de la compréhension des processus, la place du continu et du spatial ne cesse de croître. Les arguments — ou les symptômes — ne manquent pas, en effet¹ :

- le temps des processus cognitifs, ou plutôt leur fond temporel, est continu : c'est l'observateur, ou peut-être le sujet lui-même, qui introduit des segmentations dans ce "flot" cognitif (cf. le concept d'autochronie avancé par V. Rosenthal [1993]) ;

- la superposition de différentes échelles temporelles (non seulement celle, immédiate, du fonctionnement, mais aussi celles, non moins fondamentales, de l'adaptation, du développement, de l'évolution...) exige *a fortiori* le recours au continu² ;

- les analogies ou les correspondances entre modes d'organisation physique, biologique, ou psychologique passent généralement par les mathématiques du Continu (les systèmes dynamiques, par exemple) : on peut citer ici la théorie de la Gestalt et son analogie électromagnétique, ou bien la Théorie des Catastrophes de René Thom ;

- les évolutions combinant des phases stables et des changements brusques à la traversée de seuils critiques se conceptualisent précisément comme des *discontinuités* — et cette notion exige souvent la mise en place préalable d'un continu "de référence" ;

¹Nous nous bornerons ici à les énumérer, et ne tenterons pas de les justifier davantage. Notre intervention est en effet centrée sur la question des représentations : on espère mettre en évidence, en examinant les modèles avec une certaine précision, les possibilités nouvelles ouvertes par l'intervention croissante du Continu et de l'espace. Il n'entre évidemment pas dans nos intentions de censurer par-là les approches concurrentes qui feraient appel à une intuition discrète des mêmes phénomènes. Le dialogue reste ouvert, et nécessaire. On pourra en particulier se reporter dans cette perspective aux belles études critiques de J. M. Salanskis [1992, 1993], ainsi qu'à un numéro spécial de la revue «Intellectica» édité sous sa direction («Philosophies et sciences cognitives» [Salanskis, dir., 1993]). On pourra également chercher à transposer aux sciences cognitives les réflexions du mathématicien J. Harthong [1992].

²Il se peut que le fond temporel continu soit discrétisé pour les besoins de la modélisation. De même, les données fournies par les capteurs

physiques d'un système
peuvent bien être
numérisées d'emblée.
L'important, cependant,
est de ne pas confondre
ce micro-discret
primitif avec le macro-
discret souvent invoqué
pour rendre compte des
segmentations spatiales
ou temporelles à valeur
phénoménologique.
Bien sûr,
l'informatisation de ces
deux sortes de discret
consiste à établir entre
leurs tenant-lieux
informatiques un simple
rapport fini,
combinatoire, de
compilation. Nous nous
bornerons ici à
constater que ce
rapport est
généralement
incompréhensible sans
un apport massif de
mathématiques
continuistes, comme en
témoigne toute la
littérature sur la
reconnaissance des
formes.

- si le principe de l'évolution fait appel à un schème de valeurs spécifique du système ou de l'organisme (pulsion, affect), ce schème est souvent conçu à la façon d'une grandeur intensive susceptible de variation continue. Il en est de même, plus généralement, pour toute modulation des mécanismes attentionnels, dès le moment où on les représente grossièrement, en première analyse, comme des actions de masse ;

- toute interaction avec un environnement réel mobilise évidemment une physique (par exemple une physique du signal), et donc les mathématiques appropriées ;

- les linguistiques cognitives (Langacker, Talmy, Lakoff, Vandeloise, Herskovitz...) s'appuient sur une diagrammatique complexe, évocatrice de schématismes spatio-temporels menant à une sorte de topologie cognitive, volontiers continuiste ; de même, les recherches sur l'imagerie mentale, ou de façon plus spéculative, sur les modèles mentaux, s'accordent de plus en plus franchement sur la nature irréductiblement spatiale des représentations étudiées ou postulées ;

- tous les phénomènes graduels (saillance relative, probabilité, valeurs intensives...) laissent supposer un arrière-plan continu ;

- la séparation des fonds et des formes repose sur la discontinuité des qualités sensibles (au voisinage des contours) et sur leur continuité (à l'intérieur des formes, et sur tout le fond) ;

- enfin, le *topologique* est la principale alternative au *logique* pour une définition formelle des structures, et donc pour la construction d'un univers de représentations.

Une précision s'impose ici. Le Continu et l'Espace dont il est question ne sont pas à prendre pour des intuitions pures, dont les schèmes uniques et éternels se retrouveraient à chaque fois intacts dans toutes les occasions où l'on croit devoir les invoquer. Tout au contraire, leur généralisation et leur association indéfiniment variables dans le cadre de la topologie en ont fait éclater depuis un siècle la diversité des acceptions, tout en les rattachant au discret et au numérique. Les invoquer ne signifie pas non plus que tout domaine cognitif doive s'approcher comme un champ de variation continu. Mais le discret, quand bien même il s'imposerait parfois comme intuition première, notamment dans les sciences du langage, doit aussi être recompris comme un effet de discontinu. C'est là le sens le plus général de cet élargissement conceptuel : la manipulation adéquate de structures discrètes peut exiger qu'on étende leur espace de traitement en un espace continu et/ou contraint par une topologie. C'est l'expression même des lois et régularités cherchées qui est à ce prix : saisir les possibilités originales du continu et de l'espace mathématiques dans la description *qualitative* des domaines cognitifs.

Il est facile de trouver, pour chacune des rubriques de la liste ci-dessus, plusieurs travaux marquants en modélisation cognitive, où l'une ou l'autre de ces dimensions du continu et du spatial joue un rôle central, notamment (mais pas exclusivement) à travers l'instrument connexionniste, et plus encore, à travers les abstractions mathématiques, topologiques et dynamiques, qui le fondent. Il n'est évidemment pas question de passer en revue tous ces travaux, mais seulement d'en analyser quelques-uns particulièrement éclairants sur la question des représentations, de leurs structures et de leurs fonctions. Même avec ces restrictions, d'ailleurs, il ne sera guère possible de dépasser le niveau de l'évocation : mais nous pensons cependant qu'il est possible de livrer ici une perspective globale, un plan d'étude déjà significatif.

On voudrait produire ainsi l'esquisse d'une certaine variété d'*attitudes représentationnelles*, c'est-à-dire d'attitudes épistémologiques ou plus simplement méthodologiques, à partir desquelles nous *décidons* de comprendre les sujets, les organismes, ou les systèmes, en les finalisant doublement : comme substrats pour un système de formes abstraites, d'une part, et comme agents manipulant en *conformité* avec elles une certaine extériorité, d'autre part³.

Nous analyserons donc d'abord certains aspects de la modélisation (psycho)linguistique dans la ligne du groupe *PDP*, bien connu dans ces domaines depuis la parution en 1986 de son livre-manifeste *Parallel Distributed Processing*. La première moitié de notre intervention y sera consacrée. Nous évoquerons ensuite rapidement ses différences avec l'approche, bien plus rigoureuse sur le plan linguistique, des grammaires harmoniques de P. Smolensky, G. Legendre, et Y. Miyata [Smolensky *et al.*, 1992], ainsi qu'avec les modèles de polysémie de B. Victorri.

Quittant alors la sémantique linguistique proprement dite, sans pour autant perdre toute motivation de ce côté, nous présenterons les modèles d'ancrage dans la perception, principalement ceux de J. Petitot ou de T. Regier, qui tentent de rattacher certaines fonctions symboliques (par exemple l'identification des rôles actantiels), à une structuration qualitative de l'espace perçu, à travers une perception catégorielle généralisée d'objets, de processus, d'événements.

Nous terminerons en mentionnant un ensemble de travaux en robotique, aussi important que peu développé encore, où l'on aborde la question des représentations à partir d'une épistémologie "anti-représentationnaliste", selon laquelle seule une interaction constante avec un environnement est susceptible de *construire* les fonctions assurant la viabilité : il ne s'agit plus alors de rechercher l'ancrage d'une symbolique préexistante dans un espace de perception solipsiste, mais plutôt de

³C'est précisément cette conformité postulée qui constitue la simple extériorité en une objectivité propre au système de formes internes construit par l'attitude représentationnelle.

réaliser l'émergence de fonctions perceptives ou symboliques très rudimentaires à partir de l'espace-temps d'une interaction.

2. La modélisation (psycho)linguistique dans la ligne du groupe PDP

On connaît bien maintenant l'argumentaire, devenu classique, du courant connexionniste PDP. A titre de rappel télégraphique, voici quelques-uns de ses traits les plus saillants :

- importance du contexte (non nécessairement symbolique) dans tout processus symbolique : paradigme perceptif plutôt qu'inférentiel ;
- un associationnisme complété par des couches intermédiaires structurantes : les processus sont donc architecturés ;
- comme dans tous les modèles connexionnistes : complétion associative, gradualité des résultats, apprentissage ;
- conception topologique, et non pas logique, de l'espace des représentations : les signifiants et les signifiés sont des *formes* ;
- gradualité des objets :
 - représentations distribuées et superposées,
 - continu mathématique : poids des connexions, activités des unités ;
- rejet de l'atomisme et du logicisme, rejet de l'autonomie de la syntaxe, rejet des formes strictes de compositionnalité en sémantique ;
- interactions permanentes entre les différentes modalités cognitives, et donc architectures à modules non cloisonnés ;
- régularités de type statistique, parfois analysables en termes de distance à des valeurs de référence (prototypes) : donc rejet des "types" discrets des langages formels ;
- introduction d'expédients techniques pour donner aux représentations une structure interne ou *constituance* : représentation par produit tensoriel de Smolensky, assemblage de sous-structures par compression de données (Pollack, Chalmers), analyse de données sur les représentations induites par apprentissage révélant leurs composantes fonctionnelles ;
- complexification des architectures : réseaux récurrents et interactions de modules ;
- partage de représentations entre modules, et donc coadaptation de ces représentations à plusieurs tâches menées en parallèle ;

- thèmes de prédilection des travaux dans la lignée de PDP :
 - polysémie,
 - interactions valeurs lexicales/rôles casuels,
 - construction de paradigmes à partir d'analyses distributionnelles,
 - reconnaissance de scénarii et de schémas.

Le bilan, en forme d'objection, de ce programme de recherche est presque aussi connu :

- les représentations, lorsqu'elles sont précodées, restent la plupart du temps trop proches des structures classiques (le cas des représentations induites par apprentissage est différent) ; plus généralement, les modèles procèdent de linguistiques sommaires, qu'ils ne contribuent guère à faire évoluer ;

- les architectures et les processus sont encore trop simples ;

- les modèles couvrent des ensembles de données restreints, ou bien concernent des phénomènes ou des "tâches" très spécifiques ;

- l'interaction souhaitée entre perception, action et langage n'a pas progressé dans son principe, faute d'avoir fait appel aux linguistiques et aux théories de la perception et de l'action susceptibles de l'étayer ;

- parler d'émergence (d'une organisation symbolique) dans ces modèles relève essentiellement de la métaphore, puisqu'il n'y a pas d'environnement à proprement parler, et que l'activité modélisée se réduit à l'exécution d'une tâche préspecifiée ; ce faisant, on n'observe aucun changement d'échelle dans la description, ni aucune modification radicale des structures et des processus pertinents, qui signaleraient le passage d'un niveau d'organisation à un autre⁴ ;

- il manquerait à ces modèles une théorie nouvelle de la *constituance* et de la *compositionnalité* en linguistique, une théorie véritablement *topologique* et *dynamique*, qui offre une alternative réelle aux théories logico-symboliques de la structure (critique moderniste). Certes, on obtient la gradualité et la contextualité des effets, mais dans son fond, la *systematicité* des représentations construites ne s'est guère écartée de celle des modèles classiques. Dans ces conditions, pourquoi se priver de la facilité et de la clarté combinatoire des approches classiques (critique traditionaliste) ?

⁴On peut cependant admettre le terme d'émergence dès le moment où un mode d'organisation nouveau apparaît par apprentissage : dans cette mesure, le passage à une organisation géométrique des représentations, dans le cadre d'une adaptation à une tâche qui en est initialement dépourvue, constitue une forme d'émergence qui se rencontre dans ces modèles (cf. infra).

Discussion

Nous nous concentrerons exclusivement sur la dernière objection de la liste ci-dessus dans la suite de cette section. Plus précisément :

— Il est clair qu'aucun des modèles de PDP ne s'appuie sur une théorie linguistique élaborée de l'une quelconque des dimensions de la langue étudiées (lexique, structure sémantique de la phrase, structure de texte) ; quelques modèles très spécialisés, cependant, résistent à l'épreuve de corpus assez importants (analyse des valeurs lexicales de *over*, comme chez C. Harris [1990] ou attachement de groupes prépositionnels, comme chez Touretzky [1988] ou Lehnert, en 1988). Les auteurs n'ont généralement pas eu le souci d'élaborer de nouvelles méthodologies de recueil des données et de construction de l'objet propre de l'analyse linguistique (le cas de Paul Smolensky est exceptionnel à cet égard).

— Il ne se dégage pas vraiment de ces modèles de nouvelles caractérisations *dynamiques* des structures symboliques : en effet, il faudrait *d'abord* que s'opère une jonction avec une problématique abstraite posée en termes de systèmes dynamiques, et il faudrait disposer d'algorithmes d'apprentissage adéquats, servis par une puissance de calcul qui égale, enfin, celle que l'on investit habituellement en analyse d'images ou en mécanique des fluides. Nous n'y sommes pas encore. Pourtant, cette approche dynamique avait été largement inaugurée dans le recueil PDP lui-même, à propos de modèles de schémas, d'organisation de la mémoire, ou de reconnaissance de mots. Elle n'a pas été poursuivie par ces mêmes auteurs à propos des structures syntaxico-sémantiques. Notons également, à titre de restriction supplémentaire, la faible signification temporelle des modèles les plus dynamiques de ce courant : les dynamiques sont en effet construites en fonction de leurs équilibres asymptotiques (attracteurs), et non pour leurs états transients. Elles sont donc censées implémenter des formes ultra-rudimentaires d'homéostasie, et non modéliser une *synthèse temporelle* de représentations par un jeu subtil de délais, d'amortissements, d'oscillations, etc.

— Par contre, nous estimons que les modèles du courant PDP ont clairement contribué à promouvoir de nouvelles conceptions *topologiques* de la systématité des processus et de la constituance des représentations symboliques.

Développons ce dernier point, en prenant tout particulièrement appui sur les travaux de [Elman, 1991, 1993], [St John & McClelland, 1990] et [Mikkulainen & Dyer, 1991]. Tous ces auteurs ont en commun d'avoir utilisé des architectures de réseaux (faiblement) récurrentes, ce qui, sans être absolument nécessaire à notre propos, le rend plus significatif encore. En effet, ces architectures permettent de combiner chaque nouvelle entrée à la trace (recopiée) d'une partie de l'état interne du réseau à l'instant précédent. De cette façon, on peut contextualiser les traitements sans préjuger de la taille des fenêtres temporelles nécessaires, tout en facilitant l'apprentissage des aspects invariants par translation dans le temps.

Les tâches constituant la pression adaptative sur les réseaux comportent justement un élément d'anticipation (prévoir le mot suivant dans un texte, ou bien prévoir une description casuelle complète de la phrase courante), dont la réussite est en fait conditionnée par la détection de certains traits caractéristiques du contexte passé. Dans ces modèles, les codages sur les couches d'entrée et de sortie sont préspecifiés : ils sont locaux et discrets (une unité = un mot, un "concept" ou un "trait", selon les cas). Par contre, les représentations induites sur les couches intermédiaires se répartissent dans leur espace d'état, globalement un espace vectoriel ; un espace dont le codage fait appel au continu numérique des nombres réels, et la description aux concepts de la géométrie euclidienne. Ce sont ces représentations qui constituent l'apport principal de ces modèles à la construction d'un nouveau type de systématique symbolique.

En effet, leur analyse par différentes techniques de *clustering*, analyse discriminante ou en composantes principales, révèle des regroupements ou des répartitions, c'est-à-dire des caractéristiques géométriques, qui conditionnent le rôle fonctionnel de ces représentations dans le processus. De fait ce dernier *dépend continûment* de la position des représentations dans l'espace d'état, ce qui est déjà une très forte contrainte ; et de plus la pertinence de l'analyse de données dans l'analyse globale de cette dépendance a été démontrée avec une certaine généralité [Baldi & Hornik, 1989], [Gallinari *et al.*, 1991]. La topologie de l'espace d'état, et la continuité des processus qui s'y appliquent, déterminent donc une systématique particulière, *qui identifie les rapports fonctionnels à des rapports de position entre représentations*. Ainsi, la modification (permutation) de tel ou tel aspect d'une représentation se traduit par le déplacement d'une région à une autre, par exemple dans une des directions détectées par l'analyse de données : comme le dirait de mauvais gré J. Fodor, si la représentation de $R(A, B)$ est possible, il en va de même de celle de $R(B, A)$, et le passage de l'une à l'autre manifeste une certaine dépendance systématique, que l'on doit décrire en termes géométriques. *De même, la possession d'un aspect ou constituant commun n'est rien d'autre que l'appartenance à une région commune*. Dans le cas où les approximations linéaires de l'analyse de données peuvent s'appliquer, cela signifie qu'une composante de représentation, c'est-à-dire l'analogue d'un constituant, peut s'identifier à sa projection sur tel ou tel sous-espace discriminant. Par exemple, si l'on s'intéresse au rôle *agent* dans une analyse casuelle, on pourra déterminer un sous-espace sur lequel les représentations de la structure casuelle globale se projettent en formant des agrégats bien distincts, correspondant respectivement aux valeurs *Marie, Jean, chien, chat...* pour ce rôle. Et la représentation de ces valeurs, en tant que facteurs distinctifs pour les configurations globales qui les

contiennent, est à chaque fois la projection obtenue. Cependant, il faut souligner que les *transformations* entre composantes représentationnelles, quoique systématiques, ne se laissent pas réduire à l'action d'un opérateur primitif et universel : il faut en effet les "réapprendre" pour tout espace de représentations spécifique. La représentation d'un constituant, qui dans une approche classique serait à chaque occurrence identique à lui-même, c'est-à-dire à son type formel, est donc ici éminemment variable selon les contextes. Ainsi les projections qui constituent les représentations vectorielles de *Marie* ne sont pas *a priori* les mêmes pour les rôles *agent* ou *patient*⁵.

⁵On pourra consulter sur ce point l'étude théorique et expérimentale de M. Crucianu [1993].

Remarquons également que ces répartitions géométriques se laissent souvent redécrire intuitivement, quoique de façon très approximative, à partir de systèmes de traits ou de classes déjà attestés en syntaxe ou en sémantique (sujet/complément, nom/verbe, animé/inanimé, humain/animal...). On peut d'ailleurs, à l'inverse, commencer par fixer un tel trait avant de rechercher ses meilleurs axes discriminants. Il y a en somme construction *géométrique* de classes ou de dimensions à valeur syntaxico-sémantique, et corrélation systématique entre les phénomènes grammaticaux et la répartition des représentations relativement à ces classes ou dimensions. Autrement dit, la géométrie des paradigmes découle, par apprentissage, de l'effectuation d'une tâche le long de la syntagmatique d'un corpus : *le lien entre paradigme et syntagme est dans un accord profond avec celui que le structuralisme le plus canonique revendique*, même si, dans les faits, ces modélisations ont été engagés jusqu'ici au service de linguistiques peu rigoureuses. Et de surcroît, le lien structural entre paradigme et syntagme est *généralisé* à d'autres tâches que celle de *commutation* : la construction des paradigmes peut ici dépendre — en théorie — d'une multitude de critères, et en particulier du format topologique des espaces internes alloués pour déployer les distinctions pertinentes (autrement dit, le paradigme construit dépend de son substrat géométrique, il peut être plus ou moins "compressé"). Ainsi par exemple, on pourrait envisager — mais cela demanderait des moyens de calcul très importants — de construire par les mêmes méthodes des paradigmes *textuels*, et non plus seulement lexicaux ou grammaticaux : les représentations de (sections de) textes se répartiraient alors en fonction de certaines opérations de paraphrase, de résumé, ou de classifications thématiques, dont les résultats seraient codés sur les couches de sortie.

Comparons alors cette conception de l'organisation structurale avec celle de l'ordre symbolique logico-formel qui est, comme l'on sait, tout à la fois :

— le soubassement de la théorie fonctionnaliste classique des représentations mentales ;

- la base interprétative des sémantiques formelles en linguistique ;
- et la forme sous-jacente à tous les langages de représentation de l'IA classique.

La structure logique est quant à elle une structure en constituants discrets, finis et détachables, où les constituants s'agrègent par l'intermédiaire d'un répertoire fini de relations. Elle appelle donc des traitements compositionnels par démontage, modification combinatoire des constituants, et remontage. Les traitements ne dépendent que de la forme (logique) des constituants et des relations, c'est-à-dire de leur *type* computationnel, qui ne comporte aucune modification dans chacune de ses occurrences. Il y a une différence capitale entre les constituants primitifs ou atomiques, qui ne présentent aucune diversité interne pertinente, et les constituants complexes ou moléculaires, dont la structure importe : sans cette différence de nature, le traitement récursif des structures s'engagerait en effet dans des boucles infinies. De plus, si l'on veut faire jouer à la forme syntaxique un rôle explicatif *causal* — le philosophe J. Fodor insiste beaucoup sur ce point à propos des supposées représentations mentales — le rapport entre type et occurrence doit être un rapport entre une forme abstraite détachable et une matière indifférente ; et de la même manière que la forme du type est syntaxiquement analysable, la matière de l'occurrence doit être dissociable selon les lois d'une méréologie naïve. Une occurrence de structure complexe comporte donc, à l'état de *parties* effectivement détachables, des occurrences de ses constituants formels. De manière comparable, les processus syntaxiques formels doivent, eux aussi, être en correspondance stricte avec les processus matériels qui les instancient sur chaque occurrence : c'est de cette façon seulement que la syntaxe et les règles formelles peuvent être dites *causer* un comportement.

Qu'en est-il maintenant pour les représentations de type PDP ?

• Le rapport entre type et occurrence est profondément transformé : un type, pour autant qu'on veuille toujours en reconnaître, est une région de l'espace d'états, tandis qu'une occurrence est un point de cet espace. Les régions-types se recouvrent partiellement, ne serait-ce qu'à leurs frontières, et les représentations-occurrences ne relèvent donc pas nécessairement d'un (seul) type bien défini. Il peut se faire, cependant, que les régions-types ne servent qu'à marquer le territoire d'un point de référence particulier, un "prototype" si l'on veut, dont seule l'identification compte *in fine*. Il faut alors procéder à une normalisation des résultats, de façon à leur associer la référence la plus proche, ou bien interpréter la distance aux références des représentations-occurrences comme un indice de modalisation (probabilité, saillance...) : c'est par exemple ce qu'il convient de faire dans l'interprétation des couches de sortie, lorsqu'elles codent de façon locale des catégories discrètes.

- L'image méréologique naïve de la constituance (et donc de toute compositionnalité) ne peut plus être soutenue : qu'est-ce qu'une partie de point, en effet ? Il convient de la remplacer, en première analyse, par une image en termes de composantes vectorielles, c'est-à-dire en termes de projection sur des sous-espaces bien choisis. Notons qu'alors un "constituant", c'est-à-dire une composante, n'est pas toujours dans une relation évidente avec les composantes représentant ses sous-constituants : il n'en dépend pas nécessairement à travers une loi simple, comme une addition vectorielle — tout dépend ici des espaces de projection respectifs, voire d'opérations plus complexes, non linéaires, effectuées par des réseaux spécialisés (Pollack, Chalmers). Non seulement, donc, une occurrence de structure complexe ne *contient* pas ses éventuels constituants (cela n'a plus guère de sens), mais plus encore elle ne les révèle pas nécessairement avec une grande facilité. La reconstruction de constituants peut être coûteuse. Ce point a été bien analysé par van Gelder, qui a proposé d'opposer dans cette perspective les notions de compositionnalité concaténative et fonctionnelle.

- Si la reconstruction de constituants, ou plus exactement de composantes, peut être coûteuse, elle n'est pas toujours nécessaire : dans la plupart des modèles étudiés, le traitement s'applique à la représentation telle quelle, sans qu'il soit besoin de modéliser explicitement, dans une architecture de modules spécialement dédiée, les opérations de montage-démontage typiques des approches compositionnelles. C'est indirectement, par le biais de l'apprentissage, que les composantes intéressantes spécifiques au problème sont détectées. Dans ce cas, l'identification de ces composantes se fait dans l'après-coup, par des moyens d'observation externes ; à strictement parler, elle ne joue aucun rôle causal dans le processus, même si elle contribue à en clarifier le fonctionnement pour nous. Les descriptions qualitatives de la systématité coïncident donc rarement avec les explications processuelles de cette même systématité : cette conclusion est tout particulièrement soulignée par Smolensky. Il paraît difficile de soutenir dans ces conditions qu'il s'agirait toujours d'un calcul *sur* des représentations, même si, évidemment, l'informatisation des modèles plonge le tout dans une algorithmique.

- Il apparaît alors que la notion de *constituant* perd de son inhérence, elle résulte par exemple du parti-pris d'un observateur, qui sélectionne une métrique euclidienne particulière avant de procéder à une analyse de données. Naturellement, les métriques ne sont pas sélectionnées au hasard. Cependant, ce fait, combiné à la continuité de l'espace d'état, élimine l'idée d'une analyse ultime et canonique des systèmes de représentations. On peut souligner également que la différence de nature entre structures atomiques et complexes n'est plus intrinsèque : en effet toutes les représentations sont des éléments du même espace vectoriel. Seule la

donnée d'une base vectorielle permet de ressusciter, en partie, cette distinction (on prend en général la base constituée par les unités du réseau, mais ce n'est pas obligatoire : d'où également l'obscurité des débats sur les avantages et inconvénients supposés des représentations locales ou distribuées — notions on ne peut plus dépendantes de la base de référence adoptée). Dans le domaine de la modélisation linguistique, cela ouvre en tout cas la possibilité de traiter, sous certains aspects de leur sémantique, des lexèmes ou des syntagmes plus complexes, sur le même plan que de "simples" sèmes ! On dispose donc, somme toute, d'une notion de constituance sans primitives atomiques, et toujours relative à un cadre de décomposition variable. Une contrepartie peut-être fâcheuse à cette situation serait qu'on ne peut appuyer directement sur cette constituance la "correction" syntaxique des représentations formées : elles sont toutes plus ou moins acceptables. Si l'on estime important de disposer de critères de rejet, il faut donc les apprendre séparément.

L'utilité *spécifique* de ces traductions géométriques des rapports de constituance appellerait toutefois davantage de confirmation. Une première phase de travaux s'achève, où les questions d'existence et d'intertraductibilité avec les structures classiques étaient prioritaires. Reste maintenant à valoriser ce nouveau style de représentations : en montrant d'abord comment les réinvestir hors du cadre adaptatif précis qui leur a donné naissance, tout particulièrement pour faciliter des apprentissages ultérieurs ; en montrant ensuite comment ces représentations expriment des régularités intéressantes que la simple mention des tâches et structures engagées initialement (sur les couches visibles des réseaux, contrôlées par le modélisateur) ne permettait pas de prévoir. C'est dans ce contexte que le point suivant prend alors toute son importance.

- Sans que nous soyons capables de l'articuler ici avec la précision nécessaire, il nous semble alors qu'il convient d'infléchir le point de vue, très courant, selon lequel le progrès en matière de modélisation connexionniste en linguistique ou psychosémantique dépendrait d'un progrès de la "réflexivité" des réseaux (construction en temps réel de composants ou de hiérarchies, explicitation de traits ou de règles...). Si l'introduction d'une certaine forme de réflexivité paraît en effet nécessaire, elle ne peut cependant consister en l' "extraction d'informations" que l'on détacherait définitivement de leur contexte d'origine : très probablement en leur faisant subir au passage une discrétisation destinée à les identifier strictement, en fonction d'intuitions toujours contestables, aux caractéristiques d'une nouvelle tâche. Ce serait alors couper ces "informations" de leur source adaptative, et leur ôter de surcroît la précision numérique essentielle à ce type de modélisation. Il nous apparaît plutôt que le bon usage de ces informations tirées d'une observation d'un premier réseau serait de les prendre comme *conditions initiales*, et donc

temporaires, pour un nouvel apprentissage dans le cadre d'une tâche plus complexe, associant de nouveaux modules à l'architecture courante. L'interprétation intuitive de ces conditions initiales pourrait motiver l'architecture du réseau complexe, et orienter le "recollement" des nouveaux modules, étant entendu que toute opération de recollement peut entraîner des modifications rétroactives, après apprentissage, dans la partie initiale de l'architecture. De même pour les structures en "constituants" des représentations : si l'on estime avoir affaire à une tâche fortement compositionnelle, on peut l'apprendre sur la base d'une utilisation explicite des "constituants". Mais une fois obtenue une certaine efficacité sur les "structures" globales, il faut terminer l'apprentissage, puis fonctionner en temps réel, sans redécomposer les "structures". Ainsi *l'effectuation* des tâches ne repose plus explicitement sur la stratégie (dé)compositionnelle. Seul leur *apprentissage* y ferait appel, dans le cadre d'une méthodologie de construction évolutive de représentations et d'architectures de plus en plus complexes. Les règles et structures classiques n'interviennent plus alors que pour la description et l'apprentissage, et non pour le fonctionnement compétent, conformément d'ailleurs au rôle que certaines épistémologies souhaitent leur attribuer (H. et S. Dreyfus, P. Smolensky).

Le cognitivisme du courant PDP

Nous venons d'examiner, dans ses grandes lignes, la formalité spécifique de certaines structures de représentations connexionnistes. Quoique les instruments et les schémas de modélisation ne contraignent pas absolument les épistémologies qui les enrôlent (bien au contraire, il est fréquent de voir des épistémologies inconciliables se les disputer), il est bien clair que les modélisations du courant PDP relèvent d'une approche purement mentaliste des processus cognitifs, bref d'un cognitivisme représentationnaliste particulièrement bien servi par les instruments qu'il s'est donnés. Comment peut-on alors situer ce cognitivisme par rapport à celui, plus classique, qui s'appuie sur la forme logique ? Procédons par points.

- Le représentationnalisme du courant PDP n'est certes pas niable : les réseaux ne sont pas là pour appuyer un éliminativisme neuronal immédiatement exécutoire, mais au contraire la construction d'une psychologie. Cependant, le détachement ontologique de la couche représentationnelle s'appuie sur des moyens qui en minent plus ou moins insidieusement l'autonomie. Les représentations *construites* par apprentissage sont en effet dépendantes d'une architecture (en termes plus abstraits d'un espace d'état et d'une famille paramétrée de dynamiques) et

d'une tâche. Reprenant ici, de façon métaphorique, la distinction gestaltiste entre *moments détachables* et *moments dépendants*⁶, nous dirons que *les représentations de style PDP sont des moments dépendants de tâches et d'espaces substrats*. Elles entretiennent donc avec cet entour et cette base un rapport méréologique non trivial. C'est en amont que se situent les détachements méréologiques "naïfs" qui permettent ici l'objectivation : le précodage des entrées et sorties, c'est-à-dire la mise en forme d'un préjugé sur ce qui doit focaliser l'attention, et l'isolation d'une "tâche" à effectuer. Des considérations analogues valent pour la constitution temporelle de ces représentations : le temps évoqué ici n'est certes pas celui d'une expérience ou interaction réelle, c'est celui de l'adaptation inductive à un échantillonnage, à partir de conditions initiales. Il n'en reste pas moins que les représentations sont bel et bien ici des *moments dépendants d'un apprentissage*.

- Le type de formalité de ces représentations s'écarte notablement de celui des structures logico-symboliques, comme nous l'avons montré plus haut. Insistons à nouveau sur l'irréductibilité des propriétés processuelles des représentations à la simple donnée d'un type. Alors que dans le cognitivisme classique, seule la détermination *a priori* de types idéels importe, la plupart des approches connexionnistes, au contraire, partent concrètement d'un échantillonnage d'occurrences, et soulignent ainsi la nécessité d'une *théorie des occurrences* (impliquant donc une idéalisation adéquate de ces occurrences), en même temps que d'une théorie de la *généralisation*. Bien entendu, ces considérations ont une limite évidente, insurmontable peut-être — le codage des entrées ou sorties qui relève d'une typologie classique. La gradualité et la contextualité des déterminations psychologiques, et donc des représentations qui tentent de les figurer, constituent certainement un élément fondamental de l'épistémologie du courant PDP. Une certaine forme de continuité est donc ici primordiale. Par contre, le fond spatial abstrait, l'organisation topologique, n'étaient pas posés *a priori*, ils semblent plutôt un effet d'après-coup de la problématique adaptative : l'espace est donc intervenu dans un premier temps comme en analyse de données classique, tout simplement pour *distribuer* des descriptions. Son rôle *constituant* pour un système de représentations n'a été reconnu que très progressivement.

- Peut-on encore parler de *calcul* sur des représentations, dès le moment où l'on sort du cadre logico-syntaxique ? Si l'on conçoit les représentations comme des *formes*, ne devrait-on pas, par simple dérivation, parler alors de *transformations*, et considérer les systèmes comme des *transformateurs*, et non pas des calculateurs ? L'usage généralisé du vocable *computationnel* (vision computationnelle, neurosciences computationnelles, etc.), et l'informatisation des modèles, qui les réduit à une algorithmique, ne contribuent pas à clarifier les choses.

⁶Les moments détachables sont les parties séparables et indépendantes d'un tout morcelable (les roues d'un vélo). Les moments dépendants, comme les couleurs ou les contours, ne sont pas détachables de la même façon, et entretiennent avec leur extension un rapport de dépendance. Naturellement, une méréologie raffinée comprend bien d'autres distinctions. Mais ce qui nous importe ici est plutôt de généraliser les conditions de la dépendance des moments au delà de la simple extension, pour y inclure tout le cadre de détermination ou de construction des représentations.

Bien souvent — et c'est le cas pour les travaux examinés dans cet article — le mot *computationnel* ne pointe pas vers la théorie formelle de la calculabilité. Il signifie seulement que les modèles présentent les modifications de la forme des représentations comme l'effet de certains processus, tout aussi déterminés que les représentations auxquels ils s'appliquent. *Computationnel* indique en somme la construction d'une *effectivité* processuelle complète pour les transformations modélisées.

- Le découplage, c'est-à-dire le détachement méréologique "naïf" entre un *software* et un *hardware*, caractéristique du fonctionnalisme classique, est profondément transformé. Certes, il n'est pas question dans les modèles de PDP d'un rapport réel au "substrat biologique", ni même à l'espace physique. En revanche, même si la dépendance des représentations vis-à-vis du substrat se ramène à la contrainte d'une installation sur un espace topologique abstrait, elle réintroduit le loup (l'espace) dans la bergerie des représentations. Il n'est plus possible de spécifier le rôle d'une représentation sans lui assigner une position, et les positions font système sur la base d'une topologie.

- Nous avons rappelé tout au début de cette section que l'interaction, souhaitée par les principaux auteurs du courant PDP, entre représentations, perception et action, n'avait guère eu lieu. Même s'il ne fait l'objet d'aucune promotion explicite dans l'épistémologie officielle du courant, le *solipsisme méthodologique*, ou si l'on préfère la mise entre parenthèses du monde dans la construction et la détermination fonctionnelle des représentations, reste parfaitement compatible avec les principales réalisations et leur méthodologie. Reste que les instruments sont en eux-mêmes tout à fait neutres sur cette question, et qu'à défaut d'un interactionnisme environnemental ou social, les modèles de type PDP pointent vers une conception interactionniste, mais purement internaliste, des modalités cognitives.

3. Une montée vers l'abstraction : les grammaires harmoniques de P. Smolensky et les modèles de polysémie de B. Victorri

On cherche ici à introduire une organisation topologique et dynamique sur les espaces de traits de l'analyse componentielle classique, de façon à y intégrer des déterminations grammaticales de plus en plus riches. La construction de la signification d'un syntagme s'apparente alors à la formation d'une *Gestalt*, corrélée à des mécanismes d'optimisation ou à des principes homéostatiques. On cherche donc à introduire en linguistique une notion de *bonne forme* ou forme *harmonieuse*. Cela exige de définir cette forme en termes simultanément syntaxiques et

sémantiques. De nouvelles méthodologies linguistiques sont introduites en retour, et la conception des objets en est transformée. On peut y voir également la poursuite ou la relance des approches véritablement *dynamiques* en sciences du langage et en psychologie : approches partiellement redécouvertes dans les années 80 par le courant PDP, qui les a ensuite délaissées. Nous résumerons sous forme télégraphique quelques-uns des traits les plus saillants de ces problématiques.

- Satisfaction de contraintes et optimisation pour Smolensky, systèmes dynamiques, attracteurs et bifurcations pour Victorri : les modèles connexionnistes n'ont de valeur qu'en vertu de leurs liens avec un certain cadre mathématique, mis en position d'instrument principal (c'est cependant moins net chez Smolensky, dont le vocabulaire se fait plus spontanément neuromimétique).

- Il y a en effet tentative de *schématiser* certains objets de la linguistique, c'est-à-dire d'en reconstruire les concepts dans une détermination mathématique, *avant* de les modéliser informatiquement.

- Selon Smolensky, Legendre et Miyata, les grammaires ne sont pas des dispositifs à la productivité libre et autonome : leur fonction est d'attribuer aux énoncés, c'est-à-dire aux structures qui leur sont associées, une certaine acceptabilité (à ne pas confondre avec la simple acceptabilité syntaxique qui circonscrit l'objet de la linguistique dans les approches générativistes de la compétence). L'optimisation de contraintes est alors un principe d'organisation fondamental des grammaires. La contrainte globale à optimiser, qui est la somme de multiples contraintes plus ou moins locales, est appelée *Harmonie*. Une structure linguistique est alors d'autant mieux formée que sa valeur d'harmonie est élevée.

- Les grammaires sont alors susceptibles d'une double présentation : descriptive, par règles et symboles, à un niveau qualitatif par essence approximatif ; computationnel-effectif, à travers une fonction *Harmonie* agrégeant les contraintes. Cette fonction détermine la dynamique d'un réseau, qui l'optimise (c'est là tout le rôle des modèles connexionnistes). On calcule ainsi les configurations de traits syntaxiques et sémantiques dont la combinaison satisfait au mieux les contraintes graduées de la grammaire, c'est-à-dire correspond aux *maxima* de la fonction *Harmonie*. On obtient ainsi une sorte de reconstruction connexionniste de formalismes qui ont été jusqu'ici développés dans un cadre logico-symbolique sous le nom générique de *grammaires d'unification*. Dans ce nouveau cadre, le calcul de la représentation sémantique d'un syntagme se comprend comme une sorte de *complétion associative* : on fournit un indice (le syntagme lui-même), et l'on récupère, du moins en théorie, l'ensemble des traits syntaxiques et sémantiques qui en constituent l'analyse. Ce sont donc les traits qui, combinés à l'indice, constituent la

configuration d'harmonie maximale parmi toutes celles qui contiennent possiblement ce même indice.

⁷Elle présuppose le rassemblement d'un corpus, et sa démultiplication éventuelle par des opérations de paraphrase aux résultats diversement acceptables. On produit ainsi des énoncés d'harmonie variable, qui révèlent par contraste les contraintes grammaticales. La méthode dépend donc de l'intégration explicite d'exemples-limites ou de contre-exemples.

- La méthodologie de la modélisation est alors la suivante⁷ :

— On utilise les paramètres descriptifs les plus sûrs fournis par la théorie linguistique (des positions syntaxiques aussi bien que des traits sémantiques) en les combinant éventuellement à de nouveaux traits potentiels, correspondant à des traits déjà invoqués en linguistique mais dont l'attribution intuitive est plus flottante. Ces traits seront portés par les unités "cachées" des réseaux impliqués dans la modélisation, et leur signification fonctionnelle sera partiellement déterminée par les mécanismes d'apprentissage. Chaque syntagme étant ainsi évalué, il est codé par un vecteur de 0 ou de 1.

— On fait appel aux jugements d'acceptabilité graduée des locuteurs sur un échantillonnage extrait du corpus : chaque syntagme se voit ainsi attribuer une valeur d'harmonie choisie dans un ensemble très réduit (mettons 4 valeurs possibles).

— On applique un algorithme d'apprentissage qui associe à chaque vecteur de traits sa valeur d'harmonie. On construit ainsi, par généralisation, une fonction *Harmonie* que l'on peut tester sur le reste du corpus. On peut notamment analyser l'apport exact de chaque trait à l'harmonie d'une configuration globale. La méthode a été appliquée à l'analyse de la distinction *inaccusatif/inergatif* entre verbes du français. Elle a permis de redéfinir et préciser le rôle d'un trait classiquement utilisé dans ce contexte, qui introduit une distinction entre verbes possédant un Sujet ou un Objet en structure profonde. Les analyses antérieures tendaient à en faire la détermination unique, sinon principale, de la distinction *inaccusatif/inergatif*. Le modèle harmonique de la grammaticalité permet en revanche de l'intégrer de façon très précise, mais comme une détermination partielle.

— Seul un modèle restreint a donc été conçu et implanté. Il reste à déterminer si les résultats pourront s'étendre de l'étude des valeurs d'un terme pris en contexte au traitement d'une phrase entière. L'inscription de traits syntaxiques complexes (hiérarchies, rattachements) devra faire appel, selon Smolensky, à la structure algébrique de produit tensoriel. Or cette technique aboutit fréquemment à une démultiplication importante de la taille des réseaux impliqués.

- De son côté, B. Victorri préconise une construction *gestaltique* de la signification globale des syntagmes, à partir du couplage des potentiels sémantiques attachés aux lexèmes et aux fonctions grammaticales. La forme du sens est celle d'une *dynamique*, ou plus exactement d'un *champ de vecteurs* paramétré par des indices cotextuels. Le nombre et

l'agencement de ses bassins d'attracteurs caractérisent cette forme pour l'essentiel. Chaque unité linguistique contribue à construire le sens global de l'énoncé (c'est-à-dire la structure globale d'un champ de vecteurs) ; et en même temps le sens de cette unité se voit déterminé par l'énoncé qui la contient (c'est-à-dire qu'elle se voit attribuer, par une sorte d'effet de relocalisation de la structure globale du champ, une part spécifique du paysage dynamique ainsi construit). Ainsi le phénomène de la *polysémie*, et sa modulation indéfiniment variable d'un énoncé à l'autre (ajustement qui ne relève pas seulement d'une pragmatique du discours, mais d'abord de la sémantique linguistique proprement dite), est l'indice principal de la justesse de cette approche, en même temps que son terrain d'application privilégié. Car la variété des paysages dynamiques permet de représenter avec précision toute une variété de cas de figure interprétatifs, depuis la coévocation de valeurs distinctes jusqu'à l'ambiguïté-alternative.

- Le principe général de la modélisation est repris de la théorie des Catastrophes, et mobilise donc deux espaces, dits de *forme* et de *sens* (espaces externe et interne dans la terminologie de Thom). Pour le moment, seul un modèle partiel est véritablement opératoire. Il se focalise sur le sens d'une seule unité polysémique, dans ses interactions avec le reste d'un énoncé. Dans l'espace de forme s'inscrivent alors les déterminations cotextuelles, exprimées dans les termes les moins sémantiques possibles (il y a ici un danger de circularité, cf. [Victorri & Fuchs, 1992] et [Victorri, 1994]). La structure topologique et les dimensions principales de l'espace des sens (correspondant aux *traits* de l'analyse componentielle classique) sont en relation méthodique avec les jugements de proximité, repérage et recoupement que les locuteurs peuvent exprimer à propos des énoncés rassemblés dans un corpus. Ces énoncés échantillonnent la gamme des emplois possibles du terme analysé (en l'occurrence le marqueur grammatical *encore*).

- La technique informatique de modélisation repose sur la représentation des champs de vecteurs (vus comme systèmes dynamiques) sous la forme de réseaux connexionnistes récurrents. Leur dynamique peut justement être paramétrée par les indices cotextuels, convenablement codés. Tout énoncé donné en entrée au système est donc codé sous la forme d'un tel vecteur d'indices ; il s'en déduit automatiquement une dynamique sur le réseau récurrent, dont la forme, résumée dans la collection de ses attracteurs, représente le cas de figure interprétatif étudié. Bien entendu, on fait appel pour la construction de ce système à un algorithme d'apprentissage, et à une analyse préalable d'une partie du corpus qui met en oeuvre les *traits* dégagés dans la phase précédente. On notera cependant que, tout comme dans le cas des grammaires harmoniques, un modèle global du sens de l'énoncé reste à construire. On ne dispose là que d'un modèle partiel, centré sur une seule unité

⁸Smolensky continue de parler de représentations mentales, alors que Victorri a une conception purement instrumentale de ces formalismes. C'est également notre point de vue : il s'agit d'abord et surtout d'une extension de l'appareil sémiotique investi dans les sciences du langage, que l'objet en soit considéré comme social ou psychologique.

⁹En fait, les linguistiques cognitives dont ces travaux se réclament ne sont pas en tant que telles l'objet de la modélisation. Elles en constituent plutôt une puissante motivation. Ma présentation de cette catégorie de modèles emprunte largement à un texte rédigé en 1993 avec J. Petitot à l'occasion d'un appel d'offres du PRC Sciences Cognitives (nous l'avions centré sur le thème "Constituance perceptive et analyse dynamique de scènes").

linguistique. Passer au modèle global implique de théoriser plus précisément la forme de *compositionnalité* du sens qui convient à l'approche gestaltiste invoquée.

- Le continu et l'espace jouent ici un rôle primordial dans la constitution des objets linguistiques eux-mêmes. Par des biais différents⁸, les auteurs mentionnés dans cette section fondent l'organisation grammaticale, et notamment les paradigmes lexicaux, sur une "modalisation" dynamique des espaces d'états. En profondeur, c'est une structure mathématique de champ de vecteurs (ou plus simplement une fonction potentiel appelée *Harmonie*) qui schématise les rapports entre les valeurs représentées : ainsi une représentation n'est pas seulement la donnée d'un point de l'espace d'états, elle ne prend sens que dans le cadre géométrique défini par le champ (ou dans le cadre d'évaluation défini par la fonction *Harmonie*). Dans le cas des champs de vecteurs, le Continu de l'espace substrat est la condition préalable des *discontinuités* des régimes asymptotiques introduits par la dynamique : et la description qualitative *discrète* de la langue en découle. Des considérations analogues valent pour l'approche par les fonctions *Harmonie* : la discrétisation repose alors sur la détection de *pics* de valeurs. Par ailleurs, les espaces engagés sont des espaces *construits* ; ils sont le produit d'une méthodologie particulière de description linguistique, et n'ont donc pas le statut d'un espace de l'intuition ou de l'imagination directement appréhendable. Bien loin de figurer des bases prélinguistiques ("anté-prédicatives") de la prédication linguistique elle-même, comme ils sont censés le faire dans les modèles de la section suivante, ces espaces *résultent* au contraire de l'activité sémiotique des théoriciens-locuteurs, qui les utilisent pour synthétiser dans une topologie la variété de leurs jugements (d'acceptabilité, de proximité, de repérage, de recoupement, etc.).

4. Les modèles d'ancrage dans la perception et les grammaires cognitives

On cherche ici à dégager, à l'intérieur de la sémantique des expressions locatives (verbes, prépositions), des noyaux de sens comparables à des structures perceptives virtuelles, voire à des structures de l'imagination. On les modélise alors par des schèmes iconiques et dynamiques semblables aux diagrammes des linguistiques cognitives (J. Petitot, T. Regier)⁹.

- Rappelons rapidement les postulats principaux des linguistiques cognitives (Fillmore, Talmy, Langacker, Lakoff...) :

— pas d'autonomie de la syntaxe (les catégories grammaticales sont donc définies comme des types *d'interaction* entre syntaxe et sémantique, entre forme et sens) ;

— schématicité en syntaxe et sémantique, c'est-à-dire spécificité et productivité variables des schémas grammaticaux ;

— conceptualité ouverte et encyclopédique du sens, et modèles mentaux ;

— pas de compositionnalité stricte en sémantique, pas de primitives atomiques mais des domaines fondamentaux ;

— organisation radiale des catégories et systèmes de contraintes.

• La description sémantique ne repose donc plus sur l'atomisme (liste de primitives) et la forme logique (comme mode d'assemblage des constituants). La raison principale en est que les auteurs de ce courant insistent sur l'enracinement perceptif, et plus généralement sensori-moteur et kinesthésique, de dimensions très primitives, mais fondamentales, de la sémantique des langues naturelles. Leur approche cognitive de la linguistique — comme d'ailleurs de toute fonction symbolique — les amène à postuler la fonction structurante, pour ces dimensions de la sémantique, de schèmes spatio-temporels et dynamiques, constituant une vaste généralisation des *Gestalten* de la perception visuelle. Cette thèse s'applique tout naturellement lorsqu'il s'agit de la sémantique des expressions locatives (des verbes comme *entrer, sortir, tomber, traverser...* ou des prépositions comme *au-dessus, à travers, dans, contre...*, cf. Lakoff, Brugman, Regier, Vandeloise). Leurs noyaux de sens sont représentés par des schèmes iconiques, des *images-schèmes* figurés par des diagrammes.

• Mais en fait la thèse déborde le cas particulier des champs lexicaux et morphologiques manifestement concernés par l'expression des lieux et des temps. Elle concerne en réalité toute la sémantique, comme on peut le voir par exemple chez un auteur comme Fillmore. Car toute prédication profile des entités sur un fond, elle organise la *scène* de leur interaction, en particulier (mais pas exclusivement) au moyen de schèmes configurationnels qui idéalisent les positions des actants impliqués. Ces schèmes résident dans des espaces *abstraits*, et ne doivent pas être confondus avec l'image perçue de la scène (quand elle peut l'être), ni même avec une image mentale qui l'évoque. Le schème n'en reste pas moins topologique, et topologiquement rattachable aux images proprement dites, si besoin est. D'un autre côté, il consiste essentiellement en une sorte de configuration archétype, un squelette sous-jacent à la morphologie de la scène entière, sur lequel opèrent les choix de lexicalisation et de construction syntaxique responsables de la prédication linguistique proprement dite.

¹⁰Ces positions théoriques ont évidemment suscité de très nombreuses critiques, parmi lesquelles nous citerons celles de F. Rastier [1993] :

— cette spatialisation de la grammaire se fait au nom de la spatialisation intégrale du cognitif, de son écrasement dans le schématisme d'une intuition pure ;
 — quel est le statut de ces espaces ? espace des sciences physiques ou du sens commun ? espace transcendantal ou espace d'une psychologie ?
 — la spécificité sémantique des langues se réduirait alors à la spécificité d'un point de vue sur l'un ou l'autre de ces espaces, supposés déjà construits à l'abri de toute imprégnation par le sens linguistique.
 Pourquoi nommer ancrage dans la perception ce qui n'est qu'un corrélat perceptif ou imaginaire de l'activité langagière ?

¹¹Il vaudrait mieux parler ici d'une protosymbolique : on éviterait ainsi de paraître dissoudre une fois de plus la spécificité linguistique dans la généralité du symbolique (serait-il un symbolique enrichi, qui s'oppose, par son recours à un schématisme spatio-temporel généralisé, à la réduction logico-syntaxique du symbolique). Le préfixe "proto-" annonce également une thèse phylogénétique sur le langage,

• Cette spatialisation de la grammaire (qui remplace la forme propositionnelle par le noyau scénique qu'est l'image-schème) n'était qu'amorcée, voire implicite dans les premiers travaux de Fillmore. Elle s'est développée depuis, à travers les travaux de nombreux auteurs. Ils ont ainsi étendu une forme de sémantique *localiste* à toutes les parties du discours, sous la condition d'une *abstraction* des espaces postulés – qui leur conserve toutefois une structure proche de celle d'un espace de l'imagination, en principe rattachable à l'espace perçu. On peut en faire l'une des bases explicatives de l'organisation syntaxique de surface. On aurait alors *motivé* les constructions grammaticales par une organisation topologique cognitive. Le noyau *scénique* (et non plus propositionnel) de la grammaire pointe sur, *code* littéralement, une sorte de perception catégorielle généralisée¹⁰.

• Quel que soit alors le statut conféré à ces images-schèmes (réaliste mentaliste ou descriptif instrumental), ils apparaissent comme des structures médiatrices entre le symbolique et le perceptif. Ils constituent un format intermédiaire entre la diversité infinie des formes perçues, et leur possible catégorisation à l'intérieur de répertoires à la fois discrets et adaptatifs. Les langues naturelles nous en démontrent la possibilité, et nous en offrent des modalités précises. Il convient alors de déterminer plus précisément ce format de structure, dans un cadre topologique censé représenter le type d'abstraction spatiale requise. De ce point de vue, les grammaires cognitives conduisent aux identifications structurales suivantes (Langacker, repris par Petitot) :

- les termes sont des *domaines* localisés dans des *espaces abstraits* ;
- les relations sont des *relations positionnelles* entre ces domaines ;
- les processus sont des *évolutions temporelles* de telles relations ;
- les événements sont des *interactions* entre positions ;
- les rôles sont des *types* d'interaction entre positions.

Naturellement, on se tient là très en deçà d'une sémantique linguistique : plutôt au niveau d'une *protolinguistique*, ou mieux d'une *protosymbolique* ou topologie cognitive très générale¹¹. Il s'agit avant tout de déterminer, comme structures et processus, les schèmes sous-jacents aux identifications ci-dessus.

• On peut alors schématiser mathématiquement ces structures dans le cadre général d'une *morphodynamique* : c'est l'approche de J. Petitot, qui fait confluer la tradition catastrophiste de R. Thom et C. Zeeman avec les théories de la vision de D. Marr et J. Koenderinck. Il n'est évidemment pas question de traiter un pareil "problème" dans toute sa généralité. On

choisit une situation limitée, mais déjà caractéristique : on suppose avoir affaire à une scène perceptive simple, préalablement réduite à la donnée d'une distribution d'objets, eux-mêmes assimilés à des domaines réguliers en mouvement (*blobs*). On se propose alors de modéliser le processus qui conduit de ces données perceptives simplifiées à des catégorisations de relations, de processus et d'événements compatibles avec une prise en charge lexicale simple. La langue n'est donc là qu'à titre d'indice : on cherche en fait les corrélats morphodynamiques des catégorisations qu'elle opère à l'intérieur d'un *continuum* perceptif très simplifié.

Nous ne pouvons entrer ici dans les détails techniques de cette approche, qui sont à vrai dire plutôt complexes. L'idée générale est que de telles configurations de positions dans un espace substrat peuvent être traitées comme des formes pouvant évoluer et se transformer temporellement. Les accidents morphologiques pouvant intervenir au cours de telles déformations schématisent alors des événements d'interaction actantielle. Pour rendre cette idée effective, cependant, on ne peut se contenter des contours déjà présents dans les données scéniques brutes. Pour engendrer la structure de scène proprement dite, on doit également construire une batterie de contours virtuels (que Talmy appelle *fictive*), et plus généralement des lieux *singuliers* dont les types topologiques, d'une structure plus simple que celle de la scène à construire, restent caractéristiques de son état global instantané. Ces techniques font appel, notamment, à des algorithmes de propagation et diffusion d'activité déjà attestés en vision computationnelle¹². On recherche donc en permanence une compatibilité entre le schématisme ainsi modélisé et les structures organisatrices d'une perception visuelle effective.

- Les modèles connexionnistes aident ainsi à la construction d'une *effectivité* pour le schématisme abstrait de la morphodynamique : on définit dans ce cadre des règles d'assignation et de localisation matérielles précises pour les entités modélisées (les contours virtuels), et les transformations formelles considérées (affectant ces contours) sont toutes induites par des processus.

- Le continu et l'espace intervenant dans ces modélisations sont clairement exigés par la problématique de *l'ancrage dans la perception*, qui signifie (selon nous) que l'on ne met pas nécessairement en question la validité d'une catégorialité grammaticale et conceptuelle *discrète*, mais que l'on estime qu'elle ne peut provenir et se maintenir que dans la coarticulation constante au monde perçu à travers une schématisation topologique des interactions spatio-temporelles.

Les représentations, en tant que *Gestalten* mathématisées, sont donc ici des *moments dépendants* d'un espace cognitif, lui-même identifié à une

immédiatement transformable en une thèse sur l'ontogenèse de toute énonciation. Le schématisme et l'imagination s'écrasent alors dans leurs supposées "bases" perceptives (on parle d'ailleurs d'ancrage dans la perception). Des considérations critiques s'imposent ici, mais nous les remettons à plus tard.

¹²On pourra consulter, sur le plan mathématique, l'article de synthèse de L. Alvarez et J. M. Morel [1994].

schématisation topologique de l'espace perceptif. Ces représentations réalisent un isomorphisme entre la forme de l'intériorité et celle de l'extériorité, centré sur la structure des états de choses, exactement comme dans les approches logiques. La grande différence est qu'au lieu de rattacher circulairement la forme logique à elle-même en prétendant la lire en même temps dans la forme grammaticale et la structure du monde, on l'ancre hors d'elle-même dans une "perception", qui n'est à proprement parler qu'une forme pure de l'imagination (compatible cependant avec les meilleures approches computationnelles de la vision).

En termes cognitifs, on reste clairement à l'intérieur d'un *fonctionnalisme dynamique*, où l'effectivité des processus construits concerne une perception toujours solipsiste. Les rétines virtuelles qui tiennent lieu ici de surfaces photosensibles sont en réalité, quand on y regarde de plus près, des espaces déjà phénoménologiquement valués ; ce ne sont pas des espaces physiques authentiques, chevillés à un organisme mobile. Ils supposent déjà réussie une certaine stabilisation spatiale, et s'appuient sur *l'a priori* d'une concordance pure et simple entre saillances phénoménologiques et singularités mathématiques multi-échelles. Bien sûr, cette perception schématisée est remarquablement connectée au noyau propositionnel (ou plutôt scénique) de la grammaire, et *en même temps* à une forme qualitative du monde physique. Elle échappe ainsi aux apories qui minent le fonctionnalisme classique, qui s'appuie quant à lui sur la forme logique. Mais les représentations typiques de cette nouvelle attitude représentationnelle sont d'abord, comme toute représentation, des *systèmes d'écritures* ; elles forment une *idéographie dynamique* au service de la modélisation. L'effectivité construite concerne alors la réactivité propre de ce système d'écritures, les formes minimales de sa cohérence. Elle est, dans son ordre de formalité morphodynamique, l'analogue d'une algorithmique venant déterminer des schémas inférentiels logiques.

Mais objectiver de nouveaux systèmes de signes n'est pas objectiver la *sémiose* toute entière, pas plus qu'étendre nos ressources représentationnelles n'implique *ipso facto* un progrès de la viabilité des systèmes au sein de mondes fluctuants, imprévisibles. C'est en quelque sorte l'erreur centrale de l'IA, et avec elle du fonctionnalisme, d'avoir cru trouver de cette façon les conditions suffisantes de l'autonomie sémiotique ou de la viabilité physique. Certes, les représentations étudiées dans cette section se distinguent fondamentalement de celles du fonctionnalisme classique, dans la mesure où elles sont construites avec le souci de préparer leur rattachement effectif à une perception réelle, impliquant capteurs et mesures dans l'espace physique. Elles ne "disent" pas autre chose, finalement, que l'impossibilité de détacher entièrement le sens linguistique des contraintes qualitatives de l'inscription dans un espace et un temps. Néanmoins, ce sont des représentations dépendantes d'un

schématisme topologique qui déterminent la forme qualitative du monde en même temps que la forme grammaticale qui le symbolise : elles sont donc, comme les représentations logiques, adaptées à la saisie d'un monde où les valeurs fondamentales s'explicitent dans leur forme caractéristique. Mais rien ne nous assure *a priori* qu'elles pourraient prendre sens à l'intérieur des contraintes matérielles de la *viabilité*, c'est-à-dire à l'intérieur d'un espace physique turbulent, où de surcroît les saillances morphologiques, si difficiles à reconstruire, *n'accéderaient à l'existence qu'à travers les valeurs, les prégnances, constitutives de l'habitus du système*. Seule une constitution de représentations qui fasse droit à leur genèse adaptative, dans le cadre d'un développement singulier à chaque fois, aurait quelque chance de répondre *théoriquement* — et peut-être même pratiquement — à cette exigence.

Dans nos termes méréologiques, la mise entre parenthèses du monde, et le procès de constitution des représentations qui en résulte, n'est rien d'autre qu'un détachement simple, un morcellement séparant le substrat et l'environnement, qui méconnaît donc leur essence de moments *inter-dépendants*. Seul un coengendrement du substrat et du monde à partir d'un *être-au-monde* toujours présent pourrait donner à cette codépendance une *effectivité* — d'un tout autre type que les précédentes. Les modèles évoqués dans la section suivante tentent d'en donner une image très rudimentaire : tellement rudimentaire, voire balbutiante, qu'elle ne saurait évidemment se substituer à l'une quelconque des approches précédentes sur leur terrain propre — toute question épistémologique mise à part. Le changement de paradigme est profond, et entraîne une désertion presque complète des domaines de modélisation abordés jusqu'ici.

5. Les approches non représentationnalistes : Vie Artificielle et modélisation connexionniste

Sans prétendre à une quelconque vue d'ensemble, et encore moins à un bilan critique, de ce mouvement scientifique et des nombreux travaux qu'on pourrait lui apparenter (ceux de G. Edelman, par exemple, ou bien ceux de J. Stewart et F. Varela sur le système immunitaire), nous dégagerons quelques traits remarquables de ces problématiques en suivant ici un exposé de Rolf Pfeifer et Paul Verschure [1992]. Sur toutes ces questions, on pourra également se reporter, entre autres, à H. Bersini, P. Bourguine, R. Brooks et P. Maes.

- On renonce ici à modéliser des phénomènes cognitifs de haut niveau, pour le moment inaccessibles à une modélisation "forte", c'est-à-dire une modélisation susceptible d'expliquer ou de fonder *l'autonomie* des systèmes vivants ou artificiels.

On s'intéresse alors à la *construction de boucles sensori-motrices chez des créatures artificielles* très rudimentaires, douées de "pulsions", et non de plans précis ou de connaissances logico-symboliques.

On les place dans des environnements virtuels ou réels, auxquels elles doivent s'adapter selon des processus de renforcement modulés par des schèmes de valeurs innés : c'est la problématique du courant de recherche de la *vie artificielle*.

• Divers auteurs rattachables à ce courant critiquent ce qu'ils appellent la *conception rationaliste* du cognitif, et la méthodologie de modélisation qui en dépend :

¹³Ce type de critiques identifie souvent le "rationnel" à la catégorialité discrète des approches logiques. Une question capitale est de savoir jusqu'à quel point elles se transposent à toute approche rationnelle du cognitif, quelle que soit sa formalité de référence (cf. supra la discussion à la fin de la section 4).

— préidentification d'objets et de situations aux propriétés bien définies ;

— spécification de règles s'appliquant en toute situation aux objets, sur la base exclusive de ces propriétés¹³.

Le comportement du système dépend alors de catégories propres au concepteur, indépendantes de l'interaction historique entre le système et son environnement.

• La modélisation connexionniste, dans l'ensemble, ne s'est pas écartée du schéma "rationaliste" (précodages et présegmentations, préassignation des tâches, pas de notion d'environnement...).

• Or le mouvement, par exemple, ne peut être conçu comme une suite discrète de catégorisations décidées suivant des critères métriques simples. Il faut partir, non pas de situations préspecifiées, mais d'une sorte d'*être-au-monde*.

• Il convient alors de préconiser une nouvelle approche théorique, par exemple en robotique. Elle comporte :

— des *schèmes de valeur* ou de survie, qui spécifient grossièrement, par action de masse à l'intérieur du système, les besoins fondamentaux : bouger, éviter les chocs, se rapprocher de cibles... ;

— des *réflexes* : collision à gauche → tourner à droite ; cible perçue → se rapprocher, etc. ;

— des *répertoires* : moteurs, de détection de collisions et de cibles, ou d'espace libre... ;

— des lois d'apprentissage par *renforcement*, modulées par les schèmes de valeur.

• On définit ainsi un *design* non rationaliste :

— on caractérise un environnement, une niche écologique, et non des situations préidentifiées ;

— on définit des schèmes de valeur pour des agents munis de dispositifs sensori-moteurs, et non des règles symboliques pour des agents planificateurs ;

— on laisse l'agent trouver ses parcours singuliers dans son environnement, sans prédéfinir une ontologie formelle responsable de ses comportements.

- L'environnement est donc conçu en termes "physicalistes", mais cependant non "ontologiques" : cela suppose que la "physique" sous-jacente, qui détermine les interactions, est infiniment plus riche que les ontologies, ou catégorisations, toujours provisoires, construites par les systèmes.

- De ce fait, les systèmes ne peuvent être exactement assignés à la poursuite permanente et exclusive d'un "objectif" facilement désignable : les principes de leur adaptation impliquent en effet de l'aléatoire dans leurs parcours. Leur utilité individuelle ne pourrait donc être définie qu'en termes statistiques : et seul le recours à des *populations* de systèmes permet alors d'atteindre à l'efficacité. Le nombre n'est d'ailleurs pas le seul facteur en jeu : la *différenciation* des agents intervient également, et les bases de la viabilité individuelle sont dans ce cas strictement *collectives* (cf. les modèles d'intelligence en *essaim*).

- De même, il apparaît que la viabilité recherchée ne pourrait résulter d'une adaptation purement individuelle : seule une approche darwinienne permettrait peut-être d'y accéder, en opérant par variation et sélection sur des populations importantes, à travers de nombreuses générations.

L'arsenal formel de la modélisation connexionniste, et *a fortiori* des algorithmes de renforcement ou des algorithmes génétiques investis dans ce type de travaux, diffère quelque peu de celui qui a eu jusqu'ici la faveur des modélisateurs de processus de haut niveau (généralement la rétropropagation et ses variantes). Mais l'essentiel est ailleurs dans notre perspective. Il tient dans la tentative de construire, et toujours adapter, des représentations, à la faveur d'interactions évoluant par paliers d'organisations, sur la base d'une mise en situation physique, et à partir de schèmes de valeurs établissant implicitement des gradients d'intérêt dans l'environnement. Il n'est donc plus question de commande ou de contrôle, mais de circulation à l'intérieur de domaines de viabilité (cf. les travaux de P. Bourguin et les ouvrages de F. Varela). Dans notre terminologie méréologique, de telles représentations sont évidemment des moments très dépendants : d'une histoire et d'une architecture, d'un environnement et de schèmes de valeurs... Plutôt que des représentations, ce sont des *potentiels* représentationnels, des potentiels de *construction de situations*, caricaturés en simples représentations par l'observateur, qui les réduit à leurs actualisations les plus saillantes, au lieu de tourner son regard vers le mécanisme générateur qui les fait advenir à chaque fois. Des représentations donc, si l'on y tient, mais *ancrées dans une histoire*

¹⁴Bien loin d'être gratuite, cette image procède en réalité d'une philosophie qui affirme l'inhérence d'une genèse, et plus profondément d'une histoire, pour toute autonomie. Ici l'épistémologie peut bien se faire dialectique, puisque la modélisation éclaire et légitime une certaine circularité du jeu conceptuel. Qu'est-ce que l'autonomie, en effet, du point de vue d'un observateur, sinon l'impression d'une auto-production de son histoire par le Soi d'un être vivant ? On est alors contraint d'admettre l'historicité d'un tel Soi, puisqu'il doit être dans une certaine mesure la source de ce qui lui arrive. De même pour la Conscience Primaire, qui est le nom que donne G. Edelman au couple perception-mémoire, lorsque son ouverture phénoménologique se limite au cadre d'un Présent. Dans sa conception, du moins telle que nous souhaitons la reformuler en empruntant pour cela au discours de A. Pichot, la perception (qu'il ne faut pas confondre avec la simple sensation, qui la précède de très loin dans l'évolution) est une forme de traduction phénoménologique, pour un animal, de la disjonction entre le Soi et le Non-Soi. Cette disjonction, au lieu de se traduire uniquement par des réactions d'évitement ou d'approche immédiates,

d'interactions avec un environnement. La signification paradoxale de cette dépendance historique est précisément qu'elle offre l'image, encore bien simple, d'une genèse de l'autonomie¹⁴.

Quelle est alors la portée effective de ces recherches d'épistémologie expérimentale ? Peut-être à terme une révolution de la robotique, mais plus authentiquement à l'heure présente, une indispensable réflexion sur les conditions les plus générales de la perception, de l'action, et de la coordination sémiotique en éthologie. Une perception d'abord, qui sans être nécessairement une perception par objets ou une perception actantielle, procéderait à une confrontation et à une fusion partielle entre des activités sensori-motrices immédiates, et des anticipations et remémorations qui les codéterminent et les évaluent tout à la fois. Une action, qui sans être aucunement planifiée, serait cependant finalisée par le couplage entre les répertoires de valeurs et les anticipations de cette perception. Et une sémiotique enfin, probablement très primitive dans ses formes, ses contenus, ou ses adresses (c'est-à-dire sans mécanismes hiérarchiques, sans contenu scénique diversifié, sans anticipation du destinataire), mais dont les valeurs se construiraient néanmoins dans l'interaction collective et au fil de l'évolution.

Toutefois, il nous paraît difficile d'imaginer que de telles recherches puissent se développer d'une façon purement empirique¹⁵, comme c'est principalement le cas jusqu'à présent : elles rencontreront tôt ou tard la question de comprendre les *a priori formels* qui les contraignent à leur insu. Elles devront donc à leur tour tenter de schématiser mathématiquement leurs objets : évolution, développement, adaptation, viabilité. Différentes propositions existent déjà en ce sens (v. chez Aubin, Bourguine *et al.*). Soulignons qu'une telle schématisation, pour être valide,

sanctionnées par des évaluations pulsionnelles instantanées, est au contraire perceptivement traitée. Autrement dit, le thème premier de toute perception est la mise en phénomène, pour l'animal, de son potentiel de dis-jonction d'avec l'environnement. L'espace propre à une espèce, qui est perçu au présent par chacun de ses membres, se creuse de tous les mouvements que l'on a pu — ou pourrait — faire ou sentir. La mémoire entre dans le phénomène, et apparaît à l'animal à l'intérieur de son propre champ perceptif (mais sans prendre pour autant la qualité phénoménologique d'un Passé). Dire que l'animal perçoit signifie donc qu'il perçoit aussi sa propre mémoire, qui est indissolublement mémoire de sensations et mémoire de prégnances. Le dosage de Soi et de Non-Soi dans la construction perceptive est donc une affaire critique : il faut que le Soi se retrouve dans ce qui est perçu, tout en se confrontant au "format" perpétuellement changeant que l'appareil perceptif donne au Non-Soi. Et cette étrange coalition doit en même temps survivre. La "solution" réside là encore dans les conditions d'une genèse, dont il faut imaginer, selon Edelman, les différentes modalités temporelles incarnées dans des parties distinctes du cerveau de l'animal : catégories de mémoires couplées aux valeurs du Soi, d'une part, et répertoires

devra toujours porter sur les structures d'une *interaction* entre le système et son environnement, de préférence en les installant tous deux sur un espace substrat *commun* : faute de quoi, on cède à la tentation de déterminer une fois pour toutes les paramètres de l'interaction, au lieu de les déduire du processus d'adaptation lui-même (cf. [Deffuant & Fuhs, 1993]).

Conclusion

Nous avons d'abord constaté et motivé l'essor du Continu et de l'Espace — ou plutôt de diverses acceptions de l'espace — dans les sciences cognitives. Nous avons ensuite effectué un parcours à travers les différents types de représentations qui illustrent cet essor dans le cadre de la modélisation connexionniste, en les considérant à chaque fois comme les indices de diverses *attitudes représentationnelles*. Nous avons en particulier présenté les représentations comme des *moments dépendants*, en essayant de dégager leur mode de participation "méréologique" aux ontologies, implicites ou explicites, des travaux examinés.

Si le progrès des connaissances consiste en la reconnaissance progressive de toutes les dépendances, celui des sciences repose sur la décision temporaire de les ignorer presque toutes. Qu'en a-t-il été ici ? Nous avons par exemple feint d'ignorer nos dépendances vis-à-vis du discret : non seulement celui, peut-être contingent, de l'informatique actuelle, mais encore celui, plus fondamental, qui sous-tend toute numération, et à partir de là toute mesure, tout codage, toute *explicitation* du continu. Nous avons ainsi redéfini le statut de la *description par règles* des régularités comportementales, perceptives, ou grammaticales : nous lui avons assigné un statut, subjectivement et méthodologiquement nécessaire, mais théoriquement second, de reconstruction qualitative facilitant la compréhension, et susceptible de préparer ainsi une intervention plus judicieuse du topologique et du dynamique dans la complexification des modèles. Ces modèles, à leur tour, sont présumés figurer, dans leur matérialité, le pourquoi et le comment de l'efficacité des règles en nous — de quelles façons effectives nous serions concernés par le format cognitif des règles.

Parallèlement, nous sommes amené à *transformer* les modalités de notre dépendance vis-à-vis de l'espace, pour procéder — c'est la règle du jeu souvent adoptée ici — à la mise en espace de toute conceptualité. Ce faisant, nous schématisons les dépendances complexes de notre univers

sensori-moteurs d'autre part. Leur confrontation produit la genèse permanente du Soi, à travers une recatégorisation, ou réorganisation, qui en est comme l'effet de bord. Mais cette genèse ne peut elle-même (re)commencer que sur le fond d'une histoire du Soi déjà donnée, que le Présent de l'animal réactualise dans la construction de chaque scène qui vient occuper son champ. C'est là, très brutalement et vaguement résumé, le sens du titre «The Remembered Present» [Edelman, 1989]. Cf. également J. Stewart [dir., 1993], et pour une reprise phénoménologique de ces thèmes d'épistémologie de la biologie, l'ouvrage de A. Pichot [1991].

¹⁵*Cette empirie est souvent celle des environnements virtuels de l'informatique.*

phénoménologique en les représentant par des méréologies naïves, extensives et morcelables, des *parties* d'espace. Comment cela est-il possible ? Par la montée en complexité et en abstraction des espaces et des configurations engagés, c'est-à-dire par leur détachement croissant, mais toujours *construit*, du cadre assigné par notre corps. Autrement dit, c'est en complexifiant l'espace que nous "simplifions" les méréologies. Qu'est-ce qu'un contour, par exemple, ou bien une fonction grammaticale ? Réflexion faite, nous ne pouvons pas les traiter comme des parties morcelables d'un espace perceptif ou sémantique directement intuitionnables. Mais, convenablement schématisés, le contour ou la relation deviennent des familles de lieux géométriques, c'est-à-dire des parties simples d'espaces fonctionnels abstraits. Cette opération de schématisation respecte donc, en un certain sens, les exigences, d'ailleurs conflictuelles, de *notre* spatialité naïve : disposer d'un espace absolument connexe et cohérent (mais ici il s'agit d'un espace mathématique), et en même temps le peupler d'entités aussi morcelables et indépendantes que possible. Mais cette sauvegarde toute relative du sens commun n'est nullement un objectif de l'idéalisation scientifique, comme elle est censée l'être dans le cas tout différent du cognitivisme fodorien ; elle n'a pour ainsi dire aucune dignité épistémologique relativement au phénomène modélisé. Elle pointe cependant sur une dépendance subjective, dont le statut reste toujours à décider : transcendantal ? anthropologique universel ? culturel ? En d'autres termes, l'espace du sens commun entre d'abord en conflit avec celui de la description phénoménologique réfléchie ; cette description peut être corrélée à son tour à une schématisation mathématique, qui bouleverse cependant le type des espaces impliqués ; et tout au fond de notre compréhension *subjective* de cette schématisation, certaines des exigences de la compréhension selon le sens commun se font à nouveau sentir — alors qu'en théorie elles sont doublement délégitimées.

Encore une fois, nous voudrions souligner la disparité de statut des espaces évoqués (des espaces qui peuvent être, dans l'ordre alphabétique, cognitifs, perceptifs, physiques, sémantiques...), en même temps que celle des mouvements de complexification et de réduction qui les impliquent spécifiquement. En un sens, nous avons seulement voulu préciser, modèles à l'appui, de quelles façons diverses le recours à l'espace et au continu n'était pas le fait exclusif d'épistémologies éliminativistes, qui en appelleraient directement à la construction d'une physique des substrats neuraux, sans se préoccuper de déterminer, *au niveau d'abstraction requis*, leurs finalités potentielles. La définition ou la découverte de ces finalités peut d'ailleurs assigner à la clôture opérationnelle des systèmes cognitifs, tout comme à l'interaction environnementale, une importance très variable. Mais dans tous les cas, le maintien des *attitudes*

représentationnelles en sciences cognitives, pour autant qu'on les juge souhaitables, dépend, avec une évidence croissante, d'une détermination spatiale et temporelle *abstraite* des finalités cognitives.

CNRS—*Laforia* (Paris)

Remerciements à Paul Bourguine, Jean Petitot, François Rastier, Jean-Michel Salanskis, John Stewart, et Bernard Victorri. Ce texte est simplement l'écho de certaines discussions, aussi animées que récurrentes.

Références

ALVAREZ (L.) & MOREL (J. M.)

1994, "Formalization and Computational Aspects of Image Analysis", *Acta Numerica*, p. 1-63.

ANDLER (D.) & PETTITOT (J.) & VISETTI (Y. M.)

1991, "Dynamical Systems, Connectionism, and Linguistics", *Actes du Colloque interdisciplinaire sur la compositionnalité dans la cognition et les réseaux neuronaux*, Royaumont.

BALDI (P.) & HORNIK (K.)

1989, "Neural Networks and Principal Component Analysis : Learning from Examples without Local Minima", *Neural Networks*, vol. 2, p. 53-58.

BERSINI (H.)

1991, "ANIMAT'S I", p. 456-465, *First European Conference on Artificial Life*.

BERSINI (H.) & VARELA (F.)

1990, "Immune Networks and Adaptive Control", *Parallel Processing in Nature, Proceedings of the Dortmund Conference*, Berlin, Springer Verlag.

BOCHEREAU (L.) & BOURGINE (P.)

1989, "Implémentation et extraction de règles sur un réseau neuromimétique", p. 125-142, *Actes de Neuro-Nîmes 89*, EC2.

BOURGINE (P.)

1993, "Viability and Pleasure Satisfaction Principle of Autonomous Systems", *Actes du Colloque Imagina 93*, Monte-Carlo, INA.

BROOKS (R.)

1991a, "Intelligence without Representation", *Artificial Intelligence*, n°47, p. 139-159.

1991b, "Intelligence without Reason", *IJCAI*.

CHALMERS (D. J.)

1990, "Syntactic Transformations on Distributed Representations", *Connection Science*, n° 2, p. 53-62.

CRUCIANU (M.)

1993, "Finding Structured Representations in Connection Networks", *Rapport technique du LIMSI*, Orsay.

DEFFUANT (G.) & FUHS (T.)

1993, "Vers une généralisation de la théorie de l'apprentissage : l'exemple des fibres adaptatives", p. 506-515, *Actes du 2^e Congrès européen de Systémique*, vol. 3.

EDELMAN (G. M.)

1989, *The Remembered Present : A Biological Theory of Consciousness*, Basic Books.

1992, *Biologie de la conscience*, Paris, éd. Odile Jacob (trad. de *Bright Air, Brilliant Fire : On the Matter of Mind*, Basic Books).

EHRlich (M. F.) & TARDIEU (H.) & CAVAZZA (M.), dir.

1993, *Les Modèles mentaux : approche cognitive des représentations*, Paris, Masson.

ELMAN (J.)

1991, "Distributed Representations, Simple Recurrent Networks, and Grammatical Structure", *Machine Learning*, n°7, p. 195-225.

1993, "Learning and Development in Neural Networks : The Importance of Starting Small", *Cognition*, n°48, p. 71-99.

FODOR (J. A.) & PYLYSHYN (Z. W.)

1988, "Connectionism and Cognitive Architecture : A Critical Analysis", *Cognition*, n° 28, p. 3-71.

[Gallinari *et al.*]

GALLINARI (P.) & THIRIA (S.) & BADRAN (F.) & FOGELMAN-SOULIE (F.)

1991, "On the Relations between Discriminant Analysis and Multilayer Perceptrons", *Neural Networks*, n°4, p. 349-360.

HARRIS (C.)

1990, "Connectionism and Cognitive Linguistics", *Connection Science*, vol. 2, n°s 1-2, p. 7-33.

HARTHONG (J.)

1992, "Le Continu ou le Discret, un problème indécidable", p. 350-365, *Le Labyrinthe du Continu*, ouvrage collectif sous la direction de H. Sinaceur et J.- M. Salanskis, Paris, Springer Verlag.

LANGACKER (R. W.)

1987, *Foundations of Cognitive Grammar*, vol. 1, Stanford University Press.

1991a, *Foundations of Cognitive Grammar*, vol. 2, Stanford University Press.

1991b, "An Emergent View of Constituents in Cognitive Grammar", p. 70-77, *Actes du colloque interdisciplinaire sur la compositionnalité dans la cognition et les réseaux neuronaux*, Royaumont.

MAES (P.), ed.

1991, *Designing Autonomous Agents : Theory and Practice from Biology to Engineering and Back*, MIT Press.

McCLELLAND (J. L.) & RUMELHART (D. E.), eds.

1986, *Parallel Distributed Processing : The Microstructure of Cognition*, vol. 2, MIT Press.

MIKKULAINEN (R.) & DYER (M.)

1991, "Natural Language Processing with Modular PDP Networks and Distributed Lexicon", *Cognitive Science*, n°15, p. 343-399.

MINSKY (M.) & PAPERT (S.)

1969, *Perceptrons*, Cambridge (MA), MIT Press.

PETTITOT (J.)

1991a, "Syntaxe topologique et grammaire cognitive", *Langages*, n°103, p. 97-128.

1991b, "Morphodynamical Models and Topological Syntax", *Actes du Colloque interdisciplinaire sur la compositionnalité dans la cognition et les réseaux neuronaux*, Royaumont.

1992a, "Diffusion de contours et constituance perceptive", *Actes du 2^e Colloque interdisciplinaire sur la compositionnalité dans la cognition et les réseaux neuronaux*, Royaumont.

1992b, *Physique du Sens*, Paris, Éditions du CNRS.

1993, "Phénoménologie naturalisée : la fonction cognitive du synthétique a priori", in SALANSKIS (J.- M.), dir.

PFEIFER (R.) & VERSCHURE (P.)

1991, "Distributed Adaptive Control : A Paradigm for Designing Autonomous Agents", p. 21-31, *First European Conference on Artificial Life*.

1992, "Symbols, Patterns and Behavior", *Connection Science*, vol. 4, n°s 3-4, p. 313-325.

PICHOT (A.)

1991, *Petite phénoménologie de la connaissance*, Paris, Aubier.

POLLACK (J. B.)

1990, "Recursive Distributed Representations", *Artificial Intelligence*, n°46, 1-2, p. 77-105.

RASTIER (F.)

1993, "La Sémantique cognitive et l'espace", p. 173-185, *Images et langages : multimodalité et modélisation cognitive*, Actes du Colloque Interdisciplinaire du CNRS, avril.

REEKE (G.) & SPORNS (O.) & EDELMAN (G. M.)

1989, "Synthetic Neural Modelling : Comparisons of Population and Connectionist Approaches", p. 113-139, *Connectionism in Perspective*, R. Pfeifer, Z. Schreier, F. Fogelman-Soulié et al. eds, North-Holland.

1990, "Synthetic Neural Modelling : The "Darwin" Series of Recognition Automata", p. 1498-1530, *Proceedings of the IEEE*, 78, 9.

REGIER (T.)

1988, "Recognizing Image-Schemas Using Programmable Networks", *Proceedings of the 1988 Connectionist Models Summer School*, M. Kaufman.

1992, *The Acquisition of Lexical Semantics for Spatial Terms : A Connectionist Model of Perceptual Categorization*, PHD, Berkeley, TR-92-062.

ROSENTHAL (V.)

1993, "Cognition, vie et... temps", p. 175-207 in STEWART (J.), dir.

RUMELHART (D. E.) & McCLELLAND (J.), eds.

1986, *Parallel Distributed Processing : The Microstructure of Cognition*, vol. 1, MIT Press.

SALANSKIS (J.- M.)

1991, *L'Herméneutique formelle : l'infini, le continu, l'espace*, Éditions du CNRS.

1992, "Modèles du Continu dans les sciences", *Intellectica*, n°s 13-14, p. 45-78.

1993, "Continuum, Cognition, and Linguistics", à paraître dans le volume collectif des Actes de la table ronde sur le *Continu en sémantique linguistique* organisée par C. Fuchs et B. Victorri (juin 92, Caen).

SALANSKIS (J.- M.), dir.

1993, "Philosophies et sciences cognitives", numéro spécial d'*Intellectica*, n°17.

SINACEUR (H.) & SALANSKIS (J.- M.), dir.

1992, *Le Labyrinthe du Continu*, Paris, Springer Verlag.

SMOLENSKY (P.)

1988, "«On the Proper Treatment of Connectionism», suivi de la réponse de l'auteur aux commentaires", *The Behavioral and Brain Sciences*, n°11, p. 1-74.

1989, "Connectionism and Constituent Structure", p. 3-24, *Connectionism in Perspective*, R. Peifer, Z. Schreter, F. Fogelman-Soulié *et al.* eds, North-Holland.

[Smolensky *et al.*]

SMOLENSKY (P.) & LEGENDRE (G.) & MIYATA (Y.)

1992, *Principles for an Integrated Connectionist/Symbolic Theory of Higher Cognition*, Technical Report, University of Colorado at Boulder, Department of Computer Science.

ST JOHN (M.- F.)

1992, "The Story Gestalt : A Model of Knowledge-Intensive Processes in Text Comprehension", *Cognitive Science*, n°16, p. 271-306.

ST JOHN (M.- F.) & McCLELLAND (J.- L.)

1990, "Learning and Applying Contextual Constraints in Sentence Comprehension", *Artificial Intelligence*, 46, 1-2, p. 217-257.

STEWART (J.)

1994, "Un Système cognitif sans neurones : les capacités d'adaptation, d'apprentissage et de mémoire du système immunitaire", à paraître in *Intellectica*, n°18.

STEWART (J.), dir.

1993, "Biologie et cognition", *Intellectica*, n°16.

TALMY (L.)

1983, "How Language Structures Space", in *Spatial Orientation : Theory, Research, and Applications*, H. Pick and L. Acredolo eds., Plenum Press.

THOM (R.)

1977 (1972), *Stabilité structurelle et Morphogenèse*, New York, Benjamin, 2^e éd., InterÉditions.

1980 (1974), *Modèles mathématiques de la Morphogenèse*, Paris, UGE (Nouvelle éd. Ch. Bourgois).

TOURETZKY (D.)

1988, "Connectionism and PP Attachment", p. 325-332, *Proceedings of 1988 Connectionist Models Summer School*, M. Kaufmann.

Van GELDER (T.)

1990, "Compositionality : A Connectionist Variation on a Classical Theme", *Cognitive Science*, n°14, p. 355-384.

VARELA (F.)

1989, *Autonomie et connaissance : essai sur le vivant*, Paris, Le Seuil.

VICTORRI (B.)

1988, "Modéliser la polysémie", *T. A. Informations*, n°s 1-2, p. 21-42.

1994, "La Construction dynamique du sens", article à paraître dans un ouvrage collectif en l'honneur de R. Thom.

VICTORRI (B.) & FUCHS (C.)

1992, "Construire un espace sémantique pour représenter la polysémie d'un marqueur grammatical : l'exemple de «encore»", *Linguisticae Investigationes*, p. 125-153.

VICTORRI (B.) & RAYSZ (J.- P.) & KONFE (A.)

1989, "Un Modèle connexionniste de la polysémie", p. 97-108, *Actes de Neuro-Nîmes 89*, EC2.

VISETTI (Y. M.)

1990, "Modèles connexionnistes et représentations structurées", *Intellectica*, n°s 9-10, p. 167-212.

Compositionality and Dynamics in Neural Network Representations¹

Tim van Gelder

1. Introduction

This paper considers a problem that arises at the intersection of three very broad ideas. They are (1) that higher cognition needs compositionally structured representations ; (2) that cognition is essentially a dynamic phenomenon ; and (3) that neural networks provide the best modeling environment for the study of cognition. Each of these ideas is quite appealing, though for independent reasons. If they are all true, then, together, they pose an inevitable problem : from the perspective of dynamics, can we understand how neural networks can handle compositionally structured representations ? That is, there must be, within neural network modeling, a thoroughly dynamical way of implementing compositionally structured representations. What will these representations look like ?

The claim that higher cognition demands compositionally structured representations is a key ingredient of mainstream orthodoxy in cognitive science. It dominates the practice of most traditional cognitive modeling, and has been vigorously defended by practitioners and philosophers alike. The basic idea can be crudely expressed by saying that mental representations must be like *sentences* ; that is, they must be complex entities that are systematically constructed out of tokens of a limited set of basic types. A variety of arguments have been put forward in favor of the view that higher cognition requires compositional representations, but the most plausible seems to be just that we have no good ideas about how higher cognitive processes such as reasoning, planning, and language comprehension and production might be conducted except on the basis of the manipulation of compositionally structured representations. Such processes appear to depend on large amounts of topical and background knowledge about the world, at least in any interesting applications. How can that knowledge be stored, retrieved and deployed, except in the form of sentence-like representations ?

¹A version of this paper was read at the Conference «Naturalistic Approaches to Representation and Meaning», Zentrum für interdisziplinäre Forschung (ZiF), Universität Bielefeld, 1993. Portions were presented at the conferences COMPCOG I (1991) and II (1992), Abbaye de Royaumont, France, and at the A. A. A. I. Spring Symposium Series, Stanford University, 1991. Section 2 contains joint work with Robert Port ; a version of these ideas is forthcoming in [van Gelder & Port, in press].

This argument for the dependence of higher cognitive processes on compositional representations is often extended to the conclusion that cognitive systems must have a so-called "Classical" cognitive architecture (e.g., [Pylyshyn, 1984]). But this extension can be resisted, as long as we can find some way to incorporate compositionally structured representations in *non-Classical* cognitive systems. This is why solving the problem of the intersection of compositionality, dynamics, and neural networks is an important exercise.

The second broad idea is that cognition is essentially a *dynamic* phenomenon. This has two sides. The first is that cognition is fundamentally a matter of cognitive processing, and that cognitive processes are natural processes occurring in real biological systems. As such, they take place in *real time*, where this means something more than simply that they happen "fast enough to keep up with what is going on in the environment". It means, also, that the cognitive processes themselves are changes taking place in time, which is a continuous quantity best modeled by means of the real numbers.

This rather basic point becomes interesting when we note that most traditional computational models of cognition standardly abstract away from the temporal nature of cognitive processes. The fine temporal structure of real cognitive processing is seen as a mere implementation detail ; computational models allegedly map the abstract information processing operations that the cognitive system must go through in performing a given task. Insofar as such models incorporate a notion of time at all, time is merely a matter of order or sequence of operations, and so is adequately modeled by means of the integers, which comprise the simplest and most well known ordering. Consequently, there is a whole dimension of modeling adequacy — tracking the fine temporal detail of real cognitive processes — that virtually all traditional computational models entirely ignore.

The inherently temporal nature of cognitive processing suggests that providing adequate scientific descriptions will require a mathematical and conceptual framework capable of dealing with continuous change in real time. The obvious candidate here is traditional dynamical modeling using differential equations, and its modern extension, dynamical systems theory. The second side to the idea that cognition is essentially a dynamic phenomenon, then, is that it is *dynamics* (broadly speaking) that provides the right mathematical and conceptual tools for understanding cognitive processes. Cognitive science has always been a diverse enterprise, and one of the most consistently active sub-branches, if not the most prominent, has been that of dynamical modeling. Pioneers such as [Ashby, 1952], [Grossberg, 1988] and [Thom, 1975] have been applying dynamics to the study of cognitive processes for decades. In recent years, following the

famous explosion of theory and applications of nonlinear dynamical systems theory in the 70s and 80s, rapidly increasing numbers of cognitive scientists have been applying dynamical techniques and concepts across a wide range of areas. Indeed, if connectionism was the most dramatic theoretical revolution of the 1980s, it appears that dynamics is the connectionism of the 1990s [Port & van Gelder, eds., forthcoming, 1995].

The third broad idea is that neural networks provide the most appropriate medium for the construction of models of cognitive processes. In the light of recent debates, this idea needs little elaboration (see, e.g. [Smolensky, 1988]), and so I will pass directly to the implications accepting these ideas.

I am not presupposing that these ideas are true ; merely that they are broad empirical hypotheses which, given our current state of knowledge, look attractive quite independently of one another. Now, one way of evaluating their acceptability is to test their compatibility with each other. If they turn out to be incompatible, then presumably one or more must be rejected. The problem, in short, is to reconcile compositionality, dynamics and neural networks.

Reconciling dynamics and neural networks is no great achievement. Neural networks standardly are dynamical systems governed by differential equations, and increasingly, dynamical systems theory is the mathematical framework of choice for describing their behavior. The naturalness of the fit between neural networks and dynamics has been exemplified in Grossberg's models of various aspects of cognition. Reconciling neural networks and compositionally structured mental representations is also not particularly problematic. Plenty of connectionist models of language processing have constructed complex representations of entities such as sentences by combining, by one means or another, patterns corresponding to the constituents of those entities. The infamous Fodor & Pylyshyn's criticism of connectionism [Fodor & Pylyshyn, 1988] was not that connectionists could *not* incorporate compositional representations in their models ; rather, it was that they *can*, and in order to be adequate to cognition they *must*, but in doing so their models become mere implementations of the Classical approach. However, subsequent work has demonstrated beyond any question that connectionist models can utilize compositionally structured representations *without* amounting to mere implementations (see, e.g. [van Gelder, 1990]). How this is possible turns on the analysis of the concept of compositionality, an issue addressed in more detail below.

The trouble is that there is very little overlap between that neural network research which is truly dynamical, and that research which handles compositional representations. Those neural network researchers who bring the concepts and tools of dynamics to bear on the study of cognition with any degree of sophistication are not focusing on those

problems of higher cognition which appear to demand sentence-like representations, and *vice versa*. The real difficulty, then, appears to be reconciling compositionality and dynamics within a neural network framework. In a way, this should not be surprising. Dynamics offers a vast conceptual and mathematical armory, but concepts such as *representation* and *syntactic structure* are not part of it. What we are asking, in effect, is that some central concepts of one large research framework — that of classical computationalism — be somehow reconstructed within another very different framework.

In the final part of this paper I will briefly discuss one broad approach to handling this problem that has been under independent development by the French philosopher Jean Petitot and the American connectionist linguist Robert Port (and perhaps others). The compositional representations that are constructed within this approach, however, are radically different from the more typical examples found in orthodox cognitive science. Consequently, before describing this method, which can be termed “attractor chaining”, I will explore the concept of compositionality itself, and highlight some of the very different ways in which compositionality can be implemented.

2. Kinds of Compositionality

Consider some of the obvious differences between a printed sentence and a spoken utterance of the very same sentence. Both contain the same words, but in the printed sentence they are static ink configurations, while in the utterance they are temporally extended sound patterns. In the printed case, words are combined by juxtaposition in space ; in the utterance, by juxtaposition in time. The printed words are combined in a very discrete way — there is, quite literally, space between each — and all occurrences of a given word are effectively identical, whatever words happen to surround them. In the utterance, by contrast, juxtaposition is not discrete and context free ; words flow into each other, and their physical shape is affected by their neighbors.

Printed and uttered tokens of the very same sentence, then, are physically *constructed* or *built up* in quite different ways. One way to put this is that the two tokens exhibit very different kinds of *compositionality*, despite being syntactically and semantically identical. (Note that this use of the term “compositionality” is not to be confused with another very common use, in which the term is used to refer to the kind of situation in which basic constituents make approximately the same semantic contribution in every context in which they appear.)

This point has two very important consequences. First, compositionality itself comes in at least two, and possibly many, different

kinds. Second, the *concept* of compositionality can be a subject of study, quite independently of syntax and semantics. Whereas syntax and semantics focuses on a language in the abstract, the study of compositionality focuses on particular, concrete *implementations*. How many kinds of compositionality are there ? What are the fundamental issues ? And what kinds of compositionality are typically found in mainstream artificial intelligence, in neural network research, and in natural cognitive systems ?

A representation is compositionally structured when, roughly, it is systematically constructed out of tokens of a limited set of basic compoundable units. Technically, it is more appropriate to say that a representation is compositional if it stands in certain *constituency relations* ; that is, there is :

- (a) some finite set of *primitive* types, realizable by actual physical tokens ;
- (b) a possibly unbounded number of compound types, realizable by actual physical tokens ; and,
- (c) a set of abstract constituency relations defined over these primitive and compound types.

Any particular token of a representation is compositional just in case, by virtue of belonging to a certain type, it stands in appropriate constituency relations, i.e., can be said to have constituents. (For more detailed discussion of this approach to compositionality, see [van Gelder, 1990].)

The point of specifying the concept of compositionality in this abstract way is to distance us as much as possible from the standard examples of compositionality, paradigms of which are symbolic structures such as printed sentences and LISP expressions. The printed sentence and the spoken utterance of the same sentence are both compositional representations, since they are both tokens of the same type, and that type stands in abstract constituency relations ; however, they realize that compositionality in very different ways. To understand *syntax* is to understand the abstract hierarchical structure of constituency relations among types ; to understand *compositionality* is to understand the various different ways in which concrete *tokens* actually realize those types and their constituency relations.

There are at least six key issues in understanding compositionality. It is possible to think of these metaphorically as the basic *dimensions* of an abstract space of possible kinds of compositionality. These six issues are :

1. Static vs. Dynamic
2. Digital vs. Analog

3. Arbitrary vs. Non-Arbitrary
4. Mode of Combination : Concatenative vs. Non-Concatenative
5. Static vs. Temporal Combination
6. Syntactic Conformity.

Ideally, these dimensions carve up the space of possible kinds of compositionality in the most theoretically revealing way ; to extend the metaphor, they should be something like the “principal components” of that space.

The first three issues concern properties of the tokens of the primitive types, which I will call *symboloids*, in order to emphasize that they include, but are not restricted to, standard symbols such as LISP atoms or printed words. The second three concern the way in which these basic tokens are actually combined in order to form tokens of complex types.

2. 1. Properties of Symboloids

Consider an individual word token from a particular printed sentence of English. It is a paradigm example of *symbol* (i.e., a particular kind of symboloid). Some of its obvious properties include the fact that it sits still on the page for a long time ; that it is straightforward to recognize the symbol *type* to which it belongs ; and that it has *meaning* or semantic significance, yet its physical structure bears no interesting relationship to that meaning. These relatively obvious properties correspond to the first three dimensions of compositionality.

It is worth pausing to mention that it can often be quite difficult to pin down the primitive types of a given compositional scheme of representation. For example, linguists, AI theorists and high school teachers all usually assume that individual words are the primitives of English. This is satisfactory as a first approximation. But Bolinger among others has frequently pointed out how difficult it is to nail down the actual list of the particular items that have just the right properties to count as the primitives [Bolinger, 1975]. Natural language is not just a set of morphemes plus rules of syntax. Speakers use and *know* linguistic fragments that come in many sizes : from submorphemic ideophones to words, idioms, and Bolinger’s “collocations”, as well as *clichés* and even entire sentences and paragraphs of boilerplate (as in genres like wills and academic recommendation letters). The nature of many of these mysteriously constrained yet flexible lexicalized *units* lies well beyond the grasp of current representational schemes employed in linguistic theory.

Nevertheless, in what follows I will be proceeding on the assumption that the basic types have been already adequately identified for any compositional scheme under consideration.

(a) Static vs. Dynamic

Any representation, and any symboloid, is a concrete physical item (an ink mark, a sound wave, a collection of electronic bits, a pattern of neural firing, etc.). It counts as a token of a particular type in virtue of some kind of distinctive physical configuration or *shape*. We can think of this distinctive configuration as a matter of change along key physical aspects of the representation. Thus, a printed token of the word *cat* exhibits change in the way ink is placed on the page, change that would be registered by a scanner as it moved along the page. It is precisely the details of this spatial change which makes it different than a token of the word *dog*. Notice that, by contrast, an instance of *cat* typically exhibits no significant change in physical configuration from one moment of time to the next. This is what I mean by saying that it "sits still", or is a static token.

Compare this with an uttered token of the same word. In this case, it is variation in frequency and amplitude of a sound wave over time which determines its type-identity ; change over space is irrelevant. This is an example of a dynamic primitive. The key feature of a dynamic symboloid or representation is that change over time is essential to type-identity. It follows that, in order to determine to which type the entity belongs, you have to wait long enough for its distinctive change to unfold. In the case of a static symboloid or representation, by contrast, you have (in principle, at least) enough information at any instant to determine type-identity.

Printed English words, then, are static symboloids, while spoken words are dynamic symboloids. One way to summarize the difference is to say that dynamic symboloids, but not static symboloids, happen in time.

(b) Digital vs. Analog

A key feature of most standard compositional schemes of representation is that they are digital. By this I mean that it is possible to produce symboloids in the scheme, and determine the type identity of any given symboloid with complete and unambiguous success. Another way to put this is that the symboloids are such that the reading and writing processes are positive [Haugeland, 1985].

A classic example of a process that can succeed positively is scoring in basketball ; the ball either goes through the hoop or it doesn't. There has never been a semi-basket in the history of basketball. Similarly, in standard symbol systems, such as LISP machines, the most basic production, identification and transformation processes can be carried out with unquestionable success. Barring malfunction, the system can always

tell whether the symbol in its buffer is *foo* or not ; there is simply no question of its being *somewhat foo*. Obviously, printed English is also a digital scheme, given printing presses or laserwriters to produce the words and normal human readers to type-identify them.

An analog scheme is one that is not digital ; analog symboloids are ones that cannot be produced and identified positively. This is not to say that they cannot be successfully produced and identified. Consider trying to determine who won a javelin throw by visual inspection. Most of the time, the test will indicate a winner, but sometimes the difference between the throws will be so small that it will be difficult to choose. Unlike basketball scoring, javelin competition is analog. Similarly, spoken English is analog, since, although it is usually possible to produce an identifiable token of any given word, success is always *more or less* rather than perfect. It makes sense, for example, to say that a given uttered sound was the word *hardest* because it sounded *rather more* like *hardest* than *hottest*.

Analog compositional schemes are, increasingly, cropping up in connectionist work. Consider, for example, Pollack's RAAM architecture [Pollack, 1990]. If one trains such a network to represent many sequences, the representations of various stack states become so closely packed in the activation space of the hidden units that the network cannot positively distinguish one representation from another very close to it in hidden unit activation space. "Attractor chaining" style compositional representations, to be discussed below, are schemes in which the symboloids themselves are analog rather than digital.

Intuitively, whether a set of entities is digital or analog in this sense has much to do with whether or not they are *discrete* rather than *continuous*. Thus it seems natural to say that printed words are discretely different, while the range of forms of spoken words is more continuous. But what does this actually amount to ? Roughly speaking, in this context, a set of entities is continuous if, *between* any two entities in the set, there is another entity which also belongs to the set. A discrete set is one such that there are large *gaps* between entities. Put differently, a continuous set is densely packed into the space of possible entities of that kind, while a discrete set is sparsely packed into it. However, it turns out that it is difficult to give any more precise formulation of this intuitive idea which really does effectively distinguish those schemes of compositional representation which intuition counts as discrete and those which it regards as more continuous. For this reason, together with the fact that digital vs. analog is clearly more significant from the point of view of the functional properties of a scheme of representation, I choose not to emphasize discreteness vs. continuity as a dimension of symboloids.

(c) Arbitrariness

I have noted already that individual symboloids belong to the type that they do in virtue of their distinctive physical configuration or makeup. The next dimension is a matter of how symboloids relate to each other. Are symboloids related in their physical configuration, or are they completely arbitrary relative to each other ? A particularly useful way to ask this question is : does fixing the physical configuration of one kind of symboloid in the scheme in any way constrain the configuration of any others ?

Most words of English, written or spoken, are basically arbitrary in this sense. Thus, a printed token of the word *cat* bears no interesting relationship, in its physical configuration, to a printed token of the word *dog*, over and above the fact that they are both ink marks constituted by smaller marks corresponding to letters. In particular, fixing the physical configuration of one word leaves the physical configuration of the other almost completely unconstrained (which is just to say that we could have used a quite different word in its place ; the choice is completely arbitrary). Likewise, the sound [k^hæt] bears no interesting physical relationship to the sound [dɔg].

The issue of the arbitrariness or non-arbitrariness of symboloids gets particularly interesting when a further complicating factor is introduced : the meaning of the individual symboloids. Symboloids standardly have at least two fundamentally different kinds of properties. On one hand, there are their physical properties. On the other, there is the meaning or semantic significance which each symboloid has and which contributes to the meaning of the compound representation that is formed out of it. Consequently, we can ask whether there is any interesting relationship between these two kinds of properties. Does the physical makeup of a given symboloid in any way reflect its semantic properties ? If it does, that symboloid is not arbitrary with respect to its meaning ; and since meaning is constraining physical configuration, different symboloids with different meanings must exhibit corresponding differences in their physical configuration, and hence the symboloids are not arbitrary with respect to each other.

Clearly, virtually all written or spoken words of English are semantically arbitrary ; a printed token of *cat* bears no interesting relationship to feline mammals over and above the semantic one. The same is true of symbols in any standard computational system. This is one deep source of the anxiety that many people feel about so-called "grounding" in computational systems. If the basic symbols are essentially semantically arbitrary, how can they have any *real* or *intrinsic* meaning ? Surely, it is often thought, symboloids in real cognitive systems must somehow reflect their meanings in their formal configuration more directly.

Some connectionist compositional schemes do utilize semantically non-arbitrary symboloids. For example, in McClelland and Kawamoto's model of case role assignment [McClelland & Kawamoto, 1986], input sentences are built out of vectors corresponding to distinct words, and those vectors themselves are essentially just lists indicating the presence or absence of microfeatures in the designated objects. Therefore, since the actual vector that is used to represent a given object (e.g., a hammer) is constrained by the features of hammers, these vector symboloids are semantically non-arbitrary. (Of course, the vector-elements corresponding to microfeatures are semantically arbitrary. This just shows that it is possible to construct non-arbitrary representations out of arbitrary components — which is exactly what happens in natural and formal languages when non-arbitrary sentences are constructed out of arbitrary word components.)

2. 2. Properties of the Manner of Combination

Compositional representations are obtained by combining symboloids to obtain compound wholes in accordance with the abstract constituency relations. It is natural and useful (though sometimes a little misleading) to think of this in terms of a mechanical process, in which actual individual symboloids are put in one end and out the other comes a compound representation. We can then ask : how does this process work ? What does it actually *do* in combining symboloids to obtain compound representations ?

(d) Mode of Combination

An obvious feature of printed sentences of English is that they are made up out of individual words, and that each word appears in the sentence exactly as it would if it were alone on the page. Words are taken *off the shelf*, as it were, and appear totally unaltered in the resulting sentence. A less obvious feature of *spoken* sentences of English is that this is not, in general, true : when a word appears in a sentence, it typically ends up with a somewhat different shape, depending on the words it is surrounded by, than it has when spoken alone. For example, observe the subtle changes in the pronunciation of the [t] — and hence of the whole word “cut” — in *cut Paul*, *cut some* and *cut out*.

Further, somewhat surprisingly, it is possible to systematically generate compound representations out of symboloids in a way that not only changes, but apparently completely *destroys* those symboloids.

An appropriate term for the kind of combination characteristic of printed sentences is *concatenation*. Roughly, a mode of combination is concatenative just in case it preserves the constituent symboloids in the resulting compound representation ; if, in other words, one can literally point to the constituent symboloids in the compound whole. Since this is true for both printed and uttered sentences of English, we need to go on to distinguish two different kinds of concatenation :

(a) *Pure* concatenation, in which the process of combination leaves symboloids utterly unchanged. All standard computer languages and representational schemes in *classical* cognitive science utilize pure concatenation.

(b) *Context-sensitive* concatenation, in which each constituent symboloid is recognizably present in the resulting compound, but its *shape* has been altered by the very process of combination, typically in a way that systematically reflects its context.

A *non-concatenative* compositional scheme is one in which symboloid constituents do *not* literally appear in the compound representations, even though they may have been input into the combination process. Pollack's RAAM architecture, already mentioned, is one relatively well-known kind of connectionist model which forms compound representations by non-concatenative combination.

In models with this architecture, basic symboloids take the form of vectors on the input layer, and they are systematically combined by the network to form compound (*stack*) representations in the hidden layer. The symboloid constituents are not literally present in the compound representation ; if you examine the activation vector which is the compound representation, you will not find any instance of the vector which was the input symboloid. Nevertheless, the compound representation systematically reflects, in its formal structure, its actual constituency relations. This latter point can be shown in two ways. First, further processing in the RAAM network can actually recreate the symboloid constituents. Second, the compound representations can be used in further processing which is systematically appropriate to its constituent structure (e.g. [Chrisman, 1991]).

(e) Static vs. Temporal Combination

In printed sentences of English, combination is by spatial juxtaposition, and all constituent symboloids are present simultaneously for the entire period of the existence of the whole printed sentence. In *uttered* sentences, by contrast, symboloids are combined temporally ; one follows another, and none are present at the same time. The first kind of

combination, in which all constituent symboloids are simultaneously present, is called *static* combination, while combination by temporal succession is *dynamic*.

Note that dynamic symboloids and dynamic combination are two quite independent properties of compositional representations. The first concerns the nature of the symboloids themselves, the second concerns the way symboloids are combined. Though it is quite natural for dynamic symboloids to be combined dynamically (as in speech) and for static symboloids to be combined statically (as in print), other combinations are possible ; thus, dynamic symboloids can be combined statically (a musical *chord*), and *vice versa* (as when printed words appear sequentially on a screen ; it is their temporal succession which forms a sentence).

(f) Syntactic Conformity

When basic tokens are combined to form compound wholes, to what extent are the syntactic rules of the scheme observed ? A standard assumption is that representations in a compositional scheme are, by definition, grammatically well-formed, i.e., constructed in strict accordance with certain syntactic rules. Any combination of symboloids which violates the syntactic rules is junk or *symbol salad*. However, it is often more useful to regard certain combinations as representations belonging to a given compositional scheme even if they violate some syntactical rules. Many if not most utterances of everyday spoken English are not grammatically well-formed (in the sense that their utterers would wince and then edit them if given an opportunity to examine their own speech closely), yet they are still adequate for representational and communicative functions. Their utility does, however, seem to presuppose at least some general or loose conformity to the syntax of English. The claim is that there is a rough *spectrum* of cases between those symboloidal schemes in which *strict* conformity is observed (e.g., programming languages, printed English), from schemes in which only *loose* conformity is sufficient (e.g., spoken English). It is possible that, insofar as there are compositional representations underlying general cognitive performance, these representations will, like spoken language, exhibit only loose conformity to any relevant syntactic rules.

The theoretical utility of the six dimensions outlined so far is illustrated by the following table which compares five different kinds of compositional representation : sentences of spoken and printed natural language, the representations found in a paradigm classical AI program, Hearsay 2 [Erman *et al.*, 1980] , and two kinds of connectionist schemes, RAAM networks and Port's dynamic auditory recognition model (to be described on the opposite page).

	Spoken Natural Language	Printed Natural Language	Hearsay 2	RAAM [Pollack, 1990]	Dynamic recognition [Port, 1990]
Static vs. Dynamic Symboloids	Dynamic	Static	Static	Static	Dynamic
Analog vs. Digital	Analog	Digital	Digital	Analog	Analog
Arbitrary	Arbitrary	Arbitrary	Arbitrary	Non- arbitrary	Non- arbitrary
Mode of Combination	Context- Sensitive	Pure	Pure	Non- concatenative	Context- sensitive
Static vs. Temporal Combination	Temporal	Static	Static	Static	Temporal
Syntactic Conformity	Loose	Strict	Strict	Strict	Loose

Table 1. A comparison of five different schemes of compositional representation along six dimensions of compositionality.

Note that each dimension separates these five kinds roughly evenly into different groups. Note also that printed English and the representations in the classical AI system, Hearsay 2, share exactly the same properties, and that spoken natural language and the representations in Port’s dynamic recognition model (columns 1 and 5) are also highly similar. This suggests that these two clusters of implementations of compositionality are in fact natural kinds. When compositional schemes of representation are developed, it is natural on one hand to have static, digital and arbitrary symboloids combined statically by pure concatenation in strict syntactic conformity ; on the other hand, it is natural to have dynamic, analog symboloids temporally combined by context-sensitive concatenation and in loose syntactic conformity. I will describe the first kind of compositionality, characteristic of printed natural language, as *symbolic*, and the second kind, characteristic of spoken language, as *dynamic*.

As mentioned above, these six properties of implementations of compositionality can be thought of as dimensions of a large space of possible kind of compositionality. While research in cognitive science has sampled representational formats from various regions of this space, clearly it is generically symbolic implementations of compositionality that have been dominant. In the third section of this paper I will be suggesting

that compositionality, dynamics, and neural networks might best be reconciled by means of *dynamic* compositionality.

3. Dynamic Compositionality in Neural Networks

Paul Smolensky once characterized the problem of connectionist representation as that of finding an appropriate mapping from the set of entities that one wishes to represent into a vector space [Smolensky, 1990, p. 168]. If the entities to be represented are symbolic structures, then the problem is to find a mapping from the set of those symbolic structures into the vector space. The vectors that are the output of this mapping can then be seen as compositionally structured representations themselves.

Why is this mapping into a *vector* space ? Well, obviously, because these vectors correspond to the pattern of activity over the neural units of some connectionist network at a given time. Note that a pattern of neural activity, in this sense, is a *static* representation. Change over time is not essential to the type-identity of this representation. Only the particular distribution of activity values over the units *at a time* determine what representation you have. The changes that occur *over* time as the network states evolve are *processing* of representations ; they are transitions from one representation to the next.

If we consider neural networks as dynamical systems, then a pattern of activity across the network is the *state* of the system at a given point in time. The notion of the *state* of a system is the most elementary concept of dynamical systems theory. Indeed, many, perhaps most connectionists who have conceptualized the problem of representation as that of finding a mapping simply into *states* of the system have then gone on to study those representations and their transformations in largely non-dynamical way, paying only lip-service to the idea that their networks are dynamical systems.

From the dynamical systems perspective, the more general problem of neural network representation is to find an appropriate mapping from the entities to be represented into some set of dynamical objects existing in a neural network system. At the outset, these objects might be anything specified by means of the conceptual repertoire of dynamical systems theory. In practice, proposals have dealt only with system *states*, *trajectories*, and *attractors*.

A well-known alternative to taking network representations to be simply system *states* is to take them to be fixed point attractors. Hopfield networks are a standard example ; the local *minima* of the energy function of such networks are fixed point attractors, and the network comes to represent a given input by taking that input as its initial conditions and settling into the attractor into whose basin of attraction those initial

conditions fall [Hopfield, 1982]. This approach may be adequate for certain kinds of problems, but there is a very serious difficulty confronting anyone wanting to use this approach to develop a scheme of *compositional* representations. The problem is the essentially *productive* nature of compositional schemes. The point of having a compositional scheme is to provide a *potentially unbounded* set of representations that nevertheless bear systematic structural similarities to one another. If representations are to be fixed point attractors of a network dynamical system, then we need a network whose dynamics is sufficiently complex that it already has an arbitrarily large number of distinct fixed point attractors, one for each compound representation. As far as I know, there are currently no practicable proposals for such a system.

The deep problem here, in short, is to find a way to incorporate the *productive* nature of compositional schemes of representation into a genuinely dynamical network framework.

A radical approach is to reject the standard connectionist paradigm in which representations are conceived as points in network activation space, and to think of them instead as *trajectories* : that is, as essentially temporally extended, dynamic entities. Further, one should not suppose that the network dynamics has every possible compound representation somehow already pre-configured ; rather, the network dynamics must provide the *resources* by means of which arbitrary compound representations can be constructed.

One method of doing this is what Jean Petitot calls "attractor syntax" [Petitot, forthcoming, 1995] and is also known as "attractor chaining". Consider a dynamical system with certain state variables and parameters. Change in system state is governed by its evolution equations which depend on system parameters. Another way to put this is that the *dynamics* of the system is fixed by the current specification of the system parameters. We can thus think of the parameters as *controlling* the behavior of the system. For any given set of parameter specifications the system possesses a certain landscape of attractors. As control parameter specifications vary, this landscape changes also. Significant qualitative changes in the attractor landscape are known as bifurcations. The essence of attractor chaining is to take the representation of a complex structure to be the trajectory that results as the state of the system evolves under the influence of a changing attractor landscape. Compositional structure is imposed on the trajectory by bifurcations that occur as control parameters vary.

Currently the most mathematically sophisticated treatment of the attractor chaining approach is that under development by Jean Petitot. A dramatic change in system state as the system undergoes a bifurcation is known as a *catastrophe*. René Thom developed the mathematical theory of catastrophes and suggested the application of this theory in many domains including linguistics [Thom, 1975, 1983]. Petitot is engaged in

combining the mathematics of catastrophe theory with Langacker's *cognitive grammar* [Langacker, 1987] in a neural networks framework.

A more standard connectionist approach is work by Bob Port and his students on dynamic recognition for auditory sequences ([Anderson & Port, 1990], [Port & Cummins, forthcoming, 1995], [Port, 1990]). In this work, a simple recurrent network is trained to recognize auditory sequences. The network is a dynamical system which, at any given time, has a single global point attractor whose location is determined by the current setting of the control parameters. The system state is always heading in the direction of this point attractor ; if there is no change in the parameters, it will settle into that point. Things get interesting when the location of the current global attractor is changed by variation in the control parameters. The state of the system is pulled first in one direction, then another. The resulting *trajectory* reflects in its temporal structure the sequence of fixed point locations. Distinctive trajectory shapes correspond to each such sequence (see Figure 1, below).

Figure 1 caption
A schematic illustration of a trajectory formed under the influence of a changing dynamical landscape. The points A, B and C represent the location of a single global attractor under corresponding parameter settings. The system state drifts in the direction of the global attractor as long as it is active. The temporally extended trajectory itself is the compositional representation. The axes are the first and second principal components of the system state-space. From [van Gelder & Port, in press]. For more detailed explanation, see [Port & Cummins, forthcoming, 1995].

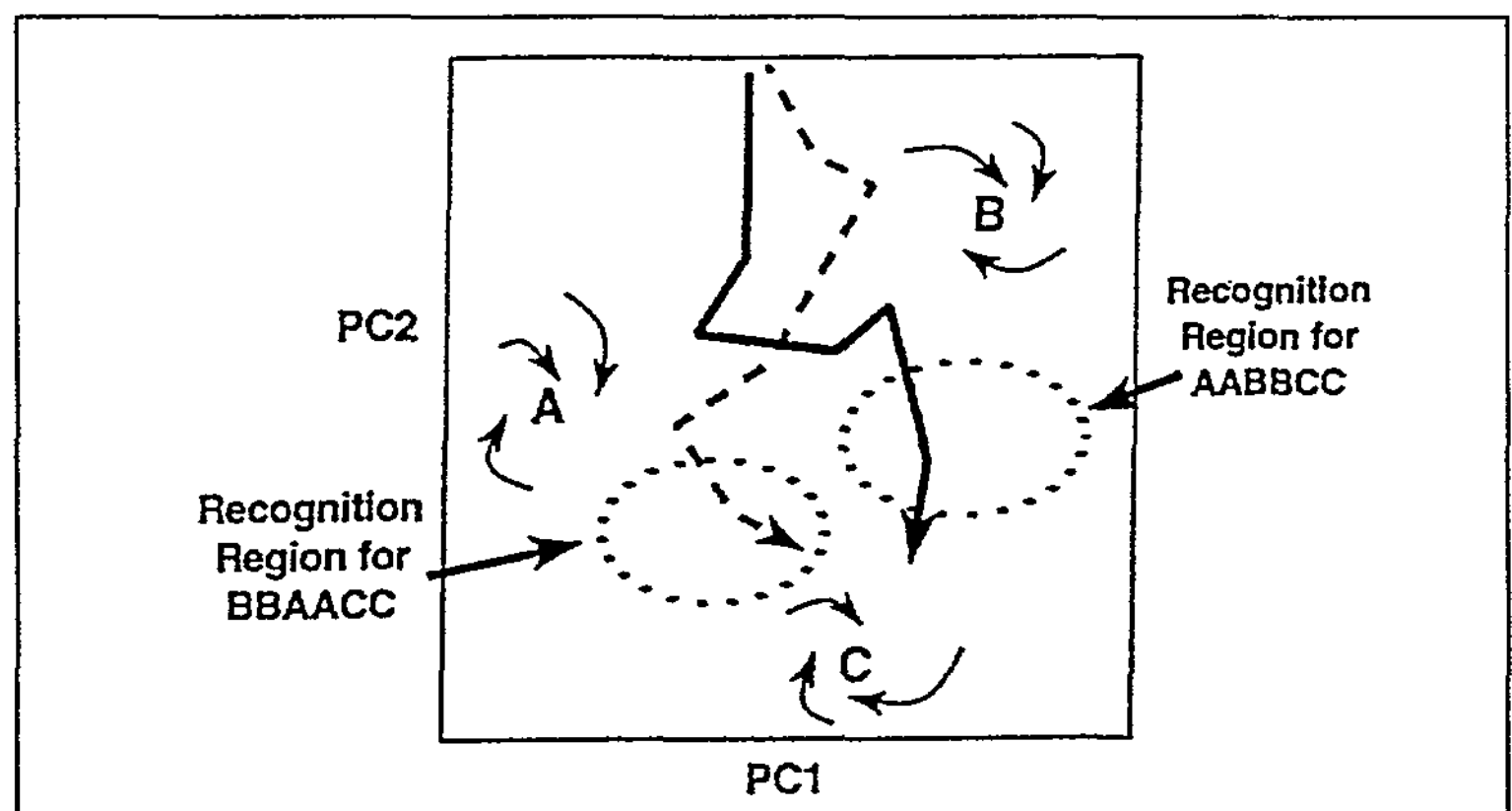


Figure 1

How are these trajectory-forming parameter changes actually implemented ? The Port recurrent network model has an input layer (see Figure 2, on the opposite page). A pattern of activation across this layer corresponds to the acoustic spectrum as the network is exposed to another element in the auditory sequence. In gaining a truly dynamical understanding of how these networks work, however, it is crucial *not* to think of the relationship between the input layer and the recurrent network in traditional connectionist terms. It is a mistake to see the input pattern as transformed into another pattern in the recurrent network. Rather, the activation levels of the input layer units are parameters which subtly influence the *dynamics* of the recurrent network considered as a stand-alone dynamical system. The input pattern does not become another

pattern across the network units ; rather, it shapes the way the state of the recurrent network evolves by molding the dynamical landscape.

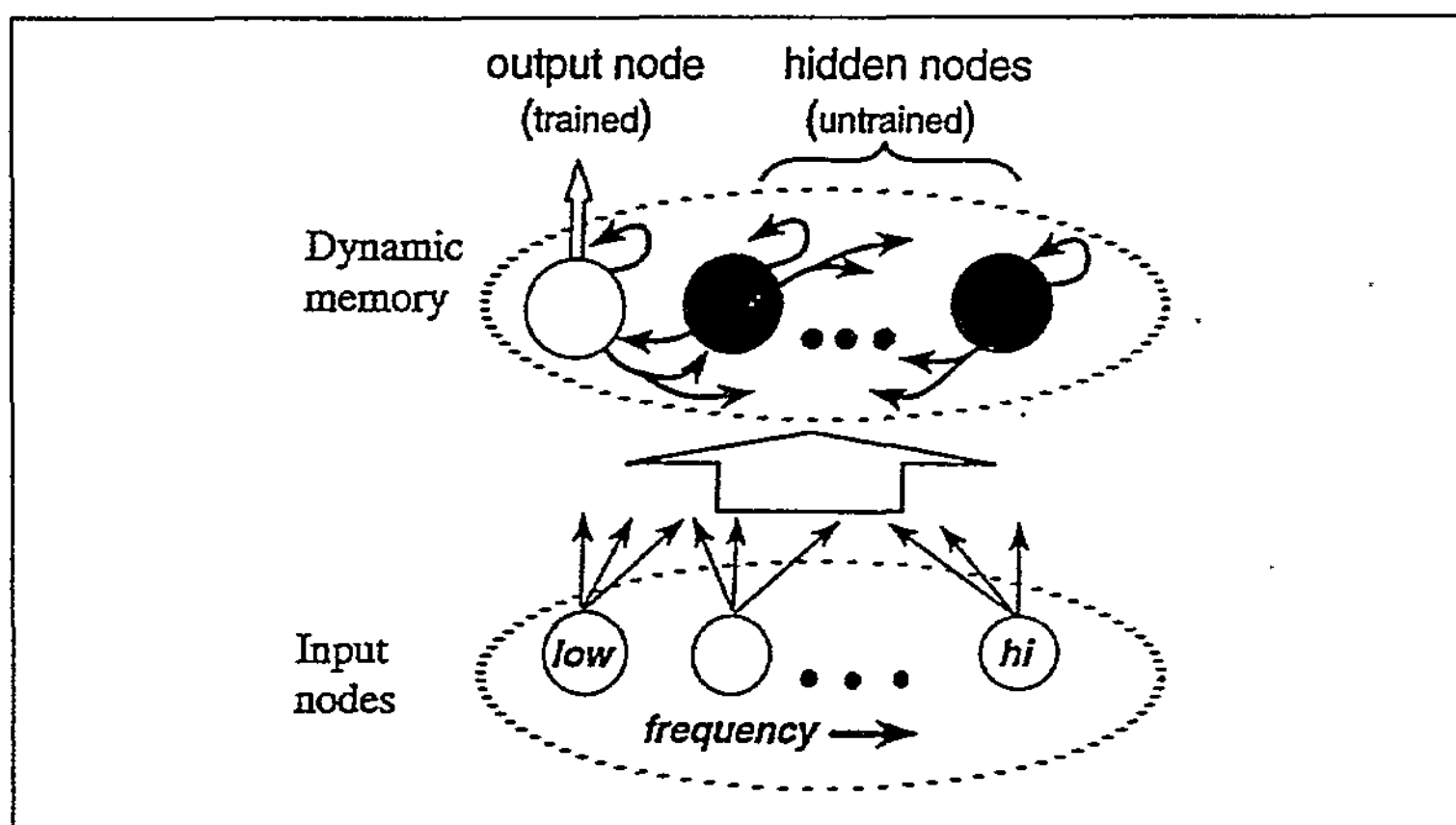


Figure 2

The Port work has focused on perception of simple auditory sequences. It is not yet anything like a practical general solution to the problem of incorporating compositional representations in network dynamical systems. It is however a modest implementation of the attractor chaining approach to constructing complex representations. It therefore provides a useful prototype for the construction of general schemes of dynamical compositional representation. In this prototype, representations are trajectories systematically constructed out of basic compoundable units. Those basic units are nothing like the static, context-free symbols of standard implementations of compositional representation. Rather, they are themselves trajectory segments — stages of processing in which the system state evolves in the direction of a particular global attractor. The compositional representation is the context-sensitive, temporal concatenation of these trajectory segments.

In general, the kind of compositional representation found here is much more similar to spoken natural language than it is to printed natural language or to *LISP* structures. In order to see these entities as compositional representations — and, more generally, in order to understand how dynamics, compositionality and neural networks might be combined — we must have a suitably enriched conception of compositionality and how it can be implemented.

Figure 2 caption
The architecture of the Port dynamic auditory recognition model. Each tone in an auditory sequence is input as a pattern of activation across the input nodes. Activation levels at this layer function as parameter settings controlling the dynamics of the main system, a fully recurrent network. As a sequence of tones is input, the attractor landscape in the main system changes; the activation trajectory then reflects the structure of this sequence ; from [van Gelder & Port, in press]. For more detailed explanation, see [Port & Cummins, forthcoming, 1995].

Research School of Social Sciences
Australian National University (Canberra ACT 0200 Australia)
tv@coombs.anu.edu.au

References

- ANDERSON (S.) & PORT (R.)
1990, "A Network Model of Auditory Pattern Recognition", *Technical Report* (Indiana University), n° 11.
- ASHBY (R.)
1952, *Design for a Brain*, London, Chapman and Hall.
- BOLINGER (D.)
1975, *Aspects of Language*, New York, Harcourt Brace Jovanovich Inc.
- CHRISMAN (L.)
1991, "Learning Recursive Distributed Representations for Holistic Computation", *Connection Science*, n° 3 (n° 4), p. 345-366.
- [Erman *et al.*]
ERMAN (L. D.) & HAYES-ROTH (F.) & LESSER (V. R.) & REDDY (D. R.)
1980, "The HEARSAY-II Speech Understanding System : Integrating Knowledge to Resolve Uncertainty", *Computing Surveys*, n° 12, p. 213-253.
- FODOR (J. A.) & PYLYSHYN (Z. W.)
1988, "Connectionism and Cognitive Architecture : A Critical Analysis", *Cognition*, n° 28, 1-2, p. 3-71.
- GROSSBERG (S.)
1988, *Neural Networks and Natural Intelligence*, Cambridge (MA), MIT Press.
- HAUGELAND (J.)
1985, *Artificial Intelligence : The Very Idea*, Cambridge (MA), MIT Press.
- HOPFIELD (J.)
1982, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", *Proceedings of the National Academy of Sciences (USA)*, n° 79, p. 2554-2558.
- LANGACKER (R. W.)
1987, *Foundations of Cognitive Grammar*, vol. 1, Stanford University Press.
- MCCLELLAND (J. L.) & KAWAMOTO (A. H.)
1986, "Mechanisms of Sentence Processing : Assigning Roles to Constituents of Sentences", in *Parallel Distributed Processing : The Microstructure of Cognition*, J. L. McClelland and D. E. Rumelhart eds., Cambridge (MA), MIT Press.
- PETTITOT (J.)
forthcoming, 1995, "Morphodynamics and Attractor Syntax", in PORT (R.) & van GELDER (T.), eds.
- POLLACK (J. B.)
1990, "Recursive Distributed Representations", *Artificial Intelligence*, n° 46, 1-2, p. 77-105.
- PORT (R. F.)
1990, "Representation and Recognition of Temporal Patterns", *Connection Science*, n° 2, p. 151-176.

PORT (R.) & CUMMINS (F.)

forthcoming, 1995, "Modeling Auditory Recognition Using Attractor Dynamics",
in PORT (R.) & van GELDER (T.), eds.

PORT (R.) & van GELDER (T.), eds.

forthcoming, 1995, *Mind as Motion : Explorations in the Dynamics of Cognition*,
Cambridge (MA), MIT Press.

PYLYSHYN (Z. W.)

1984, *Computation and Cognition : Toward a Foundation for Cognitive Science*,
Cambridge (MA), MIT Press.

SMOLENSKY (P.)

1988, "On the Proper Treatment of Connectionism", *The Behavioral and Brain
Sciences*, n° 11, p. 1-74.

1990, "Tensor Product Variable Binding and the Representation of Symbolic
Structures in Connectionist Systems", *Artificial Intelligence*, n° 46, p. 159-216.

THOM (R.)

1975, *Structural Stability and Morphogenesis* (Fowler, D.H., Trans.), Reading
(MA), W. A. Benjamin Inc.

1983, *Mathematical Models of Morphogenesis*, Chichester, Ellis Horwood.

Van GELDER (T.)

1990, "Compositionality : A Connectionist Variation on a Classical Theme",
Cognitive Science, n° 14, p. 355-384.

Van GELDER (T.) & PORT (R.)

in press, "Beyond Symbolic : Prolegomena to a Kama-Sutra of Compositionality",
in *Symbol Processing and Connectionist Network Models in Artificial Intelligence
and Cognitive Modelling : Steps Toward Principled Integration*, V. Honavar and
L. Uhr eds., San Diego, Academic Press.

Structural Syntax : The View from Cognitive Grammar

Ronald W. Langacker

My purpose is to compare particular facets of Tesnière's **structural syntax** with the contemporary theory of **cognitive grammar** (CG) [Langacker, 1987a, 1990, 1991a]. While I have no intention of labeling Tesnière as the "first cognitive grammarian", it is worth noting the substantial affinity between the two frameworks (as well as certain differences).

¹Page numbers refer to the 1965 edition of «*Éléments de Syntaxe Structurale*».

A pivotal theoretical issue is the relation between meaning and grammar. On first examination, it appears that Tesnière shares with [Chomsky, 1957] the view that grammatical structure is distinct and autonomous *vis-à-vis* semantics : "Le plan structural et le plan sémantique sont donc théoriquement entièrement **indépendants** l'un de l'autre" (41)¹. Nevertheless, meaning is for Tesnière the ultimate basis of grammar : "(...) le sens [est] en dernière analyse la raison d'être de la structure" (40). The **structural connections** that constitute for Tesnière the very crux of grammar do not exist independently of the **semantic connections** they express : "(...) il n'existe jamais de connexion structurale sans connexion sémantique" (44). Such statements are contrary to the original vision of an autonomous syntax, but perfectly compatible with the central claim of CG, namely that meaning and grammar are indissociable.

Tesnière was not actually inconsistent on this matter. The apparent discrepancy stems from his lack of a comprehensive view of linguistic semantics. He specifically equates semantic structure (*le plan sémantique*) with ideas, thoughts, or concepts that are rich in content and exist independently of linguistic expression. He describes *les mots pleins*, for example, as "ceux dont la forme est associée directement à une idée, qu'elle a pour fonction de représenter et d'évoquer" (53). A non-linguistic conception of this sort — e.g. the "raw conception" of a horse throwing its rider — is apparently what he has in mind when he emphasizes the independence of semantics and grammar. After all, that same conceived event might be coded linguistically by either an active or a passive.

Yet there is more to meaning than just "raw conceptions", or conceptual "content". Linguistic semantics must also accommodate our multifaceted ability to **construe** such content in alternate ways. We can, for instance, abstract away from specific, detailed conceptions and portray conceived entities at varying levels of **schematicity** (e.g. *computer* > *machine* > *object* > *thing*). We can also establish **correspondences** between subparts of distinct conceptions, and portray one conception against the **background** of another (as in metaphor). Tesnière was certainly aware of these phenomena : they are simply not what he had in mind when speaking of *le sens* or *le plan sémantique*. They do however figure in his discussion of grammatical elements and syntactic relationships. Semantic correspondence, for example, is one of the notions needed to give an explicit account of his *connexions sémantiques*. Thus, if terms like **meaning** and **semantics** subsume not only conceptual content but also semantic correspondence and other aspects of construal, it follows that meaning and grammar are not at all separate or independent. In fact, Tesnière's vision does not seem incompatible with the basic tenet of CG that grammatical structure *reduces* to conventional patterns for the construal of conceptual content and its phonological symbolization.

CG posits just three kinds of linguistic structures : **semantic**, **phonological**, and **symbolic**, the latter residing in the symbolic linkage between a semantic and a phonological structure (its **semantic** and **phonological** poles). Moreover, permitted structures are limited to (parts of) **overtly occurring expressions**, to **schematizations** of permitted structures, and to relationships of **categorization** between permitted structures (e.g. the relation between a schema and a specific expression that **instantiates** or **elaborates** it). Lexicon, morphology, and syntax form a continuum, comprising only **assemblies of symbolic structures** (form-meaning pairings). Thus all grammatical elements are claimed to be **meaningful**. By and large, those symbolic assemblies considered "lexical" are fairly specific, whereas those considered "grammatical" are more schematic.

This reduction of grammar to symbolic assemblies requires a **conceptualist** semantics that accommodates construal. On the content it invokes, every symbolic element imposes a particular construal which constitutes an intrinsic and essential aspect of its conventional semantic value. Besides our ability to portray a situation at varying levels of *specificity*, and to conceptualize one structure against the *background* of another, dimensions of construal include **perspective** (e.g. *vantage point*) and various types of **prominence**. Two important kinds of prominence are **profiling** and the **trajector/landmark** asymmetry.

Profiling can be characterized as *reference within a conceptualization*. From the array of conceptual content it evokes — its conceptual **base** — every expression selects a particular substructure to serve as its **profile** (*conceptual referent*). The word *knuckle* evokes as its base the conception of a finger, within which it profiles a subpart — a joint

(Fig. 1(a))². Likewise, *intermission* has for its base the conception of a scheduled activity extending through time (*t*) and profiles a temporal interval during which that activity is temporarily suspended (Fig. 1(b)). *Knuckle* and *intermission* profile **things** (in an abstract sense of that term). Expressions can also profile **relationships**. Fig. 1(c), for example, depicts one sense of the preposition *across* (as in *the shop across the street*), where it profiles the spatial relation between two things, as reckoned from a certain vantage point (*V*). Specifically, the thing being located — termed the **trajector** (*tr*) — lies at the end of a mental path (dashed arrow) that originates at the vantage point and traverses the expanse of another entity, which serves as a spatial **landmark** (*lm*).

²Heavy lines indicate profiling.

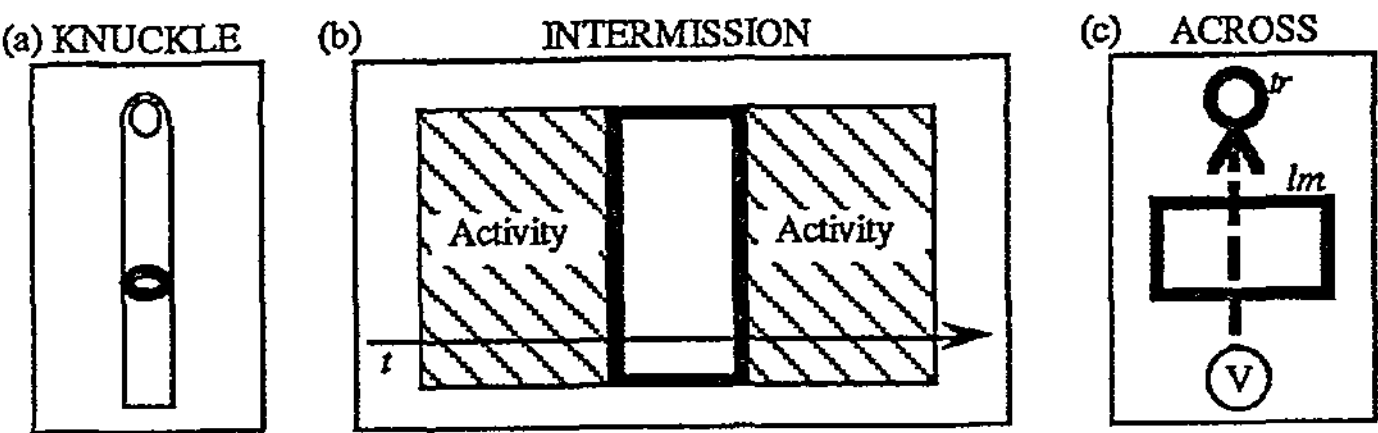


Figure 1

The terms **trajector** and **landmark** indicate a kind of prominence involving the participants in a profiled relationship. In a relational expression, one participant stands out as the focus of attention, as the entity the expression is concerned with locating, assessing, or characterizing. This element — the **trajector** — is described as the **primary figure** within the relational profile. Should the profile include an additional focus of attention, with the status of **secondary figure**, it is called a **landmark**. Trajector/landmark alignment is sometimes the only conceptual factor to which a semantic contrast is attributable. In Fig. 2, for example, *before* and *after* evoke the same conceptual content — pertaining to the temporal location of two events — and profile precisely the same relationship between them. The semantic contrast is solely a matter of which participant event is construed as a landmark for purposes of locating the other.

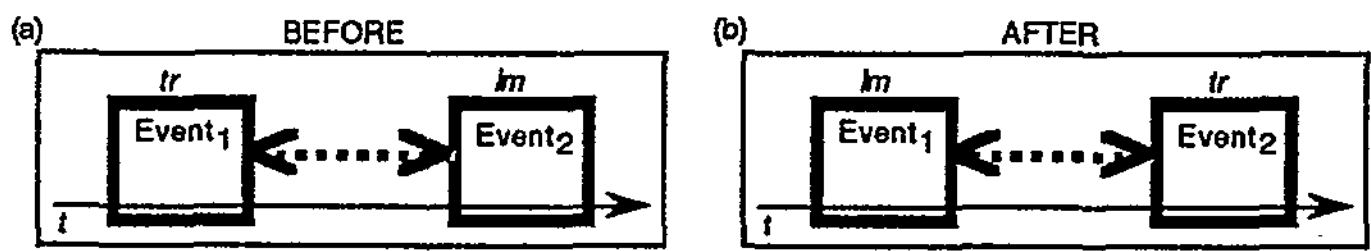


Figure 2

Symbolic structures are considered “grammatical” by virtue of being schematic, so that content is less important than construal. Tesnière’s failure to fully appreciate this aspect of meaning is apparent in his discussion of grammatical classes and markers. With respect to categories, he posits “une opposition principielle entre le plan catégorique et le plan sémantique” (49). But even as he emphasizes their fundamental difference, Tesnière observes the close connection between “semantic” (i.e. conceptual) and grammatical categories. The latter are abstracted from specific ideas, thereby allowing them to be grasped, arranged, and classified: “Les catégories grammaticales (...) sont les chefs sous lesquels viennent s’ordonner dans chaque langue (...) les idées qu’elle a à exprimer (...). Le plan catégorique est celui des idées générales susceptibles de servir de classificateurs” (49). This reference to “idées générales” is compatible with my own view that grammatical elements are *schematic* with respect to the content of specific expressions.

I claim that basic grammatical classes can be characterized semantically in terms of both a **prototype**, describing *central* category members, and a highly abstract **schema**, valid for *all* class members. The respective prototypes for nouns, verbs, and adjectives, for example, are physical objects, actions, and properties. To accommodate all category members, schematic characterizations must be independent of any specific conceptual *content*, residing instead in ways of *construing* such content. An expression’s grammatical class depends on the nature of its profile. A noun profiles a **thing**, defined abstractly as a conceptual “region” [Langacker, 1991b]. A verb profiles a **process**, defined as a relationship followed sequentially in its evolution through time. Such classes as adjectives, adverbs, prepositions, and participles profile various sorts of **atemporal relations**, conceived holistically (i.e. their temporal evolution is not in focus). Abbreviatory notations are given in Fig. 3.

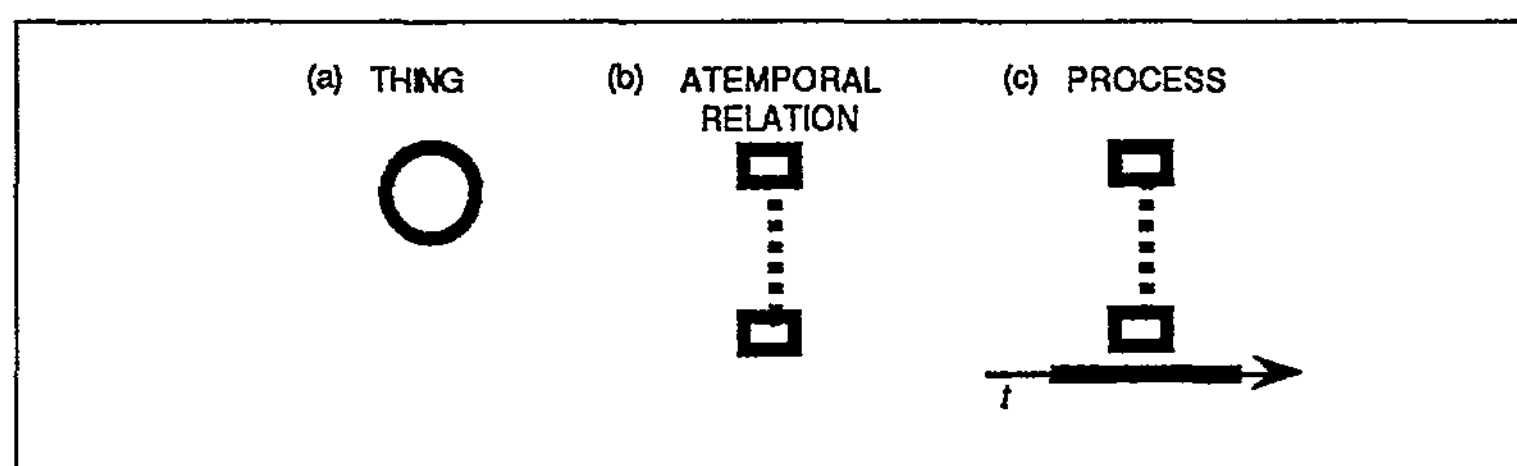


Figure 3

Tesnière likewise distinguishes between specific and more schematic structures — *mots pleins particuliers* (such as *cheval*) vs. *mots pleins généraux* (e.g. *quelqu’un*) — and identifies schematic notions as the basis for categorization. He further defines the class of nouns in terms of physical objects, and verbs in terms of processes (61). There are however

important differences. Tesnière's definitions do not apply to all class members, but just to the prototype. Additionally, he does not explicitly formulate any notion equivalent to profiling or other aspects of construal. The result — e.g. for adjectives (62) — is content-based definitions that are too vague for comfort.

Tesnière's ambivalence between a narrow and a broader conception of semantics is further evident in his discussion of "empty words". While he contrasts them quite sharply with *les mots pleins*, saying that "*les mots vides (...) ne sont pas chargés d'une fonction sémantique*" and that "*ils n'évoquent rien par eux-mêmes*" (53), he implicitly recognizes that the distinction is less than absolute. *Les mots vides* are attributed "roles" and "values" that would now be subsumed under *meaning* and *linguistic semantics*: the conjunction *et* (80), the preposition *sur* (81), the past-tense inflection — *a* (92), the negation marker *ne... pas* (218), etc. Tesnière himself observes that these "empty words" develop historically from "full words" by loss of semantic content, and indicates that the result is a matter of degree: "*(...) il est très difficile de vider complètement de sa valeur sémantique un mot plein et d'arriver à le concevoir comme un mot vraiment vide*" (55).

In CG, all such elements are seen as having at least schematic semantic value. As shown in Fig. 4, for example, a past-tense morpheme profiles a fully schematic process. Beyond this abstract conception, its only content resides in an unprofiled relationship of temporal anteriority with respect to the speech event, also schematic; see [Langacker, 1985, 1990 (ch. 12), and 1991a (ch. 6)].

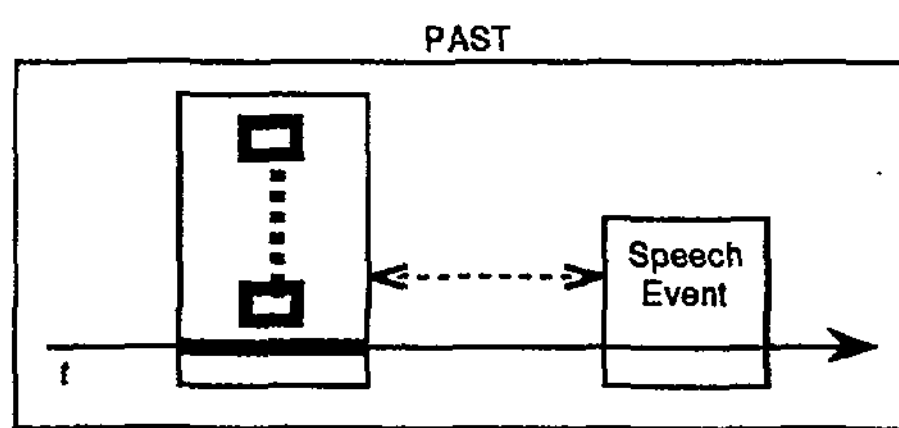


Figure 4

We must next consider grammatical constructions. For Tesnière, the syntactic organization of *Alfred frappe Bernard* is represented by the *stemma* shown in Fig. 5(a). The verb *frappe* is the *régissant*, hence placed at the top, whereas the two nouns are placed below it, each being a *subordonné*. The solid lines indicate *connexions structurales*, which are "superimposed" (42) on *connexions sémantiques*, shown as dashed lines in Fig. 5(b). *Alfred* and *Bernard* are of course *actants*. Their respective status as *prime actant* and *second actant* is reflected diagrammatically by

left-to-right ordering, or by labeling the branches with numbers, as in 5(c). Tesnière also presents a more abstract kind of representation, called a *stemma virtuel*. The common syntactic structure of *Alfred frappe Bernard* and *Charles voit Daniel* is thus depicted in 5(d), where *I* stands for a verb, and *O* a noun.

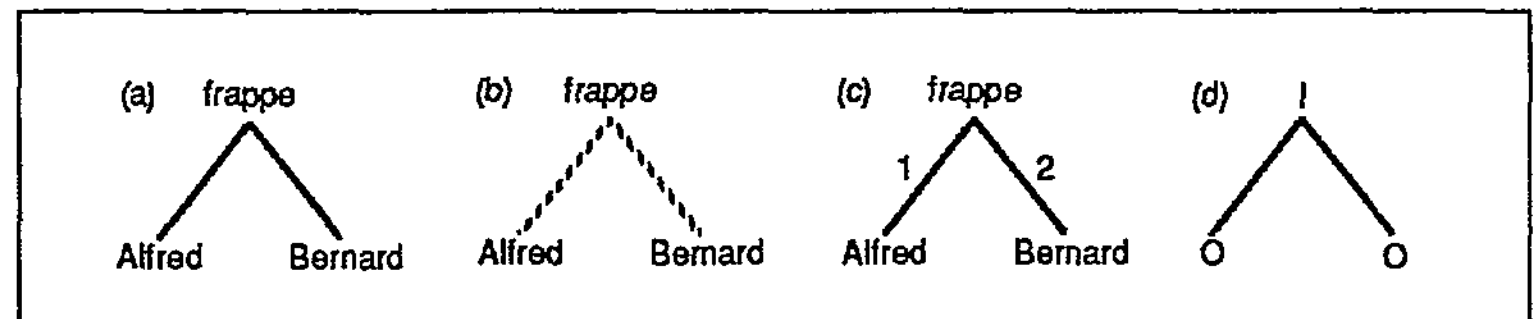


Figure 5

Apart from differences in terminology and notation, this is quite compatible with CG, which is however more explicit about the meanings of elements and the semantic basis of their grammatical connection. Consider *Alfred hit Bernard*, assuming (for sake of comparison) that it has no internal constituency grouping. Fig. 6 displays the relations among the subject, verb, and object. Note first that the diagram contains four elements, not just three. This is because the representation makes explicit, at each level of organization, the nature of the **composite structure** that emerges by virtue of the relationships among the **component structures**. The component structures *Alfred*, *hit*, and *Bernard* are shown at the bottom. They are **integrated** to form the composite structure *Alfred hit Bernard*, shown at the top. There is nothing analogous to the composite structure in Tesnière's representational system. While he recognizes (e.g. in the notion *nucleus*) that multiple elements can function as a group in their relation with other elements, he gives no explicit characterization of a group as an integral whole.

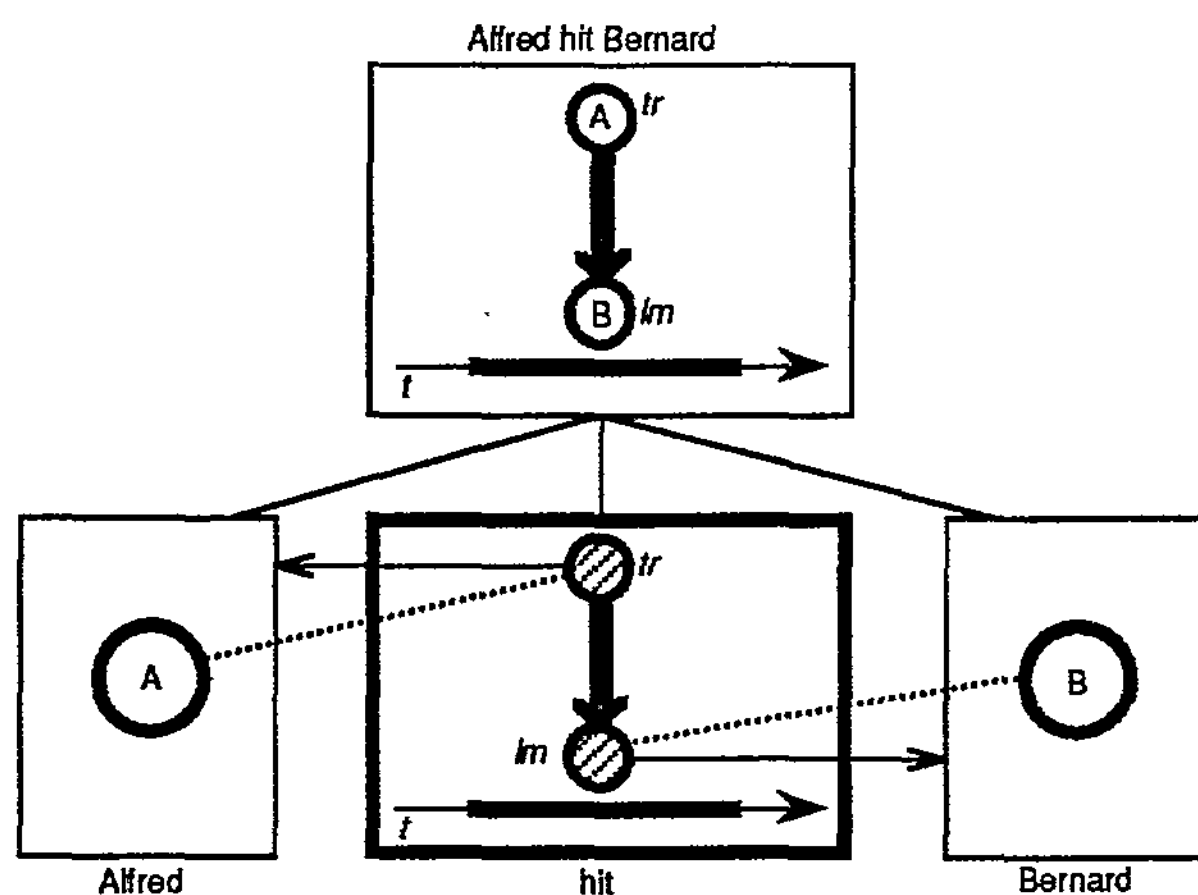


Figure 6

The component and composite structures are *symbolic* (each having a *semantic* and a *phonological pole*) and thus constitute an *assembly* of symbolic structures (or *construction*). Integration takes place at both poles, the phonological integration serving to *symbolize* the semantic integration. The informal diagrams indicate certain basic properties of the semantic structures and their integration. For the component structures *Alfred* and *Bernard*, it is merely shown that each profiles a *thing* ; the letters *A* and *B* abbreviate their other semantic specifications. The notation for the verb also abbreviates a fuller semantic value : it specifies that the designated process has two central participants, and the double arrow represents the transmission of energy.

What shows that *hit* is the head, or *régissant*, to which *Alfred* and *Bernard* are subordinated ? As a notational device, the head is enclosed in a heavy-line box. What it means in conceptual and grammatical terms for a component to be the head is that it has the same profile as the composite structure. The composite structure usually inherits its profile from one of the components, which is thus called the **profile determinant**. *Hit* is the profile determinant in *Alfred hit Bernard*, since the process it designates is also profiled by the expression as a whole. The verb is thus the head at this level of organization.

Although Tesnière states that “le sens du subordonné porte sur celui du régissant” (42), he does little by way of elucidating this essential notion. The semantic characterizations of CG make possible a reasonably precise account of these relationships. The integration of component structures is based on **correspondences** between their substructures. Represented diagrammatically by dashed lines, these correspondences are indications of the *conceptual overlap* by virtue of which the components merge to form a coherent composite conception. In Fig. 6, we see that the profile of *Albert* corresponds to the trajector of *hit*, and that the profile of *Bernard* corresponds to the landmark. The semantic specifications of *Alfred* and *Bernard* (given as *A* and *B*) are thus superimposed on *hit*'s trajector and landmark, respectively, to form the composite semantic structure.

It is usual for components to be linked semantically by correspondences between prominent substructures. It is also typical for one component structure to **elaborate** a schematic substructure of another. In Fig. 6, *Alfred* elaborates the schematic trajector of *hit*, while *Bernard* elaborates its schematic landmark. The trajector and landmark are hatched to signal their status as **elaboration sites**, and arrows represent the elaborative relationship. These correspondences and elaborative relationships make explicit the manner in which “le sens du subordonné porte sur celui du régissant”. In particular, the import of *Alfred* being the *prime actant* of *hit*, and *Bernard* its *second actant*, is that their profiles respectively correspond to the verb's trajector and landmark, which they serve to elaborate.

This construction instantiates a general pattern of composition : *Alfred hit Bernard* is structurally parallel to *Charles sees Daniel*, etc. Tesnière represents this by means of the *stemma virtuel* in Fig. 5 (d). The analog in CG is a **constructional schema**, an assembly of *schematic* symbolic structures representing the abstract commonality observable across a set of complex expressions. A constructional schema is parallel to instantiating expressions in terms of such factors as profiling, profile determinance, correspondences, etc., but is more schematic in regard to content. Thus the structure in Fig. 6 instantiates the schematic assembly depicted in Fig. 7. The head is characterized schematically as a transitive verb (i.e. it designates a process with two things as focal participants), and the other two components as nominal expressions (they profile things). X, Y, and Z refer in schematic terms to phonological sequences.

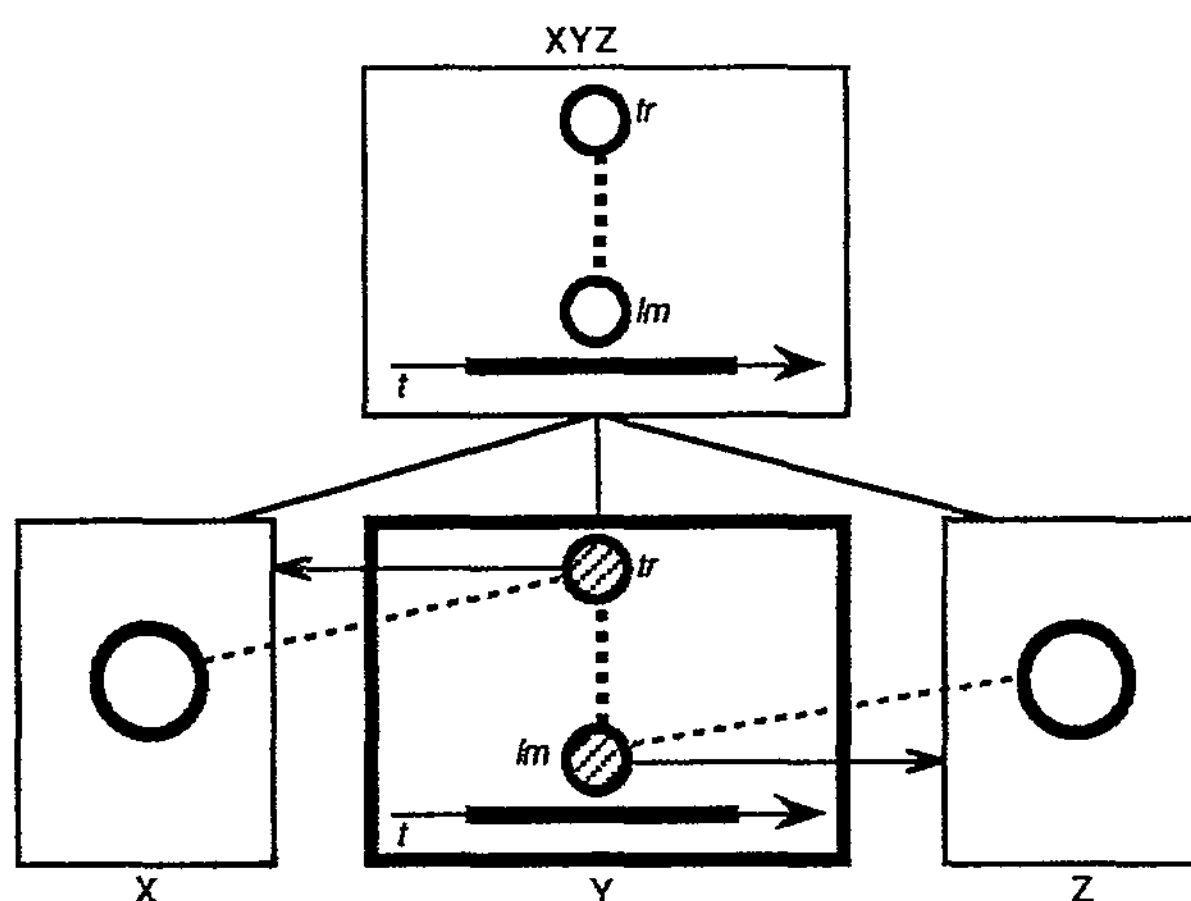


Figure 7

Tesnière's stemmas permit the characterization of diverse structures having any degree of complexity. I believe that all the valid notions and relationships he posits are straightforwardly accommodated in CG. This latter however embodies a more unified conception of various aspects of linguistic structure that Tesnière distinguishes fairly sharply. It makes no principled distinction, for example, between *mots pleins* and *mots vides*, treating them both as meaningful and as comparable in regard to how they form constructions with other elements. Nor does CG recognize *translation* as a distinct phenomenon requiring special notations and theoretical constructs.

Consider the sentence *Alfred hit Bernard before Charles arrived*. Tesnière's analysis is shown in Fig. 8. The past-tense forms *hit* and *arrived* each function as clausal nuclei, circled in the second instance to

indicate that the entire complex structure is the *régissant* with respect to *Charles*. The past-tense marker is not itself accorded the status of *nœud*, *régissant*, or *subordonné*. The clause *Charles arrived* gives rise to an adverb (*E*) by the process of *translation*, and that adverb functions as a dependent of *hit*, specifically a *circonstant*. *Before* is analyzed as a *translatif*, an “empty word” which effects the translation of a verb into an adverb.

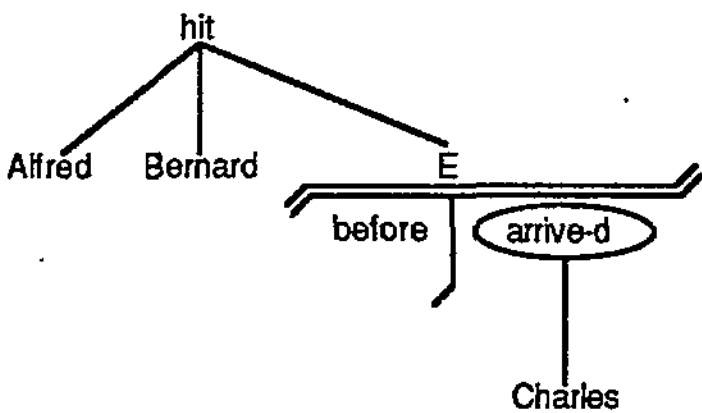


Figure 8

By contrast, the CG analysis does not involve any apparatus beyond that required for *Alfred hit Bernard*. We might begin with the past-tense marker. Zero in the case of *hit*, it is nonetheless a meaningful element. I will assume that a tense marker is the last element to be added in the formation of a finite clause. It thus combines with the remainder of the clausal structure, and is manifested phonologically on the *head* of that structure. The integration of *PAST* with the (tenseless) *Alfred hit Bernard* (Fig. 6) is diagrammed in Fig. 9.

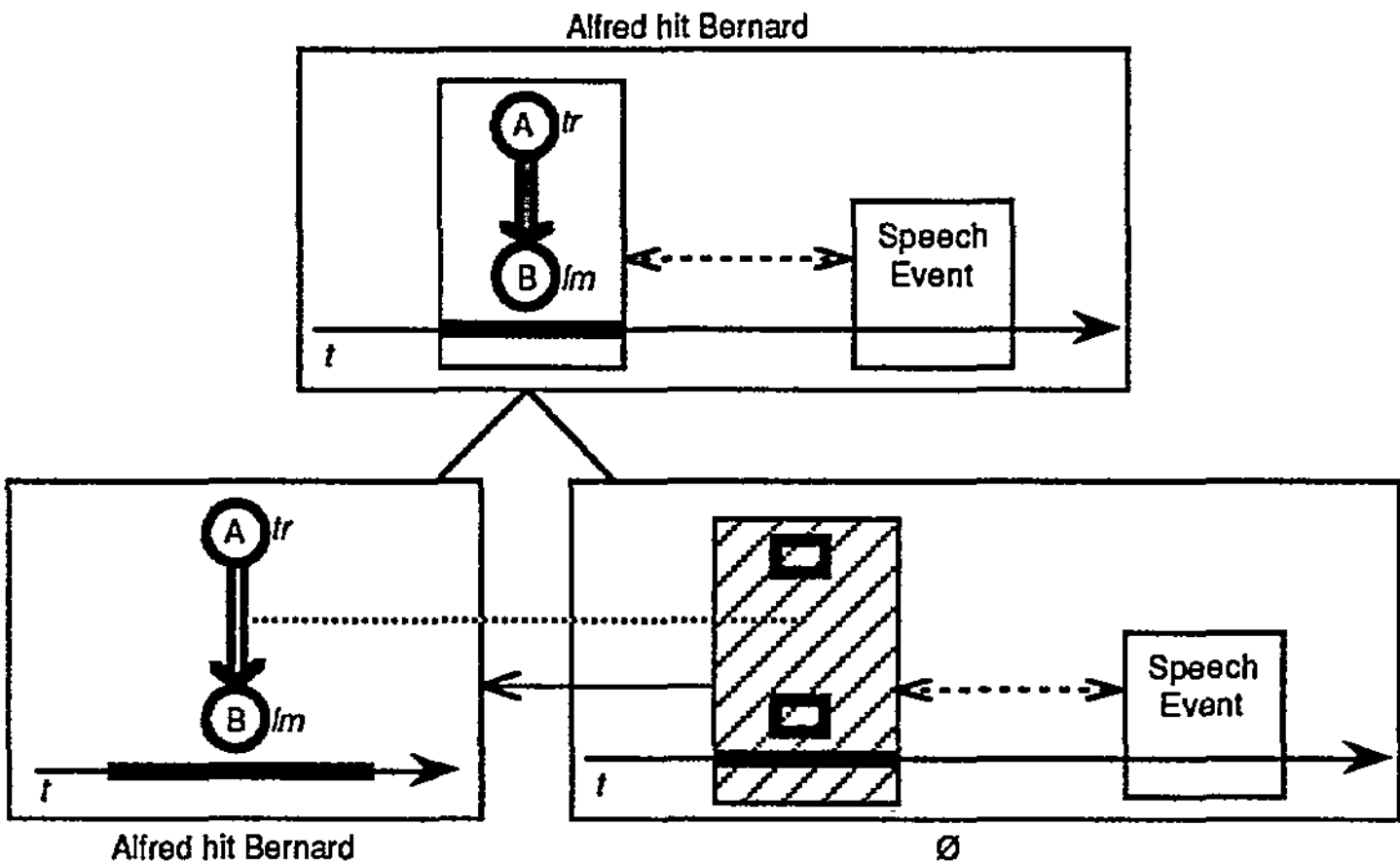


Figure 9

³Neither component structure is singled out as the profile determinant, since their profiles both correspond to the composite structure profile.

The *composite structure* at one level of organization is capable of serving as a *component structure* with respect to another, "higher" level of organization. Thus the composite structure from Fig. 6 appears in Fig. 9 as one of the two components, where it combines with *PAST* to yield a higher-level composite structure (a finite clause). The integration of *Alfred hit Bernard* and the past-tense element is effected by a correspondence between their processual profiles: the specific process designated by *Alfred hit Bernard* is equated with the schematic process that *PAST* profiles and situates prior to the speech event. This schematic process functions as an elaboration site, being specified in finer detail by *Alfred hit Bernard*³.

Fig. 10 depicts *Charles arrived*, which has two levels of grammatical organization: the "lower" level, where *Charles* combines with *arrive*; and the "higher" level, comprising the integration of *Charles arrive* with the past-tense marker. *Arrive* has only one focal participant, its trajector, which follows a spatial path (represented by a vertical arrow) and reaches a goal. *Charles* profiles a thing with certain properties (*C*). The integration of *Charles* with *arrive* involves a correspondence between its profile and the verb's schematic trajector, which it elaborates. *Arrive* is the head at this level, so the composite structure *Charles arrive* profiles the process of arriving, with Charles as the trajector. At the second level of organization, *Charles arrive* elaborates the schematic process profiled by the past-tense element. The higher-level composite structure thus designates the process of Charles arriving, temporally prior to the speech event.

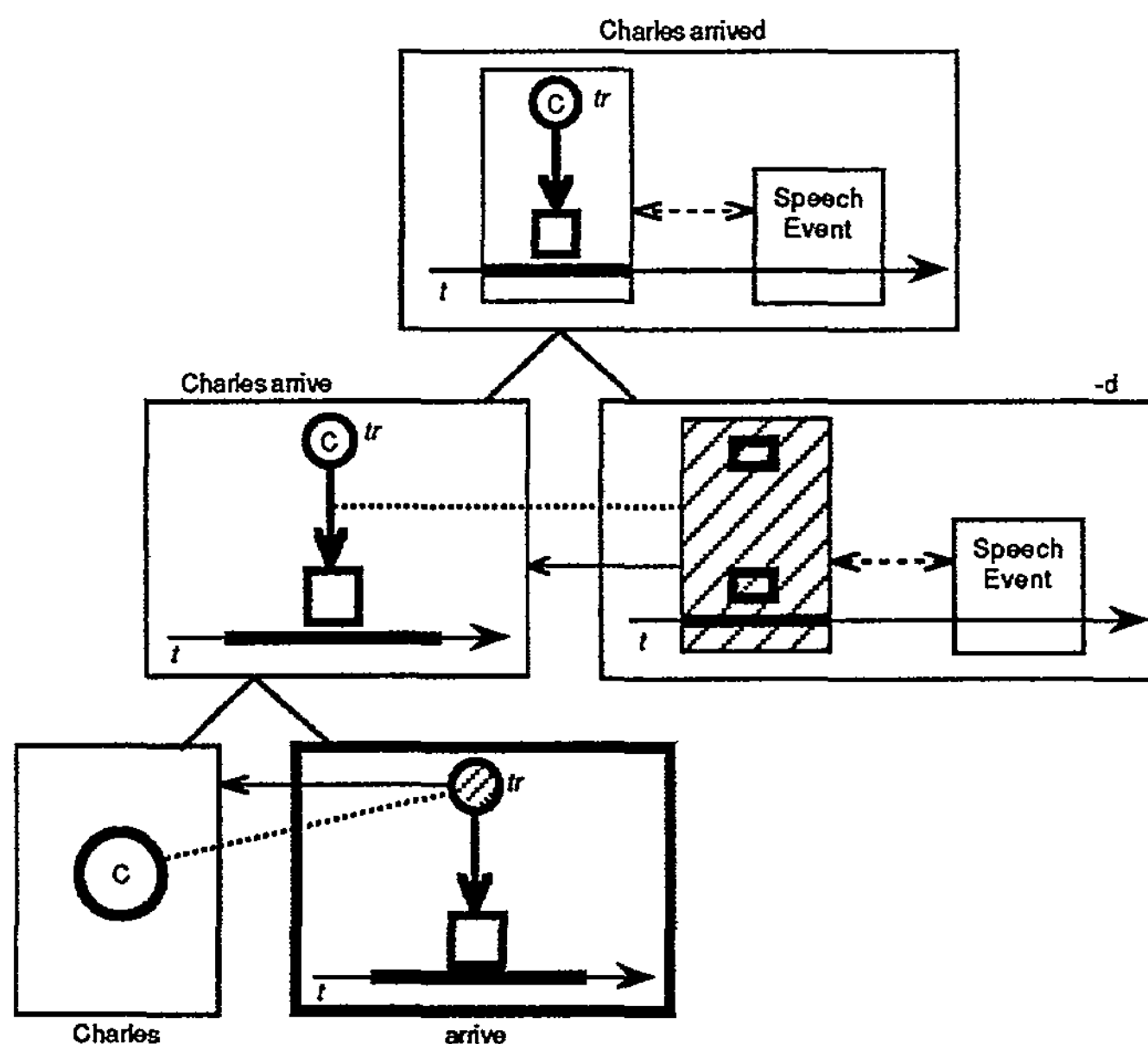


Figure 10

This finite clause combines with *before* to yield the complex adverbial expression *before Charles arrived*. Tesnière would analyze *before* as a *mot vide* and as a *translatif* that derives an adverb from a verb (Fig. 8), but it is clearly meaningful : it profiles a relationship of temporal anteriority between two events, the later event serving as landmark for locating the trajector event. I am happy to agree with Tesnière in referring to *before* as an adverb-deriving *translatif*, but not with his position that *translation* represents a distinct grammatical phenomenon radically different in character from other kinds of constructions. Indeed, the construction in Fig. 11 is quite comparable to any other examined here. *Before's* role as a *translatif* is simply a matter of its functioning as the *profile determinant* in this construction. Because it profiles a relation whose trajector is itself relational, *before* is properly regarded as a *schematic adverb*. It is integrated with a finite clause, such as *Charles arrived*, by virtue of a correspondence between its landmark and the clausal profile. This elaboration of the landmark renders the adverb specific. Conversely, *before* can be thought of as imposing its profile on the content provided by the processual component, thus converting (translating) a finite clause into an adverb.

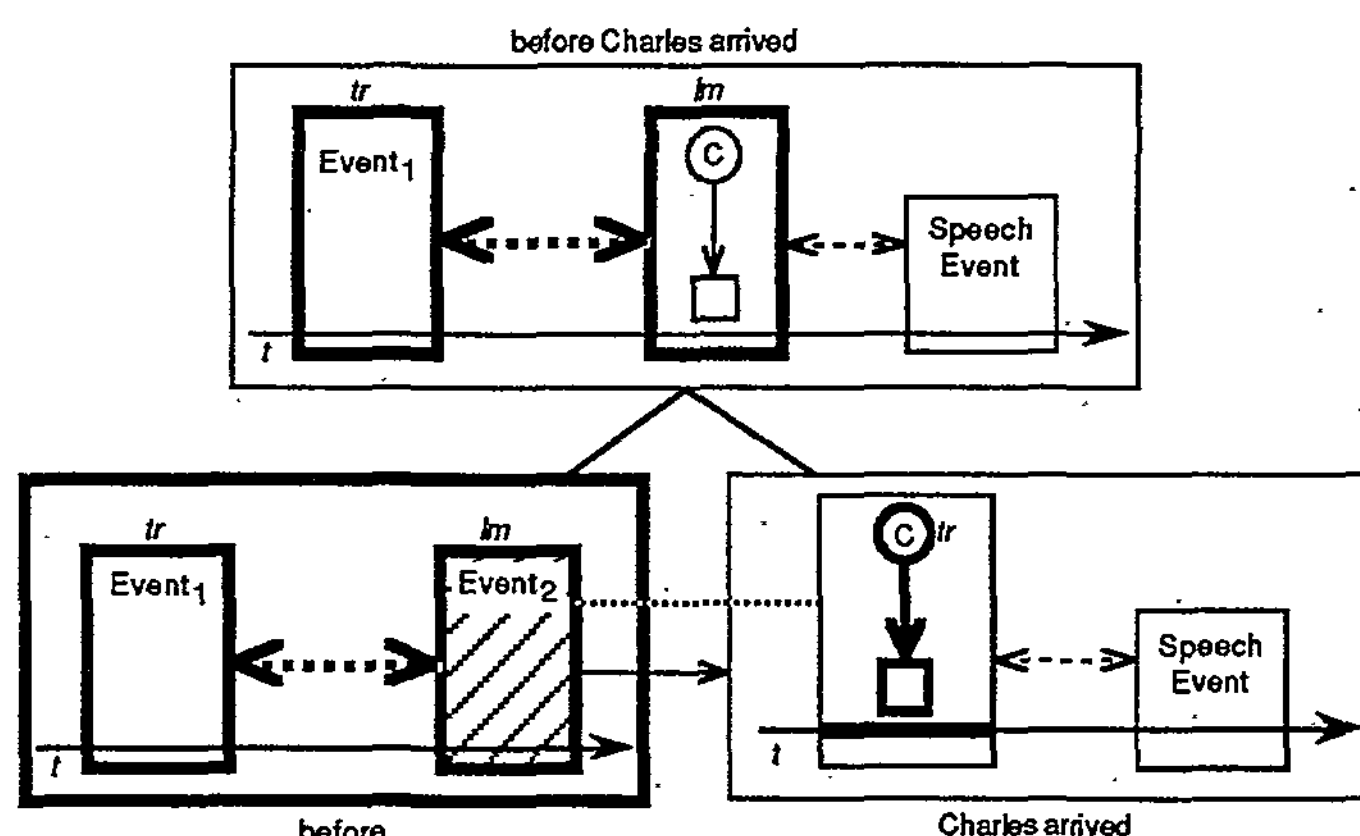


Figure 11

The last step is to integrate the finite clause *Alfred hit Bernard* with the complex temporal adverb *before Charles arrived*. Their composite structures thus function as the two component structures at this final level of organization (Fig. 12). The basic correspondence effecting their integration identifies the processual profile of *Alfred hit Bernard* with the schematic trajector of the adverb, which it elaborates. An additional

correspondence indicates that the speech events referred to by the two components are the same. *Alfred hit Bernard* is the head or profile determinant, since the process it designates is profiled by the sentence as a whole. This amounts to saying that it constitutes the main clause.

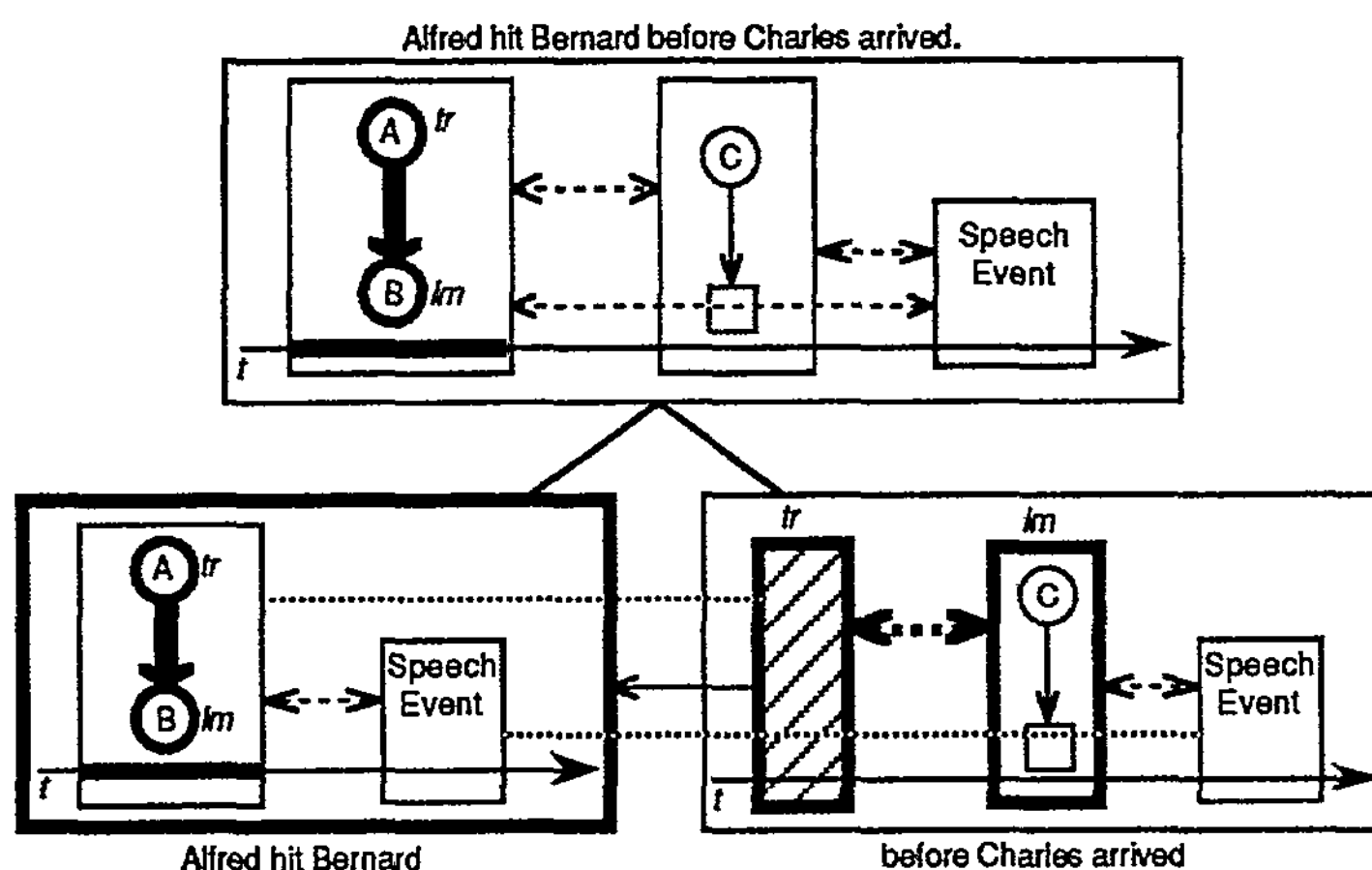


Figure 12

The grammatical organization of the entire complex sentence is summarized in Fig. 13 (see Fig. 13, on the opposite page).

The primary difference between Tesnière's structural syntax and CG is that only the latter attempts the explicit characterization of semantic structure. It describes not only the meanings of individual elements but also the *composite* semantic values arising at every level of organization as simpler symbolic structures combine to form progressively larger ones. I have tried to indicate that these semantic descriptions are necessary for a precise understanding of grammar itself. In the absence of specific descriptive constructs — such as profiling, trajectory/landmark alignment, thing vs. relation, correspondence, and profile determinance — one cannot provide an explicit account of the *semantic connections* that Tesnière recognizes as being co-extensive with *structural connections*. These same constructs figure in the semantic characterization of grammatical classes, markers, and basic notions like *head*, *main clause*, *subject*, *object*, *complement*, and *modifier*. A comparison of how the two theories deal with these classic notions, involving a discussion of Tesnière's distinctions *régissant* vs. *subordonné* and *actant* vs. *circonstant*, unfortunately cannot be undertaken in the space available here⁴.

⁴It is provided in the longer version of this paper, available on diskette.

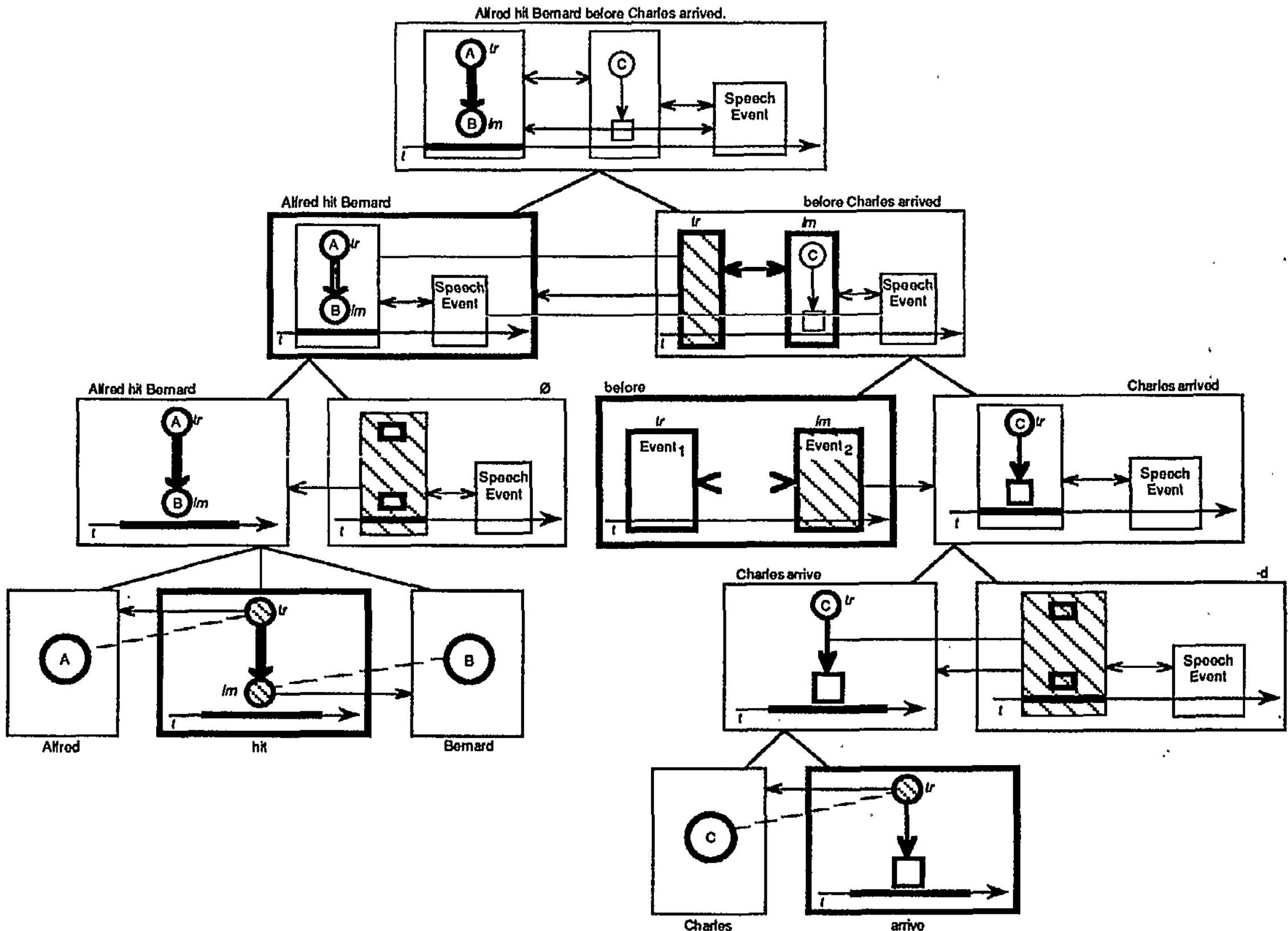


Figure 13

In a work as rich and comprehensive as Tesnière's, there is bound to be much to criticize from any modern perspective. What is far more surprising, from the standpoint of CG, is how few targets of serious criticism structural syntax offers, and how similar the two frameworks are in their basic mode of thinking about grammatical phenomena. I have tried to make it evident that Tesnière's research program was fundamentally compatible with that of cognitive grammar. Despite the difference in time and intellectual context, I feel we have been engaged in a common enterprise.

University of California (San Diego)

References

CHOMSKY (N.)

1957, *Syntactic Structures*, The Hague, Mouton, Janua Linguarum, 4.

FILLMORE (Ch. J.)

1968, "The Case for Case", p. 1-88, in *Universals in Linguistic Theory*, E. Bach and R. T. Harms eds., New York, Holt.1977, "The Case for Case Reopened", p. 59-81, in *Grammatical Relations*, P. Cole and J. M. Sadock eds., *Syntax and Semantics*, vol. 8, New York, Academic Press.

FUCHS (C.) & Le GOFFIC (P.)

1992, *Les Linguistiques contemporaines : repères théoriques*, Paris, Hachette.

GIVÓN (T.)

1984, *Syntax : A Functional-Typological Introduction*, vol. 1, Amsterdam, John Benjamins.

KEENAN (E. L.) & COMRIE (B.)

1977, "Noun Phrase Accessibility and Universal Grammar", *Linguistic Inquiry*, 8, p. 63-99.

KURODA (S. Y.)

1972, "The Categorical and the Thetic Judgment", *Foundations of Language*, 9, p. 153-185.

LAMBRECHT (K.)

1987, "Sentence Focus, Information Structure, and the Thetic-Categorical Distinction", *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, 13, p. 366-382.

LANGACKER (R. W.)

1985, "Observations and Speculations on Subjectivity", p. 109-150, in *Iconicity in Syntax*, J. Haiman ed., Amsterdam, John Benjamins.1987a, *Foundations of Cognitive Grammar*, vol. 1, Stanford University Press.1987b, "Grammatical Ramifications of the Setting/Participant Distinction", *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, 13, p. 383-394.1989, "Absolute Construal", p. 65-75, in *Worlds Behind Words : Essays in Honour of Prof. Dr. F. G. Droste on the Occasion of His Sixtieth Birthday*, F. J. Heyvaert and F. Steurs eds., Leuven University Press.1990, *Concept, Image, and Symbol : The Cognitive Basis of Grammar*, Berlin, Mouton de Gruyter.1991a, *Foundations of Cognitive Grammar*, vol. 2, Stanford University Press.1991b, "Noms et Verbes", p. 103-153, in *Sémantique Cognitive*, Cl. Vandeloise éd., Paris, Seuil.1992, "The Symbolic Nature of Grammar : The Meaning of *of* and *of of*-Periphrasis", p. 483-502, in *Thirty Years of Linguistic Evolution : Studies in Honour of René Dirven on the Occasion of his Sixtieth Birthday*, M. Pütz ed., Amsterdam, John Benjamins.

RICE (S.)

1987, *Towards a Cognitive Model of Transitivity*, doctoral dissertation, San Diego, University of California.

SHIBATANI (M.)

1985, "Passives and Related Constructions : A Prototype Analysis", *Language*, 61, p. 821-848.

SMITH (M. B.)

1987, *The Semantics of Dative and Accusative in German: An Investigation in Cognitive Grammar*, doctoral dissertation, San Diego, University of California.

TESNIERE (L.)

1965 (1959), *Éléments de Syntaxe Structurale*, Paris, Klincksieck.

Van HOEK (K.)

1992, *Paths Through Conceptual Structure : Constraints on Pronominal Anaphora*, doctoral dissertation, San Diego, University of California.

Dynamic Semiotics or the Case for Actantial Case

Franson Manjali

In this paper we shall attempt to sketch two theoretical approaches that view signification in terms of the constraints imposed on the structure of language by the essentially dynamic and constantly transforming nature of the world. These are : (1) the semiotic theory of René Thom based on the concept of morphogenesis, and (2) the *karaka* theory of the Sanskrit grammarians, particularly that of Bhartrhari (7th Century A.D.). By highlighting these approaches we are suggesting that semiosis can be seen as a matter of reflecting/revealing the infinite dynamism of the world by means of a finite variety of basic sentence structures. And, at the same time, we are pointing to the inadequacy of the static semiotic paradigms.

Ferdinand de Saussure : Discrete Signs

We shall start with a reference to F. de Saussure who proposed a notion of sign as an inseparable bipartite entity constituted of the signifier and the signified. Saussure saw the *langue* (language system) as made up of discrete signifying units or signs defined in terms of their relations and mutual differences, and which enter into acceptable combinations in language use (*parole*). While the sound-form and the thought, mediated by language, are continuous and 'nebulous' in nature, language in itself is constituted of discrete signs. Saussure excludes from the realm of language the undivided streams both of thought and of sound-form. Important to this conception of language was the discretization of both the signifiers and signifieds and the modes of reconstitution of the formal and semantic continuity by means of syntagmatic combinations. Syntagmatic and associative/paradigmatic relations "are two forms of our mental activity, both (of which) are indispensable to the life of our languages" [Saussure, 1974 (1916), p. 123].

We notice that Saussure is upholding the widely held belief that language is a rule based system of discrete symbolic units and their combinatorial behavior. Thus, even while insisting on the complete autonomy of language Saussure readily accepts the view that "language, in a manner of speaking, is a type of algebra consisting solely of complex terms" [Saussure, 1974 (1916), p. 122].

As regards the nature of the combinations of signs, Saussure appears to be far less committed. The temporal order of the spoken language imposes on it a character of linearity, and this necessitates the sign units to be "linked together." Syntagms are "combinations supported by linearity" [Saussure, 1974 (1916), p. 123]. Here, indeed there is a paradox that Saussure himself reveals to us : while syntagms are combinatorial constructs defined by reciprocal occurrence, "the sentence is the ideal type of syntagm" [Saussure, 1974 (1916), p. 124]. However, the latter belongs to speaking and not to the language system. Thus, at the level of combinatorics Saussure perceives a continuum of more or less constructional rigidity, the least rigid syntagmatic unit being the sentence, which indeed is not a unit of the language system, but of speaking. Saussure's solution is as follows :

"(...) in the syntagm there is no clear-cut boundary between the language fact, which is sign of the collective usage, and the fact that belongs to speaking and depends on individual freedom. In a number of instances it is hard to class a combination of units because both forces have combined in producing it, and have combined in indeterminable proportions" [Saussure, 1974 (1916), p. 124].

Just as he has an excellent sense of the sign as the basic, independent unit of language, Saussure is also conscious of the coexistence of signs in a totality :

"Language is a system of interdependent terms in which the value of each term results solely from the simultaneous presence of the others (...)" [Saussure 1974 (1916), p. 114].

A language-totality is thus the sum of all its sign-units, and their relations, both syntagmatic and paradigmatic.

Perhaps, by overstating the autonomy of the language structure, Saussure remains insensitive to the specific structuring of the sentence, both at the syntactic and semantic levels. While concentrating on a description of the individual signs and the language-totality, Saussure appears to have paid less attention to the syntactico-semantic constitution of the sentence. The notion of sentence, we know, has been central for linguistics of the classical period, both in European and Indian traditions. In Europe — for those who insisted on its centrality — the sentence had been seen as the *minimal unit of expression of a complete thought*,

containing the subject and the predicate components. In India, there were profound and meaningful debates between scholars who held that sentence conveyed undivided meaning (*akhanda-pakshavada*) and those who held that sentence meaning is a result of the combinatorics of word-meanings (*padavada*). Bhartrhari, we know, was a firm adherent of the former position.

Though the understanding of language structure in terms of syntagmatic and paradigmatic relations is extremely useful, it is still important to perceive the hierarchical organization of the language units, which the generative grammar of Chomsky can best capture. Language is a system where the multiple levels of organization of form and meaning are masked by a surface linearity. Sentence is not only the highest level of this hierarchy, but also, in relation to thought, the bounding structural unit. (Beyond the sentence, of course, there is the textual level which may also have its hierarchical organization, for instance, of the narrative units.) Etymologically speaking, a 'sentence' expresses what is felt or thought ('sentir'). In the Aristotelian conception, language is a mode of representing or imitating reality (*mimesis*), involving the use of a subject-predicate structure. Other modes of representation such as painting, music or drama do not have this particular structure, and hence cannot be evaluated in relation to the truth or the falsity of the representation.

Lucien Tesnière : The Actantial Paradigm

While this unit-to-unit correspondence between language and the world has been the main parameter of truth in the Greek tradition, we can also speak of a figure-like adequation of language in relation to reality. Thus, in addition to the logical/propositional value of the sentence implicit in the former, philosophers and linguists have considered sentence as a mode of reflecting events in the world in a somewhat pictorial manner.

Lucien Tesnière, for instance, has proposed such a view in his *Éléments d'une syntaxe structurale* [Tesnière, 1959]. The so-called 'dependency' grammar of Tesnière is based on an implicit notion of 'action' which was well known to the Indian grammarians. For Tesnière, the meaningfulness of a sentence was due to the central organizing role of the predicate verb which represented an action, and functioned as the highest syntactic node of the sentence. The verb is the complete and the independent term of a sentence. Dependent on the verb are the 'actants' which are the participants in the action (this dependency relation is to be diagrammatically represented by means of a tree-structure or 'stemma'). Tesnière viewed the sentence as representing a 'little drama' (*petit drame*) wherein the predicate represents an action (in the theatrical sense), or even a process, and the dependents of the predicate are the principal elements in

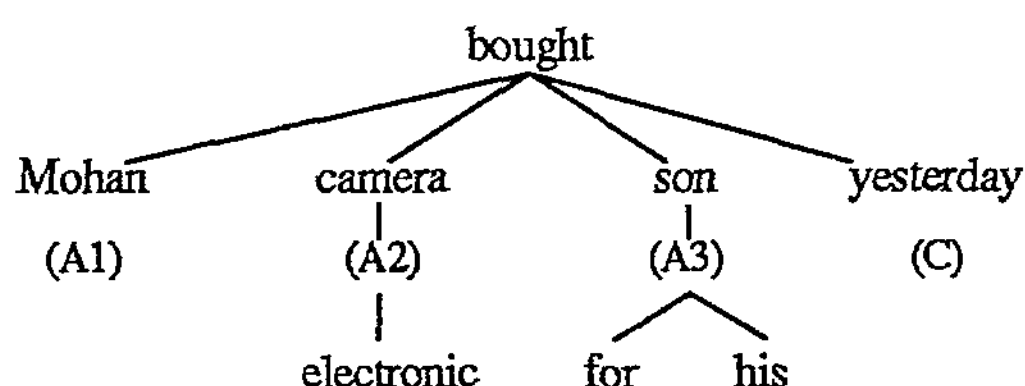
the action. Since Tesnière is distancing himself from a logical conception of grammar he is eschewing the "subject-verb-object-indirect object" type of propositional structure ; he is opting for a rather theatrical conception where the nominal elements are initially non-heterogeneous actants in participating in a process, but appearing in their functionally specialized roles as subject, object and indirect object in the context of the sentence-structure. Tesnière defines actants as "beings or things which in some capacity and in whatsoever manner, even in the capacity of mere onlookers and in the most passive manner participates in a process" [Tesnière, 1959, p. 102].

While actants are one type of dependents of the predicate (they designate characters in a anthropomorphic sense), the other type called the circumstants designate situations. According to Tesnière there can be a maximum of only three actants in a sentence while the circumstants may be several. The following example may be provide an illustration :

— Sentence :

"Mohan bought an electronic camera for his son yesterday".

— Stemmatic representation :



Here, A1 (= 'Mohan') is the Subject, A2 (= 'Camera') is the Object of the transitive verb or the Agent of the passive verb, A3 (= 'son') is the Beneficiary, and (C = 'yesterday') is the Circumstant.

Tesnière's dependency grammar is a kind of case-grammar that shows the semantic roles of sentence constituents. We may note that he had also introduced the notion of "valency" to denote the number of actants carried by a verb. Thus the valency could be zero ('rain'), one ('cry'), two ('hit') or three ('give'). (We shall note in passing that Tesnière's actantial paradigm has been adopted by A. J. Greimas in developing his semio-narrative grammar).

Despite their apparent similarity, Tesnière's stemma is different in content from Chomsky's tree-diagram. While in the latter, the connections between the nodes have no theoretical value, in the former these connections are perceived in an organic way, that is, as the connections between the participants in an action. The stemmas are the diagrammatic representation of a holistic image of the meaning of the sentence

conceived as action. They are suggestive of the sentence-meaning as some sort of dynamic gestalts.

Tesnière, we notice, is stressing on the notion of a semantic continuum which is unknown to Saussure :

“Every word which forms part of a sentence ceases itself to be isolated as in a dictionary. Between it and its neighbors, the mind perceives connections whose ensemble forms the framework of the sentence. (...) These connections are indicated by nothing” [Tesnière, 1959, p. 11].

For example in the sentence, “Mohan speaks,” there are three elements : (1) Mohan (2) speak and (3) the connection which unites the two first elements, and without which they would not form a sentence. Not to account for the connections “is to ignore the essential, which is the syntactic link.” “The connection is indispensable for the expression of thought. Without connection, we will not be able to express any thought, and we will only be utter a succession of images and indices, isolated from each other, and without any link between them” [Tesnière, 1959, p. 12]. On the importance of connections for Tesnière, Jean Petitot remarks :

“(...) a sentence is above all a system of connections which being ‘incorporeal’ (non sensible) can only be grasped by the ‘mind’. These structural connections, oriented and hierarchised, are not of logical essence, but constitute an ‘organic and vital’ principle of organization (...)” [Petitot, 1985, p. 45].

They constitute the ‘vital’ organic principle of the linguistic ‘energeia’ in the sense of Humboldt (see [Petitot, 1989, p. 182]).

While introducing that actantial perspective and the notion of structural connections, Tesnière seems to be arguing for the semantic continuum. Of course, he explicitly, supported a organicist and holist conception of the sentence. The advantage of such a position is that it permits us to think of structural space where the actants are related to each other via the activity referred to by the verb. Among other scholars who have maintained similar views is the Russian linguist, S. Katznelson who, while noting the fragmentary nature of words as against the holistic character of sentence observes that it is the “grammatical elements (...) (that) reestablish the living links which full words tend to lose when they are withdrawn from the images of coherent events” (Katznelson, 1975, p. 102).

René Thom : Catastrophe Theory

Tesnière’s fundamental ideas of actant and valency as well as the organicist perspective has much influenced the semiotic/semantic thinking

of mathematician René Thom known for his Catastrophic Theory. The central role assigned to the verb is also a common factor between the systems of ideas of Tesnière and Thom.

Thom's natural/realist philosophy is governed by the two central principles of structural stability and morphogenesis. The importance of Catastrophe Theory to Linguistics and Semiotics comes from the fact that it is most directly concerned with structures.

The theory has essentially to do with the effect of local (quantitative, micro) variations on the global (qualitative, macro) structure. Catastrophe theory involves the description of the (sudden, abrupt) discontinuities induced by the continuous local perturbations of a system. As per Thom's theorem :

"The number of qualitatively different configurations of discontinuities that can occur depends not on the number of state variables, which is generally very large, but on the control variables, which is generally very small. In particular, if the number of control variables is not greater than four, then there are only seven types of catastrophes, and in none of these more than two state variables are involved" [Saunders, 1980, p. 3].

The seven elementary catastrophes are : Fold, Cusp, Swallowtail, Butterfly, Elliptic Umbilic, Hyperbolic Umbilic and Parabolic Umbilic ; these have their corresponding topologies.

It is indeed possible to note the striking parallel between the elementary catastrophes and the cases found in the classical languages. We may perhaps speak of the evolution of sentence structures (that is, their morphogenesis) parallel to and part of the morphogenesis of natural forms.

Following the German biologist, Jacob von Uexküll, Thom believes that basic concepts originate as a function of the biological self-regulation involving the prey, the predator, and (sexual) partner. The more complex concepts are built upon these :

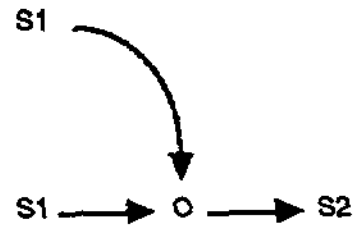
"The logos of living beings has served as the universal model for the formation of concepts".

These biologically founded and linguistically valid concepts play the role of actants in the interactions represented by the verb. The verb is an organizing centre, i.e., the event that distributes the actantial places. Verbs are identified by their own structural stability as events. They have as their source and model the simulation of elementary actantial interactions realizable in space-time. Perceptually, these interactions are constrained by the four dimensions of space-time.

Thom has identified 18 such 'morphologies' of interaction, referred to as 'archetypal morphologies' by himself, and as 'semantic archetypes' by [Wildgen, 1982] (see [Thom, 1980 (1974), p. 213] for the list of archetypal morphologies).

It is possible to describe the morphogenesis of sentence-structures by projecting the actants on a substratum space, and by assuming their

'inflections' according to the increasing in the number of actants and the evolution of interactions in time. A zero-valent verb will occupy the entirety of the perceptual space. A uni-valent verb evokes the possibility of an action continuous in time. A bi-valent verb will involve interactions, and its graph will show a discontinuity (at a zone of interaction) between its earlier and later parts of the event described in time. Thom's example of such a verb is 'capture' whose actantial graph is :



where S1 and S2 are the actants and O the point of interaction.

Bhartrhari : Karaka Theory

The actantial perspective which was unknown to Saussure, and which occupies a prominent position in the linguistic thinking of Tesnière and Thom (and also Greimas), was of preeminent importance to the Indian grammarians. The key Sanskrit term for what seems like a parallel idea is '*karaka*' (doer or actor), which has been part of the Indian grammatical vocabulary for centuries.

The *karakas* are recognized by most scholars as basic semantic notions that in fact pivot sentence constructions. They are similar to the case roles/relations proposed in the case grammars. But *karakas* are much more than these, and their crucial role as a common substratum of ontology, cognition, and grammar can be understood only if we regard them as a manner of classifying 'actions' in the real world.

It may not be however inappropriate to suggest that the *karaka* notions are conceived as properties of the world corresponding to, though independent of their grammatical/morphological manifestations. Panini himself was probably merely projecting the *karakas* (literally, 'a factor of action') from morphological occurrences in the form of cases to a set of possible actions in the world. This point has been most aptly made by a recent commentator :

"If the notion of *karakas* was perhaps derived from an observation of Sanskrit cases, Panini had raised them above the level of case values and made them intermediaries between reality and the grammatical categories. Their importance, often misunderstood, goes far beyond the syntax of cases ; next to the roots, they are the prime moving factors of the whole grammar" [Scharfe, 1977, p. 95].

Panini identifies six *karakas* corresponding to six cases, viz., the nominative, accusative, dative, instrumental, locative, and ablative. Possessive and vocative are conspicuous by their absence in Panini's

grammar. This is how Panini defines the six *karakas* (*Asthadhyayi*, I. 4. 24-54) :

1. *Apadana* (lit. 'take off') : "(that which is) firm when departure (takes place)." This is the equivalent of the *ablative* notion which signifies a stationary object from which a movement proceeds.

2. *Sampradana* ('bestowal') : "he whom one aims at with the object". This is equivalent to the *dative* notion which signifies a recipient in an act of giving or similar acts.

3. *Karana* ('instrument') "that which effects most". This is equivalent to the *instrumental* notion.

4. *Adhikarana* ('location') : or "substratum". This is equivalent to the *locative* notion.

5. *Karman* ('deed'/'object') : "what the agent seeks most to attain". This is equivalent to the *accusative* notion.

6. *Karta* ('agent') : "he/that which is independent in action". This is equivalent to the case of the subject or the *nominative* notion (on the basis of [Scharfe, 1977, p. 94]).

In his *Mahabhashya*, Patanjali defines *karakas* in relation to the notion of *kriya*, or action. Action is the 'distinctive mode of action of the accessories'. He also held that : (1) the root can be defined as something which expresses *kriya*, (2) *kriya* is different from all the accessories which play a part, direct or indirect, in its accomplishment, and (3) the action is not *pratyaksha* (perceptible), it can only be inferred.

That verbs primarily convey 'action' is an idea that goes back to Yaska's etymological studies called the *Niruktas*. In his words : *bhavapradhanam akhyatam*, "an action or process is the main meaning of a verb" (see [Subramania Iyer, 1969, p. 202]).

Bhartrhari discusses various possible definitions of 'action', but what he prefers is the following :

"Whenever something, finished or unfinished is presented as something to be accomplished (i.e., *sadhya*), then it is called 'action' because of its having acquired the form of sequence" (III. 8. 1).

In addition, he stands by Patanjali's definition as per which "action is the distinctive mode of behaviour of the accessories", and appears to be rejecting another view as per which "action is that moment immediately after which the result is produced" (e.g., in 'cooking' there is a critical moment that separates the cooked state from the raw state, of rice).

The fact that action is something which has the form of 'parts arranged in sequence' would entail that it cannot be directly perceived. It can only be *inferred* by the mind. The relevant statements in *Vakyapadiya* on this are the following :

“What is called action is a collection of parts produced in a sequence and mentally conceived as one and identical with the parts which are subordinated to it (i.e., the whole)” (III. 8. 4).

“The parts which occur in a sequence and are partly existent and partly not so cannot enter into contact with the senses like the eyes whose objects are always the existent” (III. 8. 6).

Thus the action of ‘cooking’ can be seen to involve a number of subordinate actions. One may however question, if ‘cooking’ consists of the whole sequence of parts of actions perceived as whole or only the moment of transformation of the raw rice into cooked (soft) state. Bhartrhari seems to prefer the former position.

Bhartrhari’s philosophy of grammar is accompanied by a corresponding ontology describing the universe in terms of objects, forces or powers (*shakti*) and the interactions between the objects.

The essence of the universe is understood as comprising infinite powers distributed in an infinite number of objects. Due to the powers the objects are constantly changing and are in mutual interaction with other objects, resulting in their eternal mutation. The universe is thus in a constant and perpetual state of change. Yaska had identified six basic types of transformations, namely *birth*, *existence*, *change*, *increase*, *decrease*, and *death*. Actions and processes in the world result from the changes and interactions of objects. Moreover, objects themselves are the (temporarily stable) state (*siddha*) produced by the actions and processes (*sadhya*). The objects which are distributed in space (*dik*) are themselves participants in further actions and processes. In addition to spatial location, the objects have other powers resting in them, which may be the result of past actions, and which are involved in the subsequent actions. *Sadhana* is Bhartrhari’s term for these powers.

Sentences are the linguistic mode of capturing certain particularized actions abstracted from the eternal play of forces in the universe. A sentence represents a “complex meaning (we may say, a dynamic gestalt) in which some action or process is the central element and concrete objects which cooperate in accomplishing the process are the elements associated with it” [Subramania Iyer, 1969, p. 285].

Though the powers vested in the objects may be of an infinite variety, from the linguistic point of view (that is, in language) they are *classified* into six different kinds of capacities in which an object can participate in an action represented by the sentences. *Karakas* are the names for these limited number of capacities. The six *karakas* identified in the Sanskrit grammar are *karma*, *karana*, *karta*, *adhikarana*, *apadana*, and *sampradana*; in addition there is a sundry *karaka* category referred to as *sesa*.

The main feature of Bhartrhari’s ideas is the constancy of and the omnipresence of transformations in the universe. Both word and the world are the result of manifest transformations and/or apparent differentiations

of a cosmic unity, which takes the name of *sabdabrahma*. (Brahman, or the Ultimate reality is of the nature of the word, i.e., *sabdatattva*). From an eternal point of view these transformations/differentiations are unreal and illusory. Time, as one of the properties of the unchanging cosmic unity, is the material force which produces these transformations, which are in turn perceived and cognized as activities of particular things. The *sabdabrahma* is initially differentiated into its mental and material media, and the time-force (*kalashakti*) affects both these aspects.

In Bhartrhari's view, only the sentence can completely express 'reality', and not the word which may denote objects. Moreover, "reality is expressible only in the form 'it exists' which means that a word in order to express a reality has to be compounded with a verb, namely 'exists'".

Therefore, "a verb has to be part of a sentence (...). If the verb is mentioned as expressing an action to be conveyed, nouns are required to effect the action" [Bhartrhari, 1971 trad., xxxiii]. The verb constitutes the essential and minimal content of a sentence [Bhartrhari, 1971 trad., xxxiv].

Sentence-meanings which are primarily in the nature of an action are also relative to the speaking subject :

"The grammarian makes a distinction between word-meanings which mention an object, and the meaning of the sentence which is primarily an action, effected by men through objects" [Bhartrhari, 1971 trad., xxxiii].

For the grammarian, reality is understood only through speech (language) and it is understood only in the form it is presented by speech (word/language). *But language cannot describe the intrinsic nature of things, although we know things only in the form in which words describe them* [Bhartrhari, 1971 trad., xxxiii].

Bhartrhari rejects the existence of meanings of individual words. Individual word-meaning is an illusion, according to him. Only the undifferentiated sentence-meaning is real. The sentence-meaning is not a concatenation of word-meanings as argued by the Mimamsaka philosophers, but to be understood in terms of a complex cognition. Bhartrhari compares this complex cognition with that of the cognition of a picture (*citrajnana*) :

"A cognition which embraces many objects at the same time is a complex cognition. As a cognition, it is one but because of the many objects which figure in it, one sees plurality in it, though it is indivisible" [Subramania Iyer, 1969, p. 186-187].

The relationship between sentence-meaning and word-meaning is compared to the relationship between a holistic picture and its component parts. Just as a unified perception of composite (picture) can be analyzed (into the preoccupation of component parts) depending upon which part is required to be perceived so likewise is the understanding of the meaning of the sentence and "just as a single homogeneous picture is described

through various features as being blue (green, etc.) as a result of its being perceived in different ways, similarly the sentence which is single and does not possess expectancy is described in terms of words which possess mutual expectancy" [Bhartrhari, 1971 trad., p. 38].

We shall summarize Bhartrhari's views on the sentence and its meaning in the following manner. The sentence represents/reveals at least a fragment of the eternal activity in the universe, presented from the point of view of the speaker. The verb highlights the specific character of this activity, expressed in terms of the accessories/means and their qualities. When a thing is expressed as something to be accomplished, it is *sadhya*, but when it is expressed as accomplished, it is *siddha*. The means involved in the accomplishment of an action are the *sadhanas*. The recognition of the sentence-meaning takes place by way of the *vakya sphota* implying a somewhat gestalt-like comprehension.

Subramania Iyer points out that :

"The complete meaning expressed by a sentence is a complex thing in which some process of action occupies the central position and is associated with its accessories and their qualifications, all amalgamated into an indivisible whole" [Subramania Iyer, 1969, p. 200].

And :

"(...) the indivisible sentence is the unit of communication and its meaning is understood in a flash (*pratibha*). This meaning is also something indivisible, a complex cognition in which the central element is an action or process with its accessories closely associated with it" [Subramania Iyer, 1969, p. 201].

Bhartrhari's conception of the case relations (*karakas*) can be understood in the following terms. A sentence represents/reveals the accomplishment of an action. 'Means'/accessory (*sadhana*) is the power (*shakti*) of a thing to accomplish actions (*Vakypadiyam*, III. 7. 1). The difference in the powers of objects is relative to the form that speakers (subjectively) impose on them (III. 7. 6). Each object that is involved in any action in any and at any time is seen as having a particular means or power for that time (III. 7. 12). The particular help rendered to the action is expressed by the case-markers (III. 7. 13). *Karaka* (literally, 'doer' or even an 'actant' in the sense of Tesnière) is that which helps in the accomplishment of an action by assuming different forms (thus *karaka* is different from both *hetu* (cause) and *lakshana* (sign) which are relatively more world-based). It is said that in any one object, there can be six different powers that lead to action. These powers are universal, and though they appear to be unlimited, can only be six in number (III. 7. 35-36). These six powers correspond to the six *karakas*.

Karta (agent) is the basis of all the varied activities (because power is one, but appears to be divided into six kinds according to circumstances).

There are six more 'karakas' including *sesa*, or the 'rest'. These are : *karma*, *karana*, *adhikarana*, *sampradana*, *apadana* and *sesa* (which includes *sambodhana*) (III. 7. 37-44).

The object (*karma*) 'that which is most desired to be attained' is of three kinds : product (*nirvartya*) : {He made jar out of mud} ; modification/conversion (*vikarya*) : {He converted wood into ashes} ; and destination (*prapya*) : {He saw a tree (here the object doesn't change) } (III. 7. 47-51).

Whenever, after the activity of something, the action is meant to be conveyed as accomplished, then that thing is said to be the instrument (*karana*) (III. 7. 90). Thus, the instrument is a more immediate participant in an action than the agent itself.

The factor in the act of which is sought to be reached by the thing given is called *sampradana* when he does not prohibit the giver, or request him, or gives his consent (III. 7. 129).

A starting point (*apadana*) is of 3 kinds : that in relation to which a movement is mentioned ; that in relation to which the verb expresses the movement only partly ; and that in relation to which some movement is required (III. 7. 136).

That which helps in the accomplishment of the action by holding it indirectly through the agent is called *adhikarana* (abode). The contact is the same whether the abode be sesame seed, the sky or mat, etc. But the service rendered differs according to as the objects are in contact through *samyoga* (conjunction) or through *samvaya* (inherence) (III. 7. 148-149).

In addition to the six *karakas* listed above there is discussion of a possible set of cases under the name of *sesa* (the rest or the extras). *Sesa* does not represent a *karaka* relation but may involve or be preceded by one of the *karaka* relations. Under this category, Bhartrhari discusses the possessive case, where the relation of possession is supposed to be preceded by some sort of action, e.g., 'king's man' implies an action on the part of the king which has led to the establishment of a master-servant relationship. Further, in expressions like 'branch of the tree' and 'father's son', the relations like the part and whole, and procreator and offspring are "the results of previous actions not mentioned in the sentences, actions in which these objects were accessories. That previous status lingers somewhat in the present status and that is why the present status is looked upon as a kind of *karaka*, though its relation with the action expressed in the sentence is rather remote" [Subramania Iyer, 1969, p. 325].

Sambodhana or vocative is discussed separately. Merely turning the attention of somebody already there towards oneself has been declared to be the nature of the vocative case (*sambodhana*). It is indeed, one whose attention has been attracted that is employed in some action (III. 7. 163).

References

BHARTRHARI

(7th century A.D.)

1971, *The Vakyapadiya*, Critical Texts of Cantos I and II with English Translation, Summary of Ideas and Notes by K. Raghavan Pillai, Delhi, Motilal Banarasidas.

1971, *The Vakyapadiya*, ch. III, part i, Eng. Translation by K.A. Subramania Iyer, Poona, Deccan College P.R.I.

1977, *The Vakyapadiya*, Eng. Translation with Exegetical Notes by K.A. Subramanya Iyer, Delhi, Motilal Banarasidas.

MANJALI (F.)

1982, "Ontology and Meaning : Some Considerations on the Category of Case", *International Journal of Communication*, vol. II, p. 152-169.

PETTITOT (J.)

1985, *Morphogenèse du Sens*, Paris, PUF.

1989, "On the Linguistic Import of Catastrophe Theory", *Semiotica*, 74, 3-4, p. 179-209.

SAUNDERS (P. T.)

1980, *An Introduction to Catastrophe Theory*, Cambridge University Press.

SAUSSURE (F. de)

1974 (1916), *Course in General Linguistics*, Eng. Translation by W. Baskin, Glasgow, Fontana.

SCHARFE (H.)

1977, *Grammatical Literature*, vol. V, fasc. 2, in *History of Indian Literature*, Jan Gonda ed., Wiesbaden, Otto Harrassowitz.

SUBRAMANIA IYER (K. A.)

1969, *Bhartrhari : a Study of Vakyapadiya in the Light of Ancient Commentaries*, Poona, Deccan College P. R. I.

TESNIERE (L.)

1959, *Éléments d'une syntaxe structurale*, Paris, Klincksieck.

THOM (R.)

1980 (1974), *Modèles mathématiques de la Morphogenèse*, Paris, UGE (Nouvelle édition Ch. Bourgois).

UEXKÜLL (J. von)

1982, "The Theory of Meaning", *Semiotica*, 42, 1, p. 25-82.

WILDGEN (W.)

1982, *Catastrophe Theoretic Semantics*, Amsterdam, J. Benjamins.

Frameworks for Properties

Possible Worlds vs. Conceptual Spaces

Peter Gärdenfors

1. Program

I would like to argue that the traditional kind of model theory used in intensional semantics is not the right kind of framework for its purpose. Instead, I shall put forward some ideas from what is becoming known as *cognitive semantics* and try to show that this kind of framework has a better chance of doing the job.

My discussion will be focussed on the notion of *property*. I first present the standard intensional definition of a property, which is formulated in terms of possible worlds, and outline some of the philosophical problems this definition leads to. Then I introduce the notion of *conceptual space* and show how such spaces can be used as a basis for a new criterion of what a property is. This criterion will be shown to elude the problems of the traditional approach. Furthermore, it will be argued that the criterion is useful for understanding *prototype effects* of properties and the role of properties in *non-monotonic reasoning*.

Even though I concentrate on the notion of property in this paper, I believe that conceptual spaces have a more general bearing for the development of a cognitive semantics capable of handling intensional notions. Here I will only give some general remarks in that direction.

2. The traditional definition of a property and its problems

What is a property? Intuitively, it is something that objects can *have* in common. Often we can *perceive* that an object has a specific property.

These intuitions are not mirrored in the traditional definitions of the notion of property. In the classical *extensional* type of definition, like in Tarski's model theory for first order logic, a property is defined as *a set of objects*, to wit, the set of objects which has the property. Formally, this is done with the aid of a mapping from a language *L* to a model structure,

representing a world, where each one-place predicate in L is mapped onto a subset of the individuals in the model structure.

However, it was apparent that many so-called intensional properties did not fit this definition. A typical example is 'small': an emu is a bird, but a small emu is not a small bird; hence, the property of being small cannot be identified by a set of 'small' objects.

In order to handle the problems concerning intensional properties and other intensional concepts, the classical semantic theories were extended to so-called *intensional semantics*. Pioneers in this development were [Hintikka, 1961], [Kanger, 1957], and [Kripke, 1959], and as an analysis of natural language it reaches its peak with [Montague, 1974]. Here the language L is mapped onto a *set of possible worlds* instead of only a single world. Possible worlds and their associated sets of individuals are the only primitive semantical elements of the model theory. Other semantical notions are defined as *functions* on individuals and possible worlds. For example, a *proposition* is defined as a function from possible worlds to truth values. Such a function thus determines the *set of worlds* where the proposition is true. According to traditional intensional semantics, this is all there is to say about the meaning of a proposition.

In this kind of semantics, a property is something that relates individuals to possible worlds. In general terms, a property can be seen as a many-many relation P between individuals and possible worlds such that iPw holds just when individual i has the property in world w .

In intensional semantics, functions are preferred to many-many relations. There are two ways of turning the relation P into a function. Firstly, it may be described as a *propositional function*, i.e. a function *from individuals to propositions*. Since a proposition is identified with a set of possible worlds, this means that a property is a rule which for each individual determines a corresponding set of possible worlds. But we can also turn the table around to get an equivalent function out of P : for each possible world w , a property will determine a set of individuals which has w as an element of the sets of possible worlds the individuals are assigned (cf. Figure 1, on the opposite page). This means that an equivalent definition of property is that it is *a function from possible worlds to sets of individuals*. This alternative definition shows the correspondence between the extensional and the intensional definition of a property.

I now want to show that the standard definition of a property within intensional semantics leads to a number of serious problems. First of all, this definition is highly counterintuitive since properties become very abstract things. Bealer has the following remarks:

"How implausible that familiar sensible properties are functions — the color of this ink, the aroma of coffee, the shape of your hand, the special painfulness of a burn or itchiness of a mosquito bite. No function is a color, a smell, a shape, or a feeling" [Bealer, 1989, p. 1].

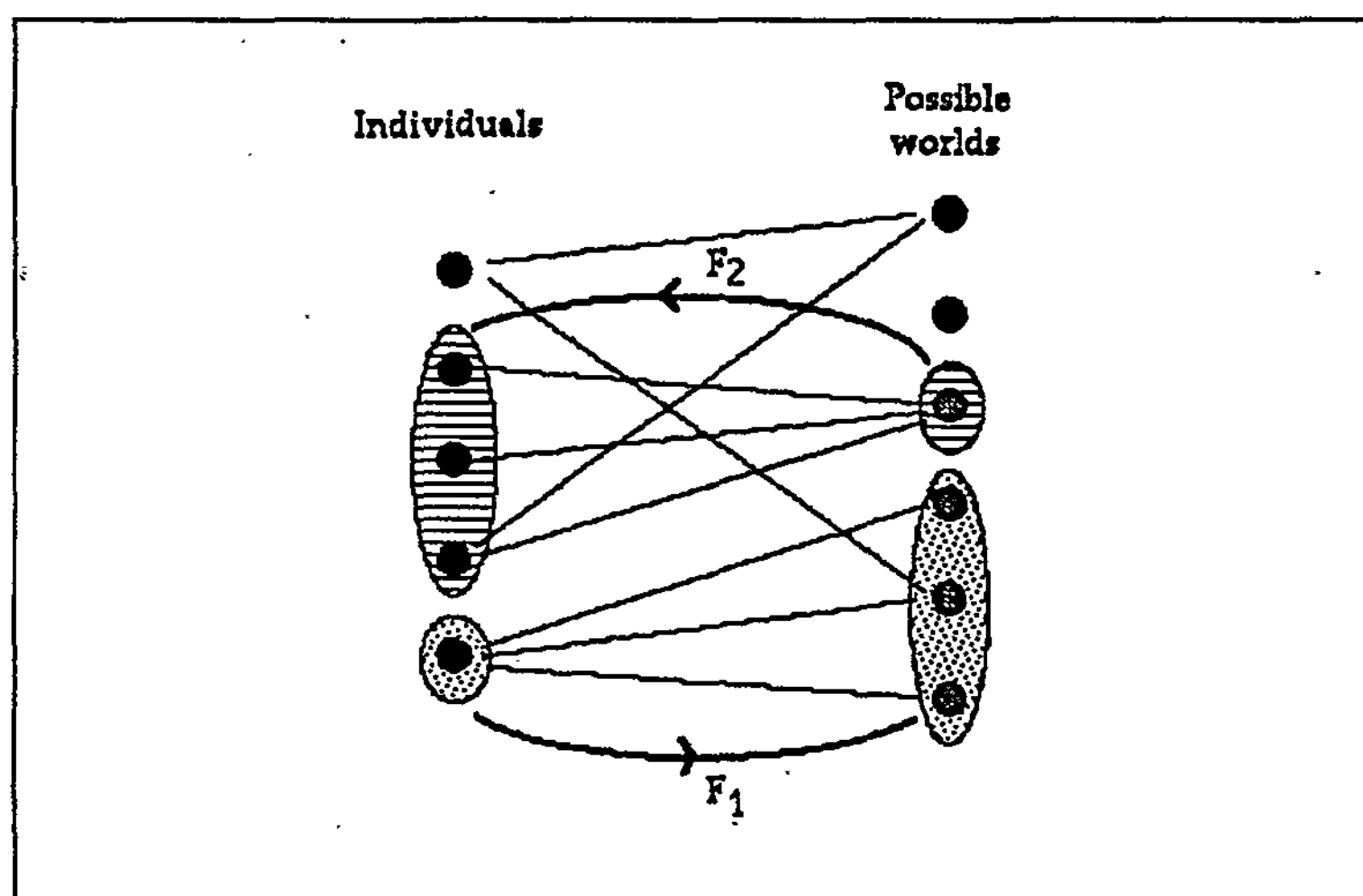


Figure 1

A property as a many-many relation between individuals and possible worlds

F_1 : 'Propositional' function mapping individuals on propositions

F_2 : 'Extensional' function mapping possible worlds on classes

The definition is certainly not helpful for cognitive psychologists who try to explain what happens when a person *perceives* that two objects have the same property in common or, for example, why certain colors look *similar*. However, the fact that the definition is counterintuitive is not a decisive argument ; it may be argued that the abstract character of properties is merely a cosmetic feature of the intensional semantics — as long as semantics produces the right results, the technical form of semantic concepts is not so important.

A related, but more serious, problem for the traditional definition of a property is that it can hardly account for *inductive reasoning*. An inductive inference generally consists in *connecting* two properties to each other, as when one concludes that all living things have chromosomes. This connection is obtained from a number of instances of individuals exhibiting the relevant properties. If a property is defined as a function from possible worlds to sets of individuals (or equivalently as a function from individuals to sets of possible worlds), then in order to determine which properties are instantiated by a particular individual (or a set of individuals), one has to determine *which functions* have the individual (or the set of individuals) as value in the actual world. Apart from problems concerning how we determine which is the actual world, this recipe will in general give us *too many* properties. For example, if we are examining a particular emerald it will instantiate a large class of Goodman-type properties like 'grue' apart from standard properties like 'green'.

¹For further criticism of the classical approach to properties in relation to inductive reasoning, cf. [Gärdenfors, 1990a].

If the only thing we know about properties is that they are some kind of abstract functions, then we have no way of distinguishing natural and inductively projectible properties like 'green' from inductively useless properties like 'grue'. What is needed is a criterion for separating the projectible sheep from the non-projectible goats. However, classical intensional semantics does not provide us with such a criterion¹.

A third problem that arises from defining properties as abstract functions concerns the difficulty of expressing an *anti-essentialistic* doctrine as has been pointed out by [Stalnaker, 1981]. Anti-essentialism is the doctrine that things have none of their properties necessarily. In that paper, Stalnaker's aim is not to defend the doctrine, but to show that on the traditional account of properties it cannot be coherently expressed.

Some kinds of essential properties are unobjectionable to an anti-essentialist : any property that is necessarily an essential property of everything, like 'being self-identical' or 'being either human or non-human' ; certain relational properties that are defined in terms of specific individuals may be essential to that individual, like 'being the same age as Ingmar Bergman' which is essential to Ingmar Bergman ; and "world-indexed" properties, like 'being an emu in the actual world' which an individual has essentially if and only if it actually is an emu. The problem for an anti-essentialist is how to find a criterion which can distinguish between such innocent essential properties and the ontologically dangerous kinds of essential properties. Stalnaker formulates the problem as follows :

"In terms of this extensional account of properties (extensional in the sense that properties are defined by their extensions in different possible worlds), what corresponds to the intuitive distinctions between referential and purely qualitative properties, and between world-indexed and world-independent properties? Nothing. All properties are referential in the sense that they are defined in terms of the specific individuals that have them. All properties are world-indexed in the sense that they are defined in terms of the specific possible worlds in which things have them. While one can, of course, make a distinction between essential and accidental attributes in terms of the standard semantical framework, one cannot find any independent distinctions corresponding to the intuitive ones needed to state a coherent version of the anti-essentialist thesis. Thus there is no satisfactory way, without adding to the primitive basis of the semantical theory, to state the thesis as a further semantical constraint on legitimate interpretations of the language of modal logic" [Stalnaker, 1981, p. 346].

Stalnaker concludes that "[w]hat the standard semantics lacks is an account of properties that defines them independently of possible worlds and individuals. (...) [A] property must be not just a rule for grouping individuals, but a feature of individuals in virtue of which they may be grouped (...)" [Stalnaker, 1981, p. 347].

The final problem that I shall point out for the functional definition of properties is perhaps the most serious one. [Putnam, 1981] has shown that the standard model-theoretic definition of "property" which has been

given here does not work as a theory of the *meaning* of properties. In proving this result, Putnam makes two assumptions about "the received view" of meaning : (1) The meaning of a sentence is a function which assigns a truth value to the sentence in each possible world ; and (2) the meaning of the parts of a sentence cannot be changed without changing the meaning of the whole sentence.

Putnam's general proof is quite technical, but the thrust of the construction can be illustrated by his example [Putnam, 1981, p. 33-35]. He begins with the sentence :

(1) A cat is on a mat.

where "cat" refers to *cats* and "mat" to *mats* as usual. He then shows how to give (1) a new interpretation.

(2) A cat* is on a mat*.

The definitions of the properties cat* and mat* make use of three cases :

- (a) Some cat is on some mat and some cherry is on some tree.
- (b) Some cat is on some mat and no cherry is on any tree.
- (c) Neither (a) nor (b) holds.

Here are Putnam's definitions :

• DEFINITION OF 'CAT*'

x is a cat* if and only if case (a) holds and x is a cherry or case (b) holds and x is a cat ; or case (c) holds and x is a cherry.

• DEFINITION OF 'MAT*'

x is a mat* if and only if case (a) holds and x is a tree or case (b) holds and x is a mat ; or case (c) holds and x is a quark.

Given these definitions it turns out that the sentence (1) is true in exactly those possible worlds where (2) is true. Thus, according to the received view of meaning, these sentences will have the same meaning. In the appendix to his book, Putnam shows that a more complicated reinterpretation of this kind can be constructed for all the sentences of a language. He concludes that "there are always infinitely many different interpretations of the predicates of a language which assign the «correct» truth-values to the sentences in all possible worlds, *no matter how these «correct» truth-values are singled out*". Thus "(...) *truth-conditions for whole sentences underdetermine reference*" [Putnam, 1981, p. 35]. Again, the underlying reason is that there are *too many* potential properties if they are defined as functions from individuals to propositions, i.e. in terms of

possible worlds and truth values. Cat* and mat* are just two examples from this large class.

Putnam's own diagnosis of the problem is that it occurs because of viewing a language as separate from its interpretation²:

²For further discussion of Putnam's theorem and its relevance to semantics, cf. [Lakoff, 1987, ch. 15].

"The predicament only is a predicament because we did two things : first, we gave an account of understanding the language in terms of programs and procedures for *using* the language (what else ?) ; then, secondly, we asked what the possible «models» for the language were, thinking of the models as existing «out there» *independent of any description*. At this point, something really weird had already happened, had we stopped to notice. On any view, the understanding of the language must determine the reference of the terms, or, rather, must determine the reference given the context of use. If the use, even in a fixed context does not determine reference, then use is not understanding. The language, on the perspective we talked ourselves into, has a full program of use ; but it still lacks an *interpretation*.

This is the fatal step. To adopt a theory of meaning according to which a language whose whole use is specified still lacks something — viz. its «interpretation» — is to accept a problem which *can* only have crazy solutions. To speak as if *this* were my problem, «I know how to use my language, but, now, how shall I single out an interpretation ?» is to speak nonsense. Either the use *already* fixes the «interpretation» or nothing can.

Nor do «causal theories of reference», etc., help. Basically, trying to get out of this predicament by *these* means is hoping that the *world* will pick one definite extension for each of our terms even if *we* cannot. But the world does not pick models or interpret languages. *We* interpret our languages or nothing does" [Putnam 1980, p. 481-482].

I have here presented four different arguments against the traditional definition of the notion of a property. The upshot is that there is something rotten in the kingdom of semantics. What is needed is a completely different way of defining properties. I shall argue that a cognitively oriented approach will do the work³.

3. Conceptual spaces as a basis for a new criterion of properties

On my view, the semantics for a language is primarily *a relation between the language and a cognitive structure*. The *meaning* of an expression is determined by what it corresponds to in such a cognitive structure. The external world enters the picture only when the relation between it and the conceptual structure is considered. This means that the *truth* of sentences is, at best, a secondary feature of a semantic theory. Questions of meaning must be answered *before* we can raise any questions about truth.

As a framework for a cognitive structure used in describing a semantics I want to put forward the notion of a *conceptual space*. A conceptual space consists of a number of *quality dimensions*. As examples of quality dimensions let me mention color, pitch, temperature, weight, and the three ordinary spatial dimensions⁴. The dimensions are taken to be

³Prime examples of works in the tradition of cognitive semantics are [Lakoff, 1987] and [Langacker, 1986]. Related versions can be found in the writings of [Jackendoff, 1983, 1990], [Johnson-Laird, 1983], [Fauconnier, 1985], [Talmy, 1988], [Sweetser, 1990] and many others. There is also a French linguistic and semiotic tradition, exemplified by [Desclés, 1985] and [Petitot, 1985, 1992], which shares many features with the American (mainly Californian) group.

cognitive and infra-linguistic in the sense that we (and other animals) can represent the qualities of objects, for example when planning an action, without presuming an internal language in which these qualities are expressed. Some of the dimensions are closely related to what is produced by our sensory receptors, but there are also quality dimensions that are of an abstract non-sensory character⁵.

The notion of a *dimension* should be understood literally. It is assumed that each of the quality dimensions is endowed with certain topological or metric structures. For example, 'time' is a one-dimensional structure which we conceive of as being isomorphic to the line of real numbers⁶. Similarly, 'weight' is one-dimensional with a zero point, isomorphic to the half-line of non-negative numbers. Some quality dimensions have a *discrete* structure, i.e. they merely divide objects into classes, e.g. the sex of an individual⁷.

At this point it is important to make a distinction between a *psychological* and a *scientific* interpretation of the quality dimensions. For example, our psychological visual space is not a perfect 3-dimensional Euclidean space, since it is not invariant under all linear (Galilean) transformations. Because of gravity, among other things, the vertical dimension is treated differently from the two horizontal dimensions. However, the scientific representation of visual space as a 3-D Euclidean space is an idealization that is mathematically amenable (there is no preferred direction, Galilean transformations preserve the structure, etc.). Similarly, our perception of the weight of objects is not fine enough to justify its representation by the full structure of the positive real numbers, but this scientific representation is motivated by the fact that the mathematics of this structure is well-known, and thus makes it possible to formulate a quantitative theory of weight which is easy to handle computationally. However, when it comes to providing a semantics for a natural language, it is, of course, the psychological interpretations of the quality dimensions that are in focus.

A psychologically interesting example of a quality dimension concerns *color perception*. In brief, our cognitive representation of colors can be described by three dimensions. The first dimension is *hue*, which is represented by the familiar *color circle*. The topological structure of this dimension is thus different from the quality dimensions representing time or weight which are isomorphic to the real line. One way of illustrating the differences in topology is by noting that we can talk about psychologically *complementary* colors, i.e. colors that lie *opposite* to each other on the color circle. In contrast it is *not meaningful* to talk about two points of time or two weights being "opposite" to each other. This simple example shows that the topological structure of the cognitive representations of perceptual qualities will have important consequences for the *semantics* of linguistic expressions used to talk about these qualities.

⁴The notion of a quality dimension is closely related to the 'domains' in Langacker's semantic theory [Langacker, 1986].

⁵Some further examples will be given in the following section.

⁶To some extent the representation of time is culturally dependent, so that other cultures have a different time dimension as a part of their cognitive structure. Cf. [Gärdenfors, 1990b] for a discussion of how this influences the structure of language.

⁷Discrete dimensions may also have additional structure as, for example, in kinship or biological classifications. The topology of discrete dimensions is further discussed in [Gärdenfors, 1990b].

⁸For further examples of perceptual quality dimensions, cf [Gärdenfors, 1990b].

The second psychological dimension of color is *saturation*, which ranges from gray (zero color intensity) to increasingly greater intensities. This dimension is isomorphic to an interval of the real line. The third dimension is *brightness* which varies from white to black and is thus a linear dimension with end points. Together these three dimensions, one with circular structure and two with linear, make up the color space which is a subspace of our perceptual conceptual space (see Figure 2)⁸.

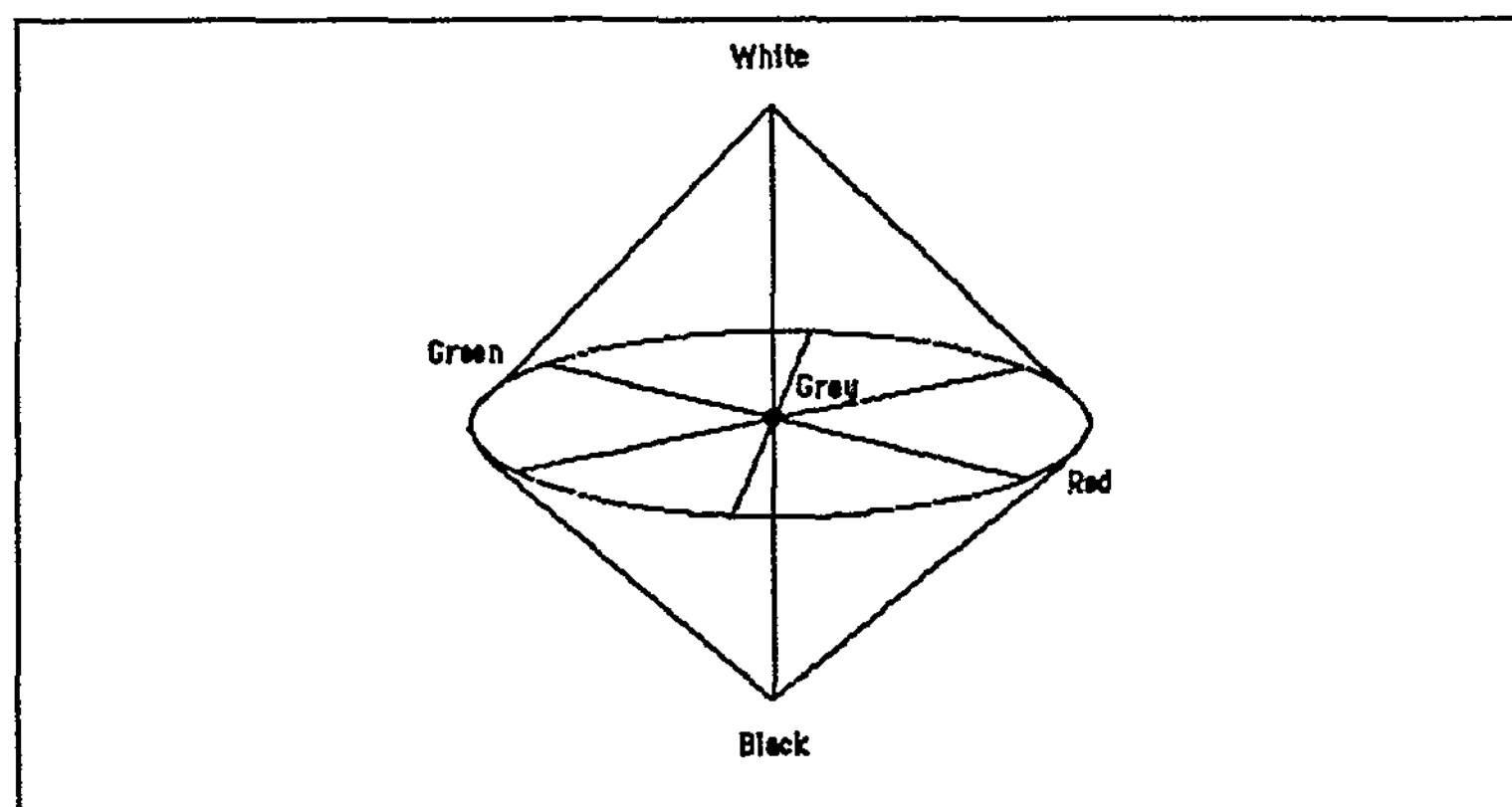


Figure 2
The full color space

I cannot provide a complete list of the quality dimensions involved in our conceptual spaces. Some of the dimensions seem to be *innate* and to some extent hardwired in our nervous system, as for example color, pitch, and probably also ordinary space. These subspaces are obviously extremely important for basic activities like finding food and getting around in the environment. Other dimensions are presumably *learned*. Learning new concepts often involves expanding one's conceptual space with new quality dimensions. *Functional* properties used for describing artifacts may be an example here. Such properties are characterized by a set of associated *actions*. Even if we do not know very much about the conceptual dimensions underlying actions, it is quite obvious that there is such a non-trivial structure⁹. Still other dimensions may be *culturally* dependent. 'Time' is a good example — in contrast to our linear conception of time, some cultures conceive of time as circular so that the world keeps returning to the same point in time, and in other cultures it is hardly meaningful at all to speak of time as a dimension¹⁰. Finally, some quality dimensions are introduced by *science*. As an example, let me mention the distinction between temperature and heat, which is central for thermodynamics, but which has no correspondence in human perception.

⁹See e. g. Vaina's analysis of functional representation [Vaina, 1983].

¹⁰Another example of the cultural relativity of time dimension is discussed in [Gärdenfors, 1990b, section 6].

(Human perception of heat is basically determined by the amount of heat transferred from an object to the skin rather than by the temperature of the object.)

This concludes my general presentation of conceptual spaces. It can be seen as a generalization of the *state space approach*, advocated among others by [Churchland (P. M.), 1986] and [Churchland (P. S.), 1986], and of the *vector function theories* of [Foss, 1988]. To some extent conceptual spaces also function like the *domains* in Langacker's semantic theory [Langacker, 1987]. The theory of conceptual spaces is a *theory for representing information*, not a psychological or neurological theory, which I believe can be applied to a number of philosophical problems in epistemology and semantics. Here, my primary aim is to show its viability as a foundation for intensional semantics.

4. Properties described with the aid of conceptual spaces

In more abstract terms, a conceptual space S consists of a class $D_1, \dots, D_n$ of quality dimensions. A point in S is represented by a vector $v = \langle d_1, \dots, d_n \rangle$ with one index for each dimension. Each of the dimensions is endowed with a certain topological or metrical structure.

A first rough idea is to describe a property as a *region* of a conceptual space S , where "region" should be understood as a spatial notion determined by the topology and metric of S . For example, the point in the time dimension representing 'now' divides this dimension, and thus the space of vectors, into two regions corresponding to the properties 'past' and 'future'. In contrast to the traditional definition in intensional semantics that was presented in Section 2, this definition presumes *neither* the concept of an individual *nor* the concept of a possible world. But the proposal suffers from a lack of precision as regards the notion of a "region". A more precise and powerful idea is the following criterion where the topological characteristics of the quality dimensions are utilized to introduce a spatial structure on properties :

Criterion P : A *natural property* is a convex region of a conceptual space.

A *convex region* is characterized by the criterion that for a very pair of points v_1 and v_2 in the region all points in between v_1 and v_2 are also in the region. The motivation for the criterion is that if some objects which are located at v_1 and v_2 in relation to some quality dimension (or several dimensions) both are examples of the property P , then any object that is located between v_1 and v_2 on the quality dimension(s) will also be an example of P . I shall argue later that this criterion is psychologically realistic. Criterion P presumes that the notion of *betweenness* is meaningful for the relevant quality dimensions. This is, however, a rather

¹¹See [Birkhoff, 1967]
for an axiomatic
analysis of "between".

¹²For an extended
analysis of this
example, see
[Gärdenfors, 1990a].

¹³A self-approved
vowel is one that was
produced by the
speaker and later
approved of as an
example of the intended
kind. An identified
sample of a vowel is
one that was correctly
identified by 75% of the
observers. The
preferred samples of a
vowel are those which
are "the most
representative samples
from among the most
readily identified
samples" [Fairbanks &
Grubb, 1961, p. 210].

weak assumption which demands very little of the underlying topological structure¹¹.

Most properties expressed by simple words in natural languages are natural properties in the sense specified here. For instance, I conjecture that all *color terms* in natural languages express natural properties with respect to the psychological representation of the three color dimensions. In other words, the conjecture predicts that if some object o_1 is described by the color term C in a given language and another object o_2 is also said to have color C , then any object o_3 with a color that lies between the color of o_1 and that of o_2 will also be described by the color term C . It is well-known that different languages carve up the color circle in different ways, but all carvings seems to be done in terms of convex sets. Strong support for this conjecture can be found in [Berlin & Kay, 1969], although they do not treat color terms in general but concentrate on basic color terms. On the other hand, the reference of an artificial color term like 'grue' will not be a convex region in the ordinary conceptual space and thus it is not a natural property according to Criterion P ¹².

Another illustration of how the convexity of regions determines properties and categorizations is the phonetic identification of *vowels* in various languages. According to phonetic theory, what determines a vowel are the relations between the basic frequency of the sound and its formants (higher frequencies that are present at the same time). In general, the first two formants F_1 and F_2 are sufficient to identify a vowel. This means that the coordinates of two-dimensional space spanned by F_1 and F_2 (in relation to a fixed basic pitch F_0) can be used as a fairly accurate description of a vowel. [Fairbanks & Grubb, 1961] investigated how people produce and recognize vowels in "General American" speech. Figure 3 (see on the opposite page) summarizes some of their findings. The scales of the abscissa and ordinate are the logarithm of the frequencies of F_1 and F_2 (the basic frequency of the vowels was 130 cps). As can be seen from the diagram, the preferred, identified and self-approved examples of different vowels form convex subregions of the space determined by F_1 and F_2 with the given scales¹³. As in the case of color terms, different languages carve up the phonetic space in different ways (the number of vowels identified in different languages varies considerably), but I conjecture again that each vowel in a language will correspond to a convex region of the formant space.

Criterion P provides an account of properties that is independent of both possible worlds and individuals and it satisfies Stalnaker's *desideratum* that a property "(...) must be not just a rule for grouping individuals, but a feature of individuals in virtue of which they may be grouped" [Stalnaker, 1981, p. 347]. However it should be emphasized that I only view the criterion as a *necessary* but perhaps not sufficient

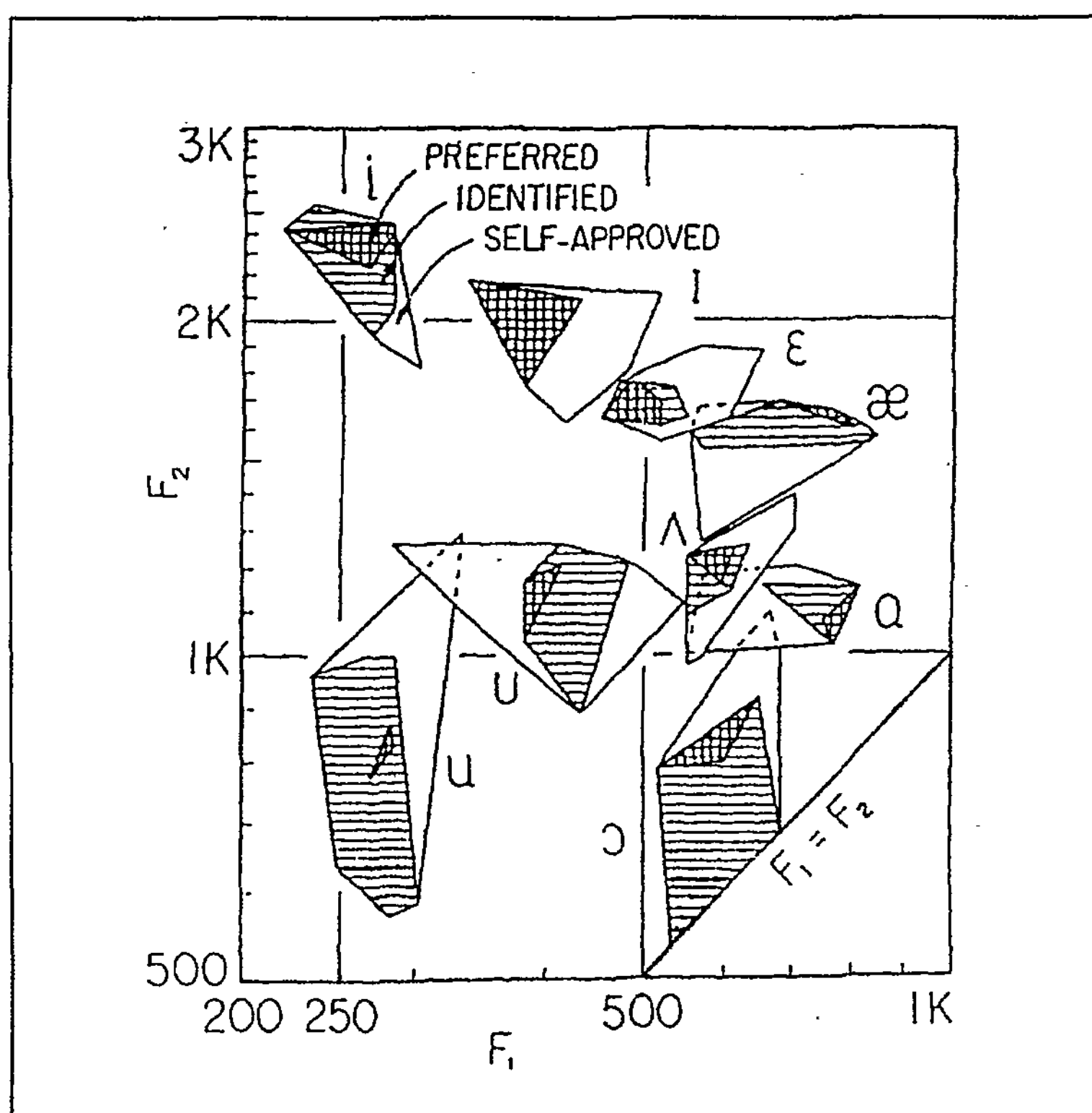


Figure 3
Frequency areas of different vowels in the two-dimensional space
generated by the first two formants (*values in cps*)
(From [Fairbanks & Grubb, 1961])

condition on a natural property. The criterion delimits the class of properties that are useful for cognitive purposes, but it may not be sufficiently restrictive.

5. Basic elements of an intensional semantics

Introducing the notion of a natural property is only the first step in developing an intensional semantics based on conceptual spaces. According to my view, semantics is a relation between language and a conceptual space. An *interpretation* for a language L consists of a mapping of the components of L onto a conceptual space. As a first element of such a mapping, *individual constants* are assigned vectors (i.e. points in the conceptual space) or partial vectors (i.e. points with some

¹⁴Abstract entities may be assigned values on a different set of quality dimensions.

¹⁵See [Gärdenfors, 1988, section 4] for an analysis of the semantics of 'secondary' properties.

¹⁶Again see [Gärdenfors, 1988, Section 4] and [Holmqvist, 1988] for some of the details.

arguments undetermined). In this way each name (referring to an individual) is allocated a specific color, spatial position, weight, temperature, etc¹⁴. If a constant is assigned to a partial vector, this means that not all the properties are known or have been determined. Following [Stalnaker, 1981, p. 347], a function which maps the individuals into a conceptual space will be called a *location function*.

As a second element of the interpretation mapping, the *predicates* of the language that denote primary properties are assigned convex regions in the conceptual space in accordance with Criterion *P*. Such a predicate is *satisfied* by an individual just in case the location function locates the individual at one of the points included in the region assigned to a predicate. Some of the so called intensional predicates, like 'tall', 'former' or 'alleged', do not denote primary properties in the sense that their regions can be described independently of other properties. Such secondary predicates, which are "parasitical" on other properties, can be described in terms of the regions assigned to the primary properties¹⁵. *Relations* (primary and secondary) can be treated in a similar way¹⁶.

If we assume that an individual is completely determined by its set of properties, then all points in the conceptual space can be taken to represent *possible individuals*. On this account, a possible individual is a *cognitive* notion that need not have any form of reference in the external world. This construction will avoid many of the problems that have plagued other philosophical accounts of possible individuals. A point in a conceptual space will always have an internally consistent set of properties — since e.g. 'blue' and 'yellow' are disjoint properties in the color space, it is not possible that any individual will be both blue and yellow (all over). There is *no need for meaning postulates* or their ilk in order to exclude such contradictory properties.

What has been accomplished here are only the first steps in the construction of an intensional semantics based on conceptual spaces. One important contrast to the traditional intensional semantics is that the new one does not presume the concept of a *possible world*. However, different location functions describe alternative ways that individuals may be located in a conceptual space. Thus these location functions have the same role as possible world in the traditional semantics. This means that we can *define* the notion of a possible world as a possible location function and this can be done without introducing any new semantical primitives to the theory (cf. [Stalnaker, 1981, p. 348]). In this way most of the constructions from traditional intensional semantics will be available, should we want them.

If we assume that the meanings of the predicates, among other things in a language *L*, are determined by a mapping into a conceptual space *S*, it follows from Criterion *P* and the topological structure of different quality dimensions that certain statements will become *analytically* true (independent of empirical considerations). For example the fact that

comparative relations like 'earlier than' are *transitive* follows from the linear structure of the length dimension and is thus an analytic feature of this relation (analytic-in- S , that is). Similarly, it is analytic that everything that is green is colored (since 'green' refers to a region of the color space) and that nothing is both red and green. Analytic-in- S is thus defined on the basis of the topological and metric structure of the conceptual space S . A consequence of this definition is that an analytic statement will be satisfied for all location functions. However, different conceptual spaces will yield different notions of analyticity.

6. How the problems for the traditional definition of "property" are avoided

I now want to argue that the criterion of "natural property" given above eludes all the problems of Section 2 that tainted the traditional definition. First of all, Criterion P makes many properties *perceptually grounded*. Since the basic quality dimensions are more or less determined by our perceptual mechanisms, there is a direct link between properties described as regions of such dimensions and perceptions. In other words, many of the elements of the location functions will be determined by perceptions. This means that the present criterion of properties will be much more useful for cognitive psychologists than the traditional definition. In the following section it will be shown that there are close connections between the idea of describing properties in terms of conceptual spaces and the *prototype theory* of categorization.

Secondly, describing properties in terms of conceptual spaces makes it much easier to understand how inductive inferences are made. The fundamental criterion is that *in inductive inferences we only allow predicates denoting natural properties* (as described here). This criterion gives us a tool for separating projectible predicates from non-projectible ones. In this way we cut down, in a non-arbitrary way, the immense class of properties (among them the Goodman-type properties) that will be available if properties are defined as functions from possible worlds to individuals.

In [Gärdenfors, 1990a] it is argued that the traditional problems in the analysis of induction have arisen because philosophers (and AI researchers) have confined themselves to linguistic representations of knowledge¹⁷. However, if we want to understand human inductive reasoning, we must go deeper than language, i.e. down to conceptual spaces. The same applies if we strive to construct computer programs that, even in a limited way, mirror the human inductive capacity. Imposing a *spatial structure* on properties, as is done here, will open up for new kinds of programming techniques. Traditional approaches to mechanized induction basically use formula manipulations. On the other hand, an

¹⁷The paper contains among other things an analysis of Hempel's and Goodman's paradoxes of induction.

implementation of the theory of conceptual spaces would use more directly "mathematical" operations, like vector calculations.

Third, the problems for the anti-essentialistic doctrine raised by Stalnaker now disappear. Much of the inspiration for the theory of conceptual spaces comes from what [Stalnaker, 1981] calls *logical space*. The main difference between his notion and mine, is that he does not emphasize the role of the topological and metric structure of the quality dimensions and he can thus not talk about "convexity" and similar notions. Here is Stalnaker's solution to the problem of expressing the anti-essentialistic doctrine :

"It should be clear that every *property* (every region of logical space) determines a propositional function in the following way : given any property, the value of the corresponding propositional function, relative to a given possible world, will be the class of individuals that have the property in that possible world — the individuals that are located in that region of logical space by the location function that represents the possible world. Thus every property determines a unique propositional function and the correspondence is one-one : distinct properties never determine the same propositional function. But it is not the case that every propositional function corresponds to an intrinsic property, for the classes of individuals selected by a propositional function in the different possible worlds need not all come from the same region of logical space. Among the propositional functions, or properties in the broad sense, that do not correspond to regions of logical space are, of course, just those that the anti-essentialist wants to distinguish from full-fledged intrinsic properties. For example, referential properties such as *being the same weight as Babe Ruth* will clearly not correspond to regions of logical space" [Stalnaker, 1981, p. 348-349].

Thus the possible worlds semantics *generated* from conceptual space semantics is rich enough to represent the distinctions needed to make sense of the appropriate kind of anti-essentialism.

Fourth and finally, the problems that Putnam's theorem causes the traditional definition of "property" dissolve into thin air under the new criterion. 'Cat' denotes a region of a conceptual space (relating to the class of possible animals, be they determined in terms of shape, biological functions, or whatever. In the next section a classification of animals in terms of shapes will be outlined). This region would, at least partly, be determined by the perceptual features of cats. We cannot create a new natural property 'cat*' by relating it to what facts are true in various possible worlds : 'cat*' as introduced by Putnam is indeed a propositional function, but definitely not a natural property and thus not an eligible candidate for an interpretation function that maps a language into a conceptual space.

7. Two further applications : Prototype theory and non-monotonic reasoning

Having argued that describing properties in terms of conceptual spaces solves all the traditional problems, I want to show in this section that such

a criterion can throw new light on other areas of research where the notion of a property is central.

As a first application I want to show that describing properties as convex regions of conceptual spaces fits very well with the so called *prototype theory* of categorization developed by Rosch and her collaborators ([Rosch, 1975, 1978] , [Mervis & Rosch, 1981], [Lakoff, 1987]). The main idea of prototype theory is that within a category of objects, like those instantiating a property, certain members are judged to be more representative of the category than others. For example robins are judged to be more representative of the category 'bird' than are ravens, penguins and emus ; and desk chairs are more typical instances of the category 'chair' than rocking chairs, deck-chairs, and beanbag chairs. The most representative members of a category are called *prototypical* members. It is well-known that some properties, like 'red' and 'bald' have no sharp boundaries and for these it is perhaps not surprising that one finds prototypical effects. However, these effects have been found for most properties including those with comparatively clear boundaries like 'bird' and 'chair'.

Now if either the extensional or the intensional version of the traditional definition of a property is adopted it is very difficult to explain such prototype effects. Either an object is a member of the class assigned to a property (relative to a given possible world) or it is not and all members of the class have equal status as category members. Rosch's research has been aimed at showing asymmetries among category members and asymmetric structures within categories. Since the traditional definition of a property neither predicts nor explains such asymmetries, something else must be going on.

In contrast, if properties are described as convex regions of a conceptual space, prototype effects are indeed to be expected. In a convex region one can describe positions as being more or less *central*. For example, if color properties are identified with convex subsets of the color space, the central points of these regions would be the most prototypical examples of the color. In a series of experiments, Rosch has been able to demonstrate the psychological reality of such "focal" colors. For another illustration we can return to the categorization of vowels presented in Section 4. Here the structure of the subjects' different kinds of responses show clear prototype effects.

For more complex categories like 'bird' it is perhaps more difficult to describe the underlying conceptual space. However, if something like Marr and Nishihara's analysis of shapes is adopted [Marr & Nishihara, 1978], we can begin to see how such a space would appear¹⁸. Their scheme for describing biological forms uses hierarchies of cylinder-like modelling primitives. Each cylinder is described by two coordinates (length and width). Cylinders are combined by determining the angle between the dominating cylinder and the added one (two polar

¹⁸This analysis is expanded in [Marr, 1982, ch. 5]. A related model, together with some psychological grounding, is presented by [Biederman, 1987].

coordinates) and the position of the added cylinder in relation to the dominating one (two coordinates). The details of the representation are not important in the present context, but it is worth noting that on each level of the hierarchy an object is described by a comparatively small number of coordinates based on lengths and angles. Thus the object can be identified as a hierarchically structured vector in a (higher order) conceptual space. Figure 4 provides an illustration of the hierarchical structure of their representations.

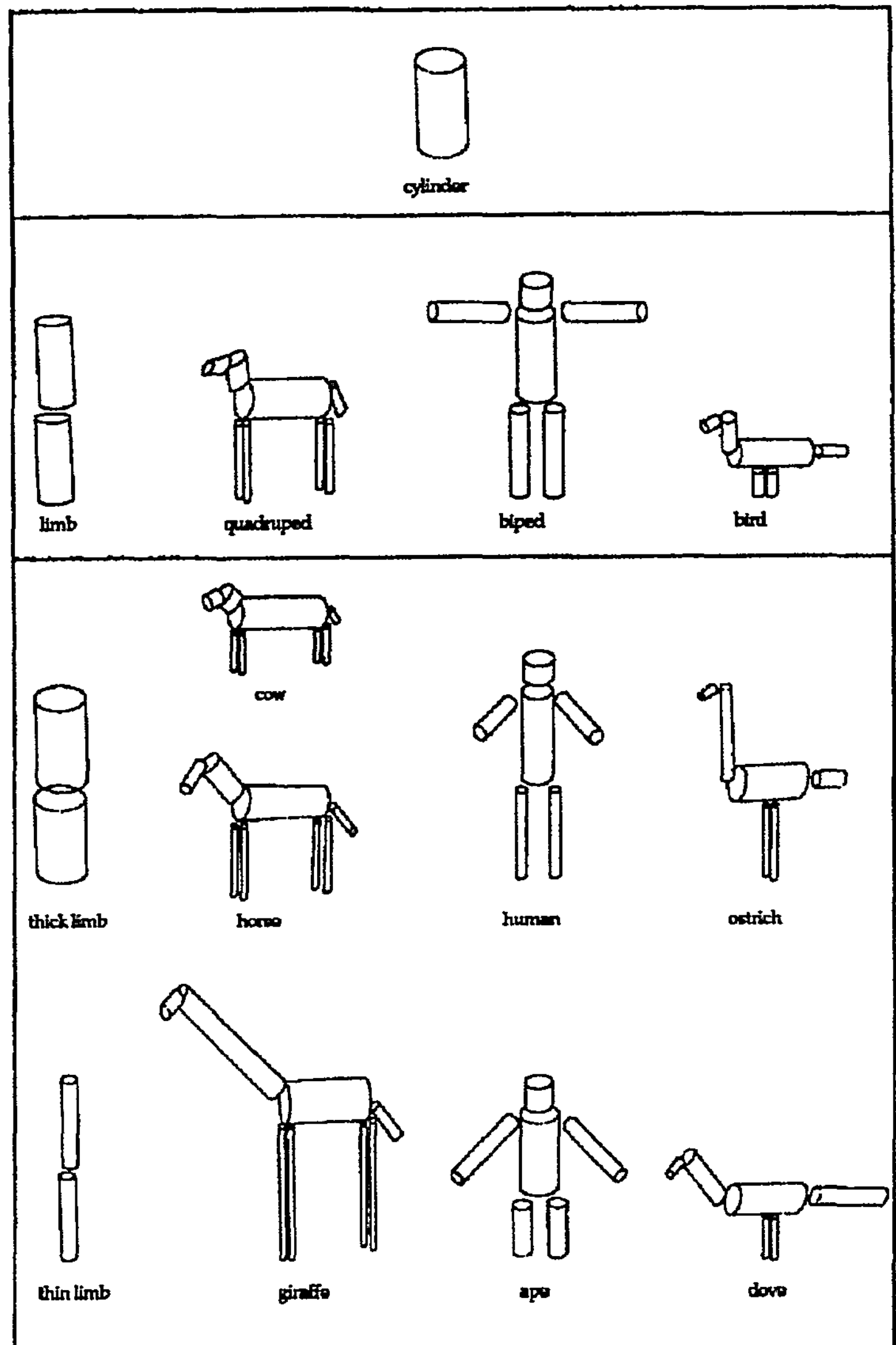


Figure 4
Hierarchical representation of animal shapes
using cylinders as modelling primitives
(from [Marr, 1982])

It should be noted that even if different members of a category are judged to be more or less prototypical, it does not follow that some of the existing objects must represent "the prototype"¹⁹. If a category is viewed as a convex region of a conceptual space this is easily explained, since the central member of the region (if unique) is a possible individual in the sense discussed above (if all its properties are specified) but need not be among the existing members of the category. Such a prototype point in the region need not be completely described as an individual, but is normally represented as a partial vector, where only the values of the dimensions that are relevant to the category have been determined. For example, the general shape of the prototypical bird would be included in the vector, but its color or age would presumably not.

¹⁹Rosch has been misunderstood on this point. Cf. [Lakoff, 1987, ch. 2].

It is possible to argue in the converse direction too and show that if prototype theory is adopted, then the representation of properties as convex regions is to be expected. Assume that some quality dimensions of a conceptual space are given, for example the dimensions of color space, and that we want to partition it into a number of categories, for example color categories. If we start from a set of prototypes $p_1, \dots, p_n$ of the categories, for example the focal colors, then these should be the central points in the categories they represent. One way of using this information is to assume that for every point p in the space one can measure the distance from p to each of the p_i 's. If we now stipulate that p belongs to the same category as the closest prototype p_i , it can be shown that this rule will generate a partitioning of the space that consists of convex areas (convexity is here defined in terms of the assumed distance measure). This is the so called *Voronoi tessellation*, a two-dimensional example of which is illustrated in Figure 5.

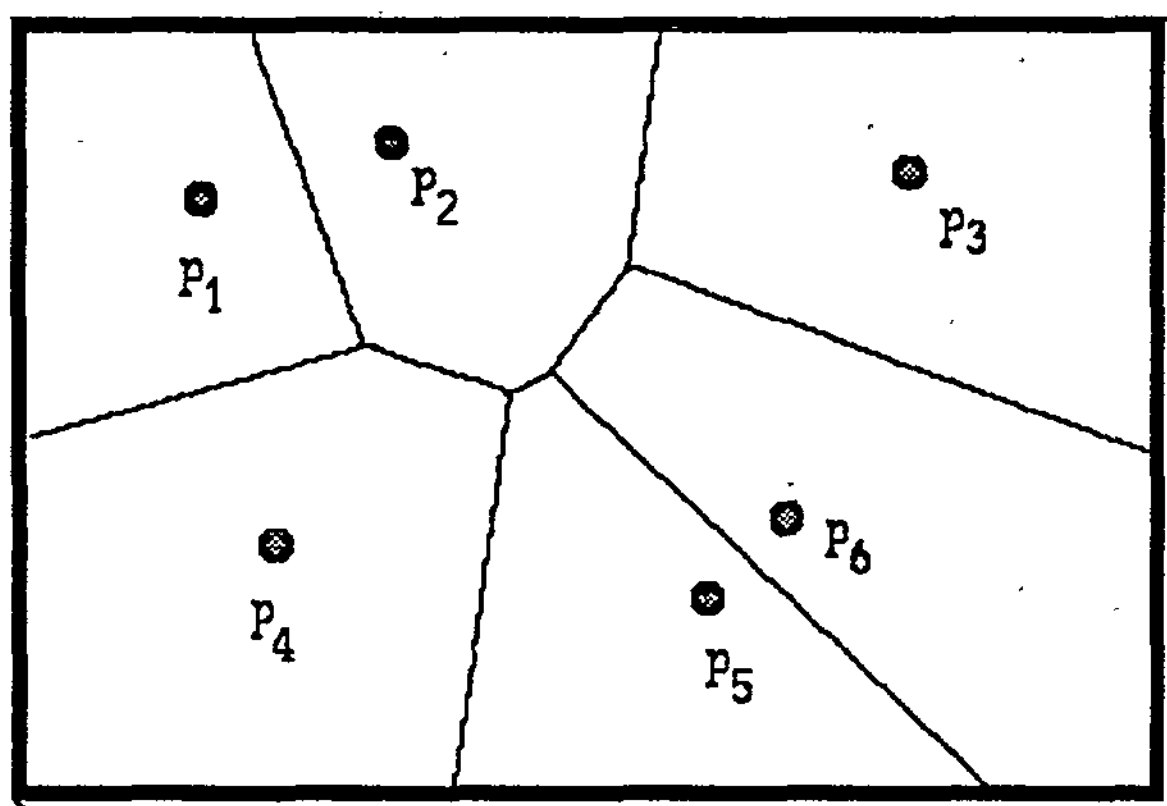


Figure 5

Voronoi tessellation of the plane into convex sets

Thus, assuming that a metric is defined on the subspace that is subject to categorization, a set of prototypes will by this method generate a unique partitioning of the subspace into convex regions. Hence there is an intimate link between prototype theory and the analysis of this article where properties are described as convex regions in a conceptual space.

A second application of the proposed criterion of properties concerns *non-monotonic reasoning*. The deductive reasoning of traditional logic is *monotonic* in the sense that when a proposition *A* can be inferred from a set *S* of sentences, then *A* can be inferred from any set that contains *S*. However, everyday reasoning, which is in general based on assumptions about what is "normally" the case, is often non-monotonic. For example, if I learn that Gonzo is a bird, then with the aid of the presumption that birds normally fly, I may conclude that Gonzo flies. But if my knowledge is expanded by the new information that Gonzo is an emu, this conclusion is no longer drawn.

Non-monotonic reasoning is one of the hottest topics within AI²⁰. However, most of the research efforts have been concentrated on finding the appropriate *logical rules* that govern non-monotonic reasoning. This means that a *propositional representation* of the relevant knowledge is already presumed. I believe (for essentially the same reasons as for the case of inductive reasoning) that in order to understand non-monotonic reasoning one must go beyond linguistic formulations of information.

Here the representation of properties as convex regions in a conceptual space may be useful. If the first thing I ever hear about the individual Gonzo is that it is a bird, I will naturally locate it in the conceptual space as a more or less prototypical bird, i.e. at the center of the region representing birds²¹. And in that area of the conceptual space, birds do fly, i.e. almost all individuals located there also have the ability to fly. However, if I then learn that Gonzo is an emu, I must *revise* my earlier location function and put Gonzo in the emu region, which is a subset of the bird region but presumably lies at the outskirts of that region [see note 6]. And in the emu region of the conceptual space almost all individuals do not fly.

This simple example only hints at how the *correlation* between different parts of a region representing a property and regions representing other properties can be used in understanding non-monotonic reasoning. For this analysis the spatial structure of properties is essential. Such correlations will only be formulated in an *ad hoc* manner if a propositional representation of information is used, where the spatial structure cannot be utilized. But I am convinced that developing the program sketched here will show how the idea of representing properties and other categories in terms of conceptual spaces can explain various phenomena of non-monotonic reasoning²².

²⁰See [Reinfrank et al., 1988] for a survey of current research areas.

²¹The relevant conceptual space may be something like a 17-dimensional hierarchical space of coordinates in Marr and Nishihara's style [Marr & Nishihara, 1978].

²²See [Holmqvist, 1989] for some further examples and ideas in this direction.

8. Conclusion

This article contains one critical and one more constructive part. In the critical part, I examined the definition of "property" that is standard within traditional intensional semantics and argued that this definition leads to serious philosophical problems, let alone problems in other areas of cognitive science.

As a remedy, I proposed, in the constructive part of the paper, that conceptual spaces be used as a framework for representing information. I outlined the first steps of an alternative intensional semantics based on conceptual spaces and, in particular, I suggested that the notion of "natural property" be described as a convex region in a conceptual space. It was then shown that this criterion avoids the problems that tainted the traditional definition. Furthermore, I argued that the new criterion is useful for understanding some other areas involving categorization, namely prototype theory and non-monotonic reasoning.

*Department of Philosophy
Lund University*

Acknowledgements

Research for this paper has been supported by the Swedish Council for Research in the Humanities and Social Sciences, by Erik Philip-Sörensens Foundation, and by Magnus Bergvall's Foundation. I wish to thank Johan van Benthem, Jaakko Hintikka, Ernest Sosa, Lucia Vaina, and the participants at the *Cognitive Science seminar* in Lund for helpful comments on the manuscript.

References

BEALER (G.)

1989, "On the Identification of Properties and Propositional Functions", *Linguistics and Philosophy*, 12, p. 1-14.

BERLIN (B.) & KAY (P.)

1969, *Basic Color Terms : Their Universality and Evolution*, Berkeley (CA), University of California Press.

BIEDERMAN (I.)

1987, "Recognition-by-components : A Theory of Human Image Understanding", *Psychological Review*, 94, p. 115-147.

BIRKHOFF (G.)

1967, *Lattice Theory*, Providence, American Mathematical Society.

CHURCHLAND (P. S.)

1986, *Neurophilosophy : Toward a Unified Science of the Mind/Brain*, Cambridge (MA), MIT Press.

CHURCHLAND (P. M.)

1986, "Some Reductive Strategies in Cognitive Neurobiology", *Mind*, 95, n°379, p. 279-309.

DESCLÉS (J. -P.)

1985, "Représentation des connaissances", *Actes Sémiotiques-Documents*, VII, 69-70, Paris, Institut National de la Langue Française-CNRS.

FAIRBANKS (G.) & GRUBB (P.)

1961, "A Psychophysical Investigation of Vowel Formants", *Journal of Speech and Hearing Research*, 4, p. 203-219.

FAUCONNIER (G.)

1985, *Mental Spaces*, Cambridge (MA), MIT Press.

FOSS (J.)

1988, "The Percept and Vector Function Theories of the Brain", *Philosophy of Science*, 55, p. 511-537.

GÄRDENFORS (P.)

1988, "Semantics, Conceptual Spaces and the Dimensions of Music", p. 9-27, in *Essays on the Philosophy of Music*, V. Rantala, L. Rowell, and E. Tarasti eds., (*Acta Philosophica Fennica*, vol. 43), Helsinki.

1990a, "Induction, Conceptual Spaces and AI", *Philosophy of Science*, n°57, p. 78-95.

1990b, "Mental Representation, Conceptual Spaces and Metaphors", forthcoming in *Synthese*.

HINTIKKA (J.)

1961, "Modality and Quantification", *Theoria*, 27, p. 110-128.

HOLMQVIST (K.)

1988, *Aspects of Parameterizing Concept Representations*, manuscript, Department of Philosophy, Lund University.

1989, «Non-monotonic» Reasoning from an Image Schema Semantics Point of View, manuscript, Department of Cognitive Science, Lund University.

JACKENDOFF (R.)

1983, *Semantics and Cognition*, Cambridge (MA), MIT Press.

1990, *Semantic Structures*, Cambridge (MA), MIT Press.

JOHNSON-LAIRD (P.)

1983, *Mental Models*, Cambridge University Press.

KANGER (S.)

1957, *Provability in Logic*, Stockholm, Almqvist & Wiksell.

KRIPKE (S.)

1959, "A Completeness Theorem in Modal Logic", *Journal of Symbolic Logic*, 24, p. 1-24.

LAKOFF (G.)

1987, *Women, Fire, and Dangerous Things*, University of Chicago Press.

LANGACKER (R. W.)

1987, *Foundations of Cognitive Grammar*, vol. 1, Stanford University Press.

MARR (D.)

1982, *Vision*, San Francisco, Freeman.

MARR (D.) & NISHIHARA (H. K.)

1978, "Representation and Recognition of the Spatial Organization of Three-dimensional Shapes", *Proceedings of the Royal Society in London, B* 200, p. 269-294.

MERVIS (C.) & ROSCH (E.)

1981, "Categorization of Natural Objects", *Annual Review of Psychology*, 32, p. 89-115.

MONTAGUE (R.)

1974, *Formal Philosophy*, ed. by R. H. Thomason, New Haven, Yale University Press.

PETITOT (J.)

1985, *Morphogenèse du Sens*, Paris, PUF.

1992, *Physique du Sens*, Paris, Éditions du CNRS.

PUTNAM (H.)

1980, "Models and Reality", *Journal of Symbolic Logic*, 45, p. 464-482.

1981, *Reason, Truth, and History*, Cambridge University Press.

[Reinfrank et al.]

REINFRANK (M.) & de KLEER (J.) & GINSBERG (M.) & SANDEWALL (E.), eds.

1988, *Non-Monotonic Reasoning*, Berlin, Springer-Verlag.

ROSCH (E.)

1975, "Cognitive Representations of Semantic Categories", *Journal of Experimental Psychology : General*, 104, p. 192-233.

1978, "Prototype Classification and Logical Classification : The Two Systems", p. 73-86, in *New Trends in Cognitive Representation : Challenges to Piaget's Theory*, E. Scholnik ed., Hillsdale (NJ), Lawrence Erlbaum Associates.

STALNAKER (R.)

1981, "Antiessentialism", *Midwest Studies of Philosophy*, 4, p. 343-355.

TALMY (L.)

1988, "Force dynamics in language and cognition", *Cognitive Science*, 12, p. 49-100.

SWEETSER (E.)

1990, *From Etymology to Pragmatics*, Cambridge University Press.

VAINA (L.)

1983, "From Shapes and Movements to Objects and Actions", *Synthese*, 54, p. 3-36.

Pour une phrastique intégrale¹

Per Aage Brandt

Nous savons d'une part que les domaines de la pragmatique, de la logique propositionnelle, de la syntaxe phrastique et de la sémantique verbale ne se réduisent pas à une seule analytique structurellement homogène ; et d'autre part, que les modèles de stratification par modules ou composantes ne rendent pas compte de la manière dont le sens passe d'un niveau à l'autre. Comment analyser ce qui à la fois semble homogène et hétérogène, homogène dans l'expérience et hétérogène dans la théorie ?

Notre propos est ici de développer une solution topologique de ce problème. Une *phrase* est en effet un événement textuel de base ou un coup dans le jeu dialogique : c'est un référent *pragmatique*, inscrit dans une structure d'énonciation et dans un scénario communicationnel. Elle fait partie d'un schématisme proprement pragmatique. En second lieu, elle représente une unité sémantique à structure *propositionnelle*, assignant avec une certaine force modale un prédicat à un thème-sujet quantifié, dont la place et la pertinence sont établies au premier niveau par ce schématisme pragmatique, impliquant une sorte d'éthique du don² inhérente à l'intersubjectivité communicative. De par son articulation prédicative, la sémantique propositionnelle constitue pourtant un schématisme spécifique, assertif. En troisième lieu, la phrase syntaxique — toujours la même phrase, mais vue de plus près — déploie une *structure casuelle* qui semble spécifier le dynamisme modal à l'œuvre dans la relation propositionnelle entre prédicat et thème³ dans un monde scénariel, figuratif et actantiel, concernant un faire ou un fait inscrits dans le temps : le casuel organise un faire comme une "réalisation" et un fait comme du "réalisé". La sémantique casuelle constitue un schématisme spécifique de réalisation. En quatrième lieu, la phrase renvoie par la spécificité de son verbe à un acte possédant une structure interne, un programme particulier avec ses circonstances particulières déterminant l'intelligibilité de la réalisation. Il y a un schématisme spécifique du verbe. Quatre schématismes sont donc à considérer dans la problématique phrastique. Nous parlerons de quatre niveaux (N1, N2, N3, N4) d'analyse et discuterons les conditions de leur intégration.

¹Ce travail est issu d'un séminaire de recherche au Centre de Recherches Sémiotiques de l'Université d'Aarhus. Le Programme de recherche en sémiotique dynamique de ce centre est financé par la Fondation Danoise de la Recherche Scientifique. L'auteur remercie en particulier M. Steffen Nordahl Lund de ses suggestions et de ses communications sur la dynamique pragmatique.

²Voir notre article "Pour une sémiotique de la promesse", in [Brandt, 1991].

³Le clivage emphatique ou contrastif — "J'habite à Copenhague" → "C'est à Copenhague que j'habite" — montre la présence, souvent explicitable, de cette articulation propositionnelle dans le phrastique en général.

Exemple :

— *Qu'est-ce que tu fais ?*

— *J'essaie de réparer ma bicyclette...*

• N 1 :

Réponse à une question (coopération pragmatique).

• N 2 :

"Mon faire est le suivant : essayer de réparer cette bicyclette".

• N 3 :

Je est un agent qui prend en charge un objet et qui *essaie* de le mettre dans un certain état par un acte nommé réparation ; *essayer* n'est pas réussir ; car l'objet résiste aux instruments appliqués ; on répare dans un certain but.

• N 4 :

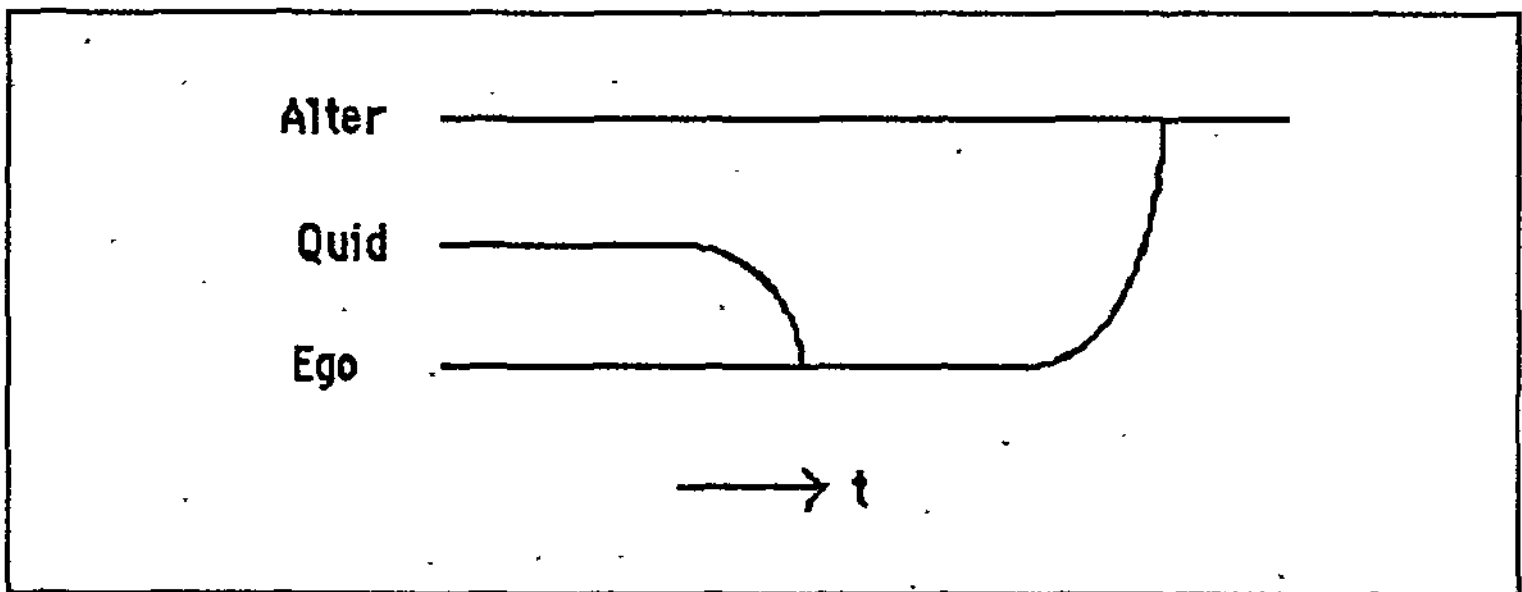
Réparer est un scénario qui concerne un objet dégradable et qui consiste à établir, après une dégradation, un état normativement identique à un état préalable à celle-ci. Réparer introduit un sujet actif intentionnel, un lieu d'activité (par ex. un garage) et un déroulement appelant une indication du style (ex.: "professionnellement"). Réparer représente une norme et peut être évalué ("Tu appelles cela réparer ? C'est du bricolage quelconque...").

Intuitivement, l'approche linguistique considère N4 comme une structure plus profonde que N3, et ainsi de suite jusqu'à N1. Elle imagine par conséquent l'analyse intégrale — solidaire de la réalité linguistique — sous la forme d'une dérivation ascendante, éventuellement générative, de N1 à partir de N4. Or, une telle démarche est vouée à l'échec, du fait que *l'enchâssement* opère dans l'autre sens — N3 enchâssant N4 est enchâssé par N2, à son tour enchâssé dans N1. Rien dans le "plus profond" ne permet de générer les structures "moins profondes" et de surface qui les enchâssent.

Nous proposons d'invertir la démarche, d'abolir toute perspective générative, et donc de considérer la phrase à partir de N1, objectivement structurée par *spécification* jusqu'à N4. Cette idée peut paraître triviale (*top-down* au lieu de *bottom-up*) ; mais sa conséquence est hautement non triviale : la linéarisation de l'énoncé ne peut avoir lieu qu'après la spécification N4, par une stématisation en une instance "plus profonde" que la sémantique casuelle, et non pas à la "surface" N1 ; comme la mise en chaîne linéaire de l'énoncé doit arriver à la fin, pour rester sensible à toutes les sémantiques, et comme la structuration spécifiante "descend", notre proposition ou hypothèse implique cette rupture avec toute démarche générativiste. La métaphore de la verticalité change alors de sens ontologique, si elle en garde un ; il s'agit plutôt, si l'on veut, d'un *zooming* qui spécifie de niveau en niveau le sens du niveau antérieur, avant d'aboutir à une structure stématique, puis linéaire, qui accueille toutes les sémantiques activées, et qui s'insère finalement dans le "quelque chose" pragmatique, pour incarner la *Gestalt* expressive. Le réel linguistique est saisi comme un processus sémantique, qui fait du *temps* de cette spécification progressive, de niveau en niveau, la dimension ontologique pertinente. Le sujet parlant peut donc avoir l'impression à la

fois de “parler” (N1) avant de “penser” (N2-3-4) et de “penser” avant de “parler” : il *marque* sa réplique dans un premier temps, mais il *articule* sa phrase à la suite d’un parcours (N1 → N4), à la fin d’un processus de spécification sémantique. Le geste arrive en effet un peu avant l’articulation linéaire et phonatoire⁴.

• N1. Le niveau de l’échange pragmatique est aussi celui de la structure énonciative. Ego, Alter et Quid interagissent dans une dynamique à trois attracteurs. A Ego, on assigne à juste titre une *assertivité*, une auto-référence ou auto-affirmation constitutive, qui contraste et alterne avec son allo-référence, l’affirmation de l’autre, de Alter, tout aussi constitutive, et que nous proposons de désigner comme son *oblativité*. Entre les pôles ou attracteurs variables de l’assertivité et de l’oblativité, le schématisme énonciatif installe une référence variable à Quid, une *objectivité* principielle qui introduit l’instance de la “troisième personne”, la “chose” au moins virtuelle qui motive la coopération déployée ou prévue entre la première personne et la deuxième, dans la morphologie de l’énonciation. Le réel situationnel de la “scène de communication” qui relie les trois instances — Ego, Alter, Quid — est bien schématisé par l’idée d’un acte de “don”, d’un transfert concernant le Quid ; mais il nous semble nécessaire de rendre compte, dans ce don, du *conflit datif* qui oppose les instances : pour pouvoir “donner” à Alter, en lui “disant” quelque chose, Ego doit d’abord “prendre” le Quid qui les sépare, et que Alter, lui, ne doit pas “prendre”. Ego déploie précisément son assertivité en “prenant” d’abord le Quid, “malgré” Alter ; ensuite, Ego bascule dans l’oblativité, non seulement en offrant le Quid en question, mais en s’offrant à l’autre, en quelque sorte (symboliquement, c’est-à-dire immatériellement) en personne, par la même occasion⁵. D’abord, Ego prend Quid malgré Alter, et l’incorpore ; ensuite, il se donne à Alter (avec le Quid incorporé). Au lieu de représenter ce parcours comme un voyage effectué par Quid, entre Ego et Alter, il semble donc pertinent de le schématiser par le graphe actantiel suivant⁶ :

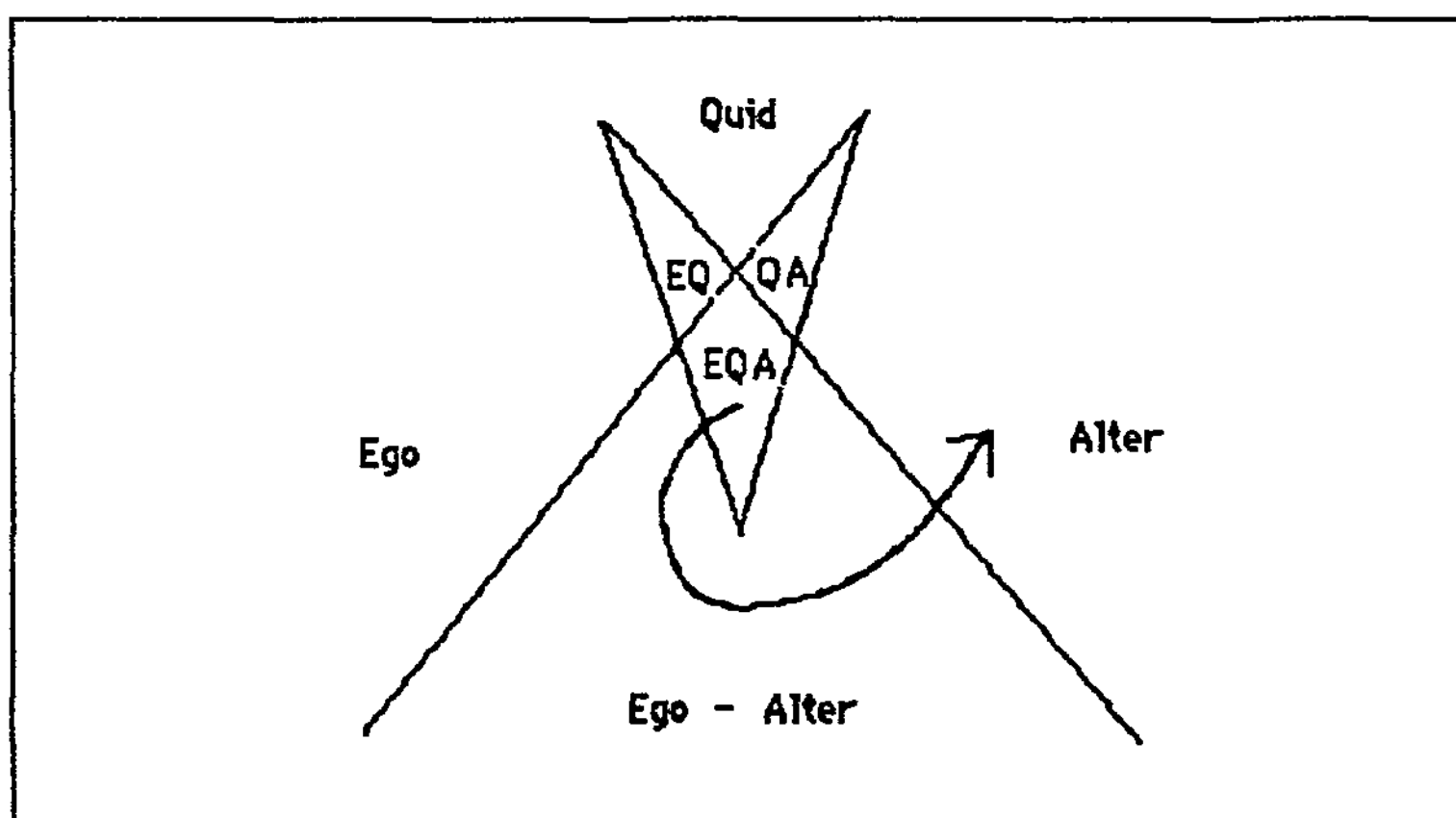


⁴Ce qui explique le fait que le bégayeur — chez qui le décalage entre geste et phonation est dramatique (probablement parce que la motricité gestuelle entrave la phonation) — s’aide en chantant ses phrases : il neutralise ainsi l’impact du geste pragmatique sur l’articulation phono-syntaxique.

⁵Ce que Marcel Mauss savait, qui décrivait le don sans oublier la charge subjectale de l’objet donné, dans l’échange “symbolique”, auquel doit appartenir le don de la parole dans le dire.

⁶Ce n’est pas une métaphore de voyage ; ce n’est même pas une métaphore du tout, mais peut-être la forme élémentaire, sui generis, du lien qu’établit la communication, lien qui crée, par l’intermédiaire d’un objet, une appartenance. Il lie Quid à Alter par un effet de sens génitif.

La situation à trois actants est suivie d'une situation à deux actants, et finalement d'une situation à un seul actant : trois strates successives structurent le scénario. La première capture schématise le geste *assertif* de Ego ; la seconde, son geste *oblatif* subséquent. Alter va donc recevoir Quid "dans la version de" Ego, comme déjà "pris" par lui, "digéré" et interprété par lui. C'est cet intervalle entre les deux captures qui est spécifiable au niveau propositionnel (N2). Le graphe actantiel considéré permet d'étudier la dynamique pragmatique dans la perspective topologique d'une catastrophe *papillon* entièrement interprétable, mais où le parcours envisagé (*donner* présuppose *prendre*) correspond à un chemin de contrôle particulier, qui traverse les trois strates :



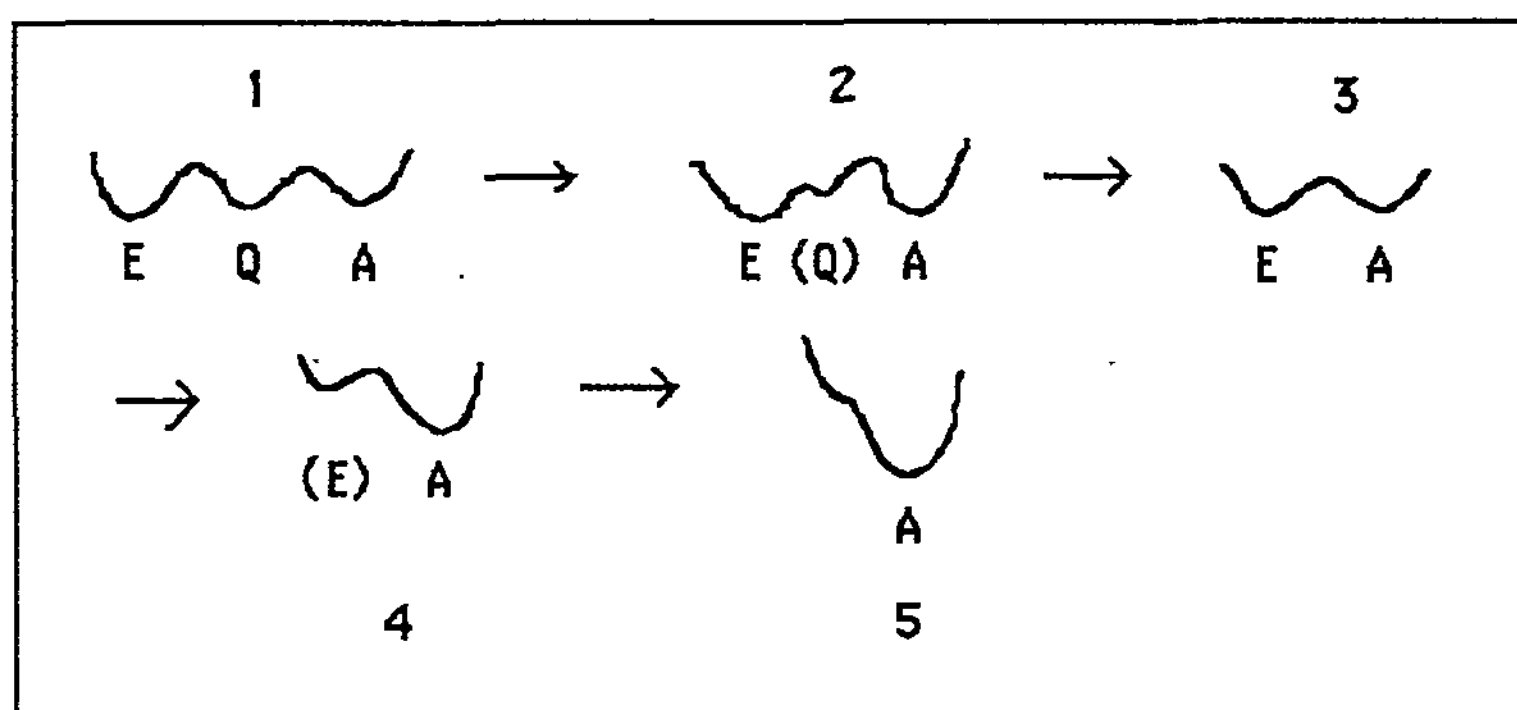
Si la conjoncture Ego — Quid — Alter (EQA) est déformée ou transformée en Ego — Quid (EQ), l'oblativité du sujet disparaît, et l'énonciation devient "narcissique" et "suffisante". Si elle est déformée en Alter — Quid (QA), l'énonciation perd toute charge assertive et devient hyper-oblatif, "prosternée", auto-effaçante. Dans les deux cas, les spécifications aux niveaux subséquents est défectueuse. Si le Quid reste seul par éclipse des interlocuteurs — $(\emptyset)Q(\emptyset)$ —, on a une situation d'adoration objectale, hyper-déictique, qui appelle le silence : le sacré, qui annule l'énonciation spécifiable. Ces pathologies énonciatives semblent pertinentes pour l'étude des dérives dialogiques et pour la compréhension des réactions régulatrices des interlocuteurs. Nous ne pouvons pas développer ici toute la richesse analytique de cette approche.

La perspective la plus intéressante de ce schématisme pragmatique est celle qui concerne la dynamique de l'enchaînement dialogique : les interlocuteurs réels assument à tour de rôle⁷ Ego et Alter, et le parcours se prolonge selon une sorte de plan qui caractérise le *genre* de l'interaction

⁷D'où le mimétisme intonational, par exemple dans les discussions violentes : une conjoncture $E\emptyset(A)$ passe de l'un à l'autre comme une insistance alternante, à moins que l'autre ne "laisse finir" l'un, en faisant un geste d'attente, en "attendant" en effet un moment (E)A pour intervenir ; les tours de parole renvoient aux boucles du chemin de contrôle.

en question. Certains genres sont "objectivistes" — $(E)Q(A)$ —, comme la conversation libre ; d'autres sont plutôt "subjectivistes" — $E(Q)A$ — dans leurs étapes finales et leurs stations de repos, ainsi du genre polémique. Tout enchaînement commence par un retour à la strate centrale $A \rightarrow A(E)^8 \rightarrow EQA$. L'interlocuteur qui "était" Alter, "devient" Ego

Dans le contexte de l'analyse phrastique, l'essentiel est ici le phénomène élémentaire de la déformation de potentiel, correspondant au chemin de contrôle standard qui représente le schéma temporel de la "communication" comme telle, schéma que manifeste déjà le geste alternativement assertif et oblatif du "dire". Il s'agit donc de la série suivante :



Le double sens de l'expression *d'assertivité* — auto-affirmation de Ego, et affirmation de quelque chose à propos de Quid — s'avère motivé. Le Quid est un thème virtuel, et sa captation par l'actant Ego ($1 \rightarrow 2 \rightarrow 3$) signifie son entrée actualisante dans un locus prédicatif, incarné par Ego.

•N2. Si Ego incarne ainsi un locus prédicatif, c'est que le corps (visible, social) de l'interlocuteur qui assume la première personne devient pour l'autre un *container*, un chorème, qui "prend" le Quid thématique à la manière dont une fonction "prend" un argument, ou dont une inférence "prend" une information, une prémisse, pour produire un résultat, une conclusion. Ce locus est un opérateur — *être-P* — qui produit un état prédicatif du thème : *T-être-P* — par l'effet de l'assertivité actualisée ; il produit une assertion⁹. Schématiquement, *T* passe dans Ego, où il devient *P*. C'est cette "corporalité" dynamique de la copule liant le "sujet logique" (le thème) à l' "attribut" (au prédicat) qui explique le fait que la valeur de vérité ou de validité, selon le cas, semble si fortement liée à l'affectivité du locuteur ; l'emphase exprime la force avec laquelle l'assertivité comme opérateur "travaille" le thème dans son locus corporel.

⁸La parenthèse indique l'attracteur faible. Le 'je' de l'énonciation renaît à chaque fois comme une bifurcation *E* de *A*.

⁹En danois 'påstand', qui renvoie à la position stationnaire du corps (le verbe 'stå', all. 'stehen'). En allemand, 'Behauptung' renvoie à la tête de celui qui affirme.

¹⁰'Eu não', disent les Portugais ('Ego non').

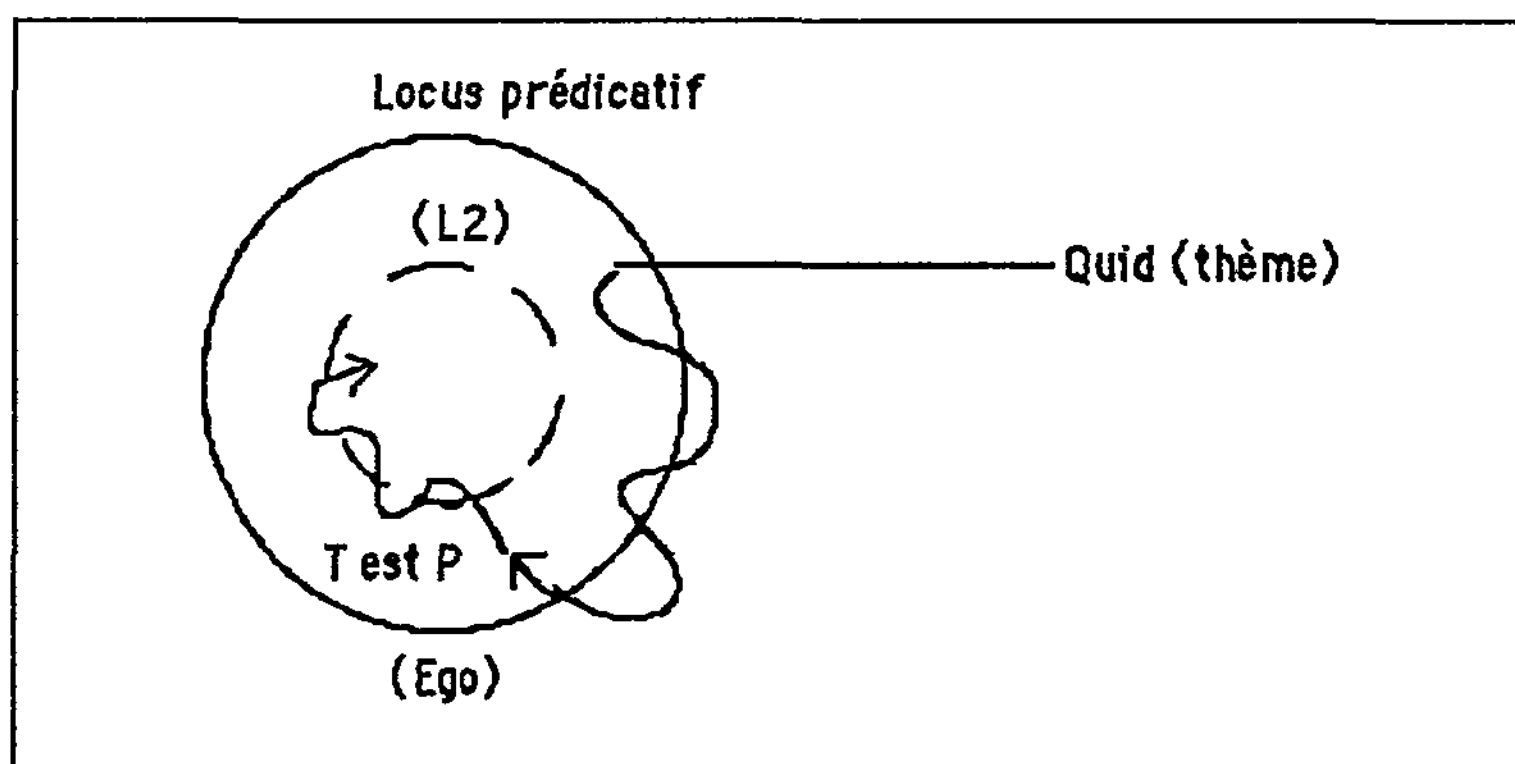
Inversement, la négation est clairement liée à un rejet gestuel, régulièrement emphatique, portant sur l'hypothèse de l'entrée du Quid dans le Ego¹⁰. La certitude et l'incertitude, la probabilité, et en général les degrés de la modalisation épistémique, expriment l'évaluation des circonstances dynamiques de cette entrée chorématique du thème dans le locus prédicatif. Quelque chose est le cas "dans une certaine mesure"...

Cette modalisation épistémique peut prendre un sens performatif, cf. l'interrogation à valeur d'exhortation dans la demande,

— Peux-tu me rendre un service ? (= Je veux que tu le fasses) — ,

si un aspect de l'autre est associé à Quid (ici son service). Dans ces cas, l'oblativité subséquente de Ego se spécifie en reconnaissance ("je suis à toi" : capture de Ego par Alter). Nous n'étudierons pas à cette occasion les détails de la sémantique mixte du performatif ($N1 \rightarrow N2 \rightarrow N1$) ; l'essentiel est que le schématisme prédicatif constitue même pour ce type de phrases une *spécification* chorématique du schématisme communicationnel. Cette transition par spécification du premier au second niveau est possible, pensable, parce que l'approche topologique garde et retient l'homogénéité des structures. Ainsi, le lien assertif entre la dynamique pragmatique et le "contenu" phrastique global est établi.

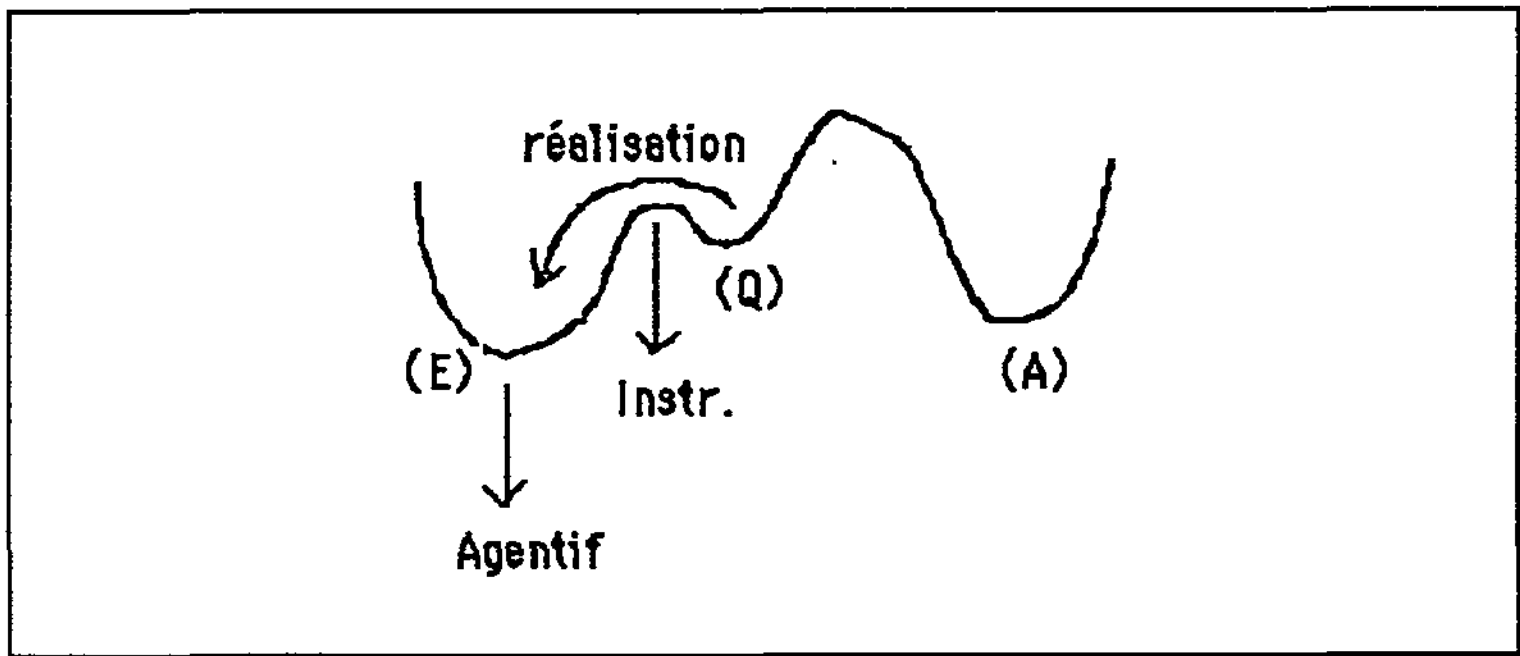
Les propriétés booléennes du propositionnel sont dues au schématisme chorématique de base (qui peut être itéré) :



¹¹Voir à ce sujet notre article [Brandt, à paraître].

• N3. La structure casuelle¹¹ spécifie ensuite les conditions de réalisation de cet "être" qui produit un état du Quid. La barrière qui entoure le chorème est quantitativement variable, et la force d'attraction de ce locus prédicatif l'est également. Tout ce qui représente une réduction de la barrière, c'est-à-dire du seuil qui sépare les puits de Quid et de Ego, relève du cas adjuvant, ou *instrumental*. L'instrumental est ainsi facteur

accompagnateur, un comitatif, associé au cas *agentif* : non pas à Ego, bien entendu, mais à ce qui “aide” Ego à avoir raison, quand il dit que l’état de Quid se réalise¹² grâce à un agent ; le sens de l’agentif est celui d’un accroissement décisif de la force réalisante du locus prédicatif dont Ego est le représentant — malgré lui, si l’on veut¹³, mais “au nom de la vérité”. L’agentif et l’instrumental sont donc complémentaires : l’un augmente la force d’attraction de l’état résultatif, l’autre réduit la résistance de la barrière qui s’y oppose. Soit :

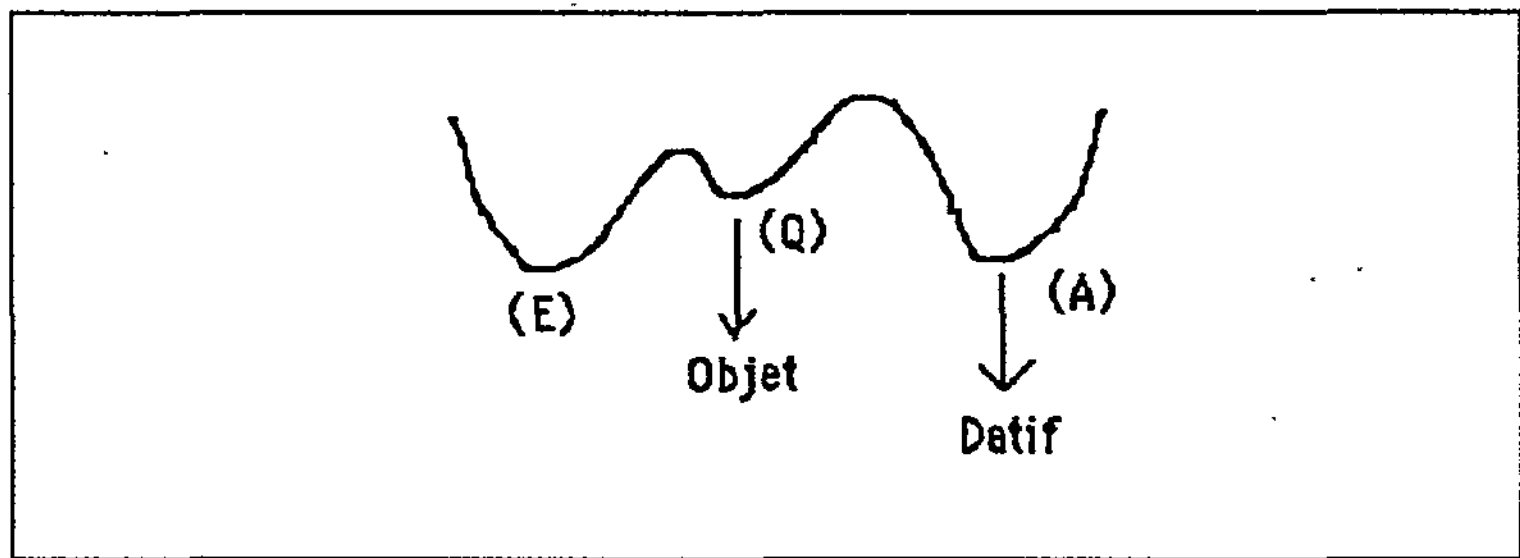


¹²Remarquer le double sens de ‘réaliser’ : ‘rendre qch. réel’ et ‘se rendre compte de qch.’ (de l’anglais). Ego réalise dans le dernier sens, et l’agentif dans le premier. L’ambiguïté est due à cette transition (N2 → N3) de la proposition à la syntaxe casuelle.

¹³Freud disait, en expliquant son principe de réalité, que la vérité est dite concessivement : on regrette que quelque chose soit le cas...

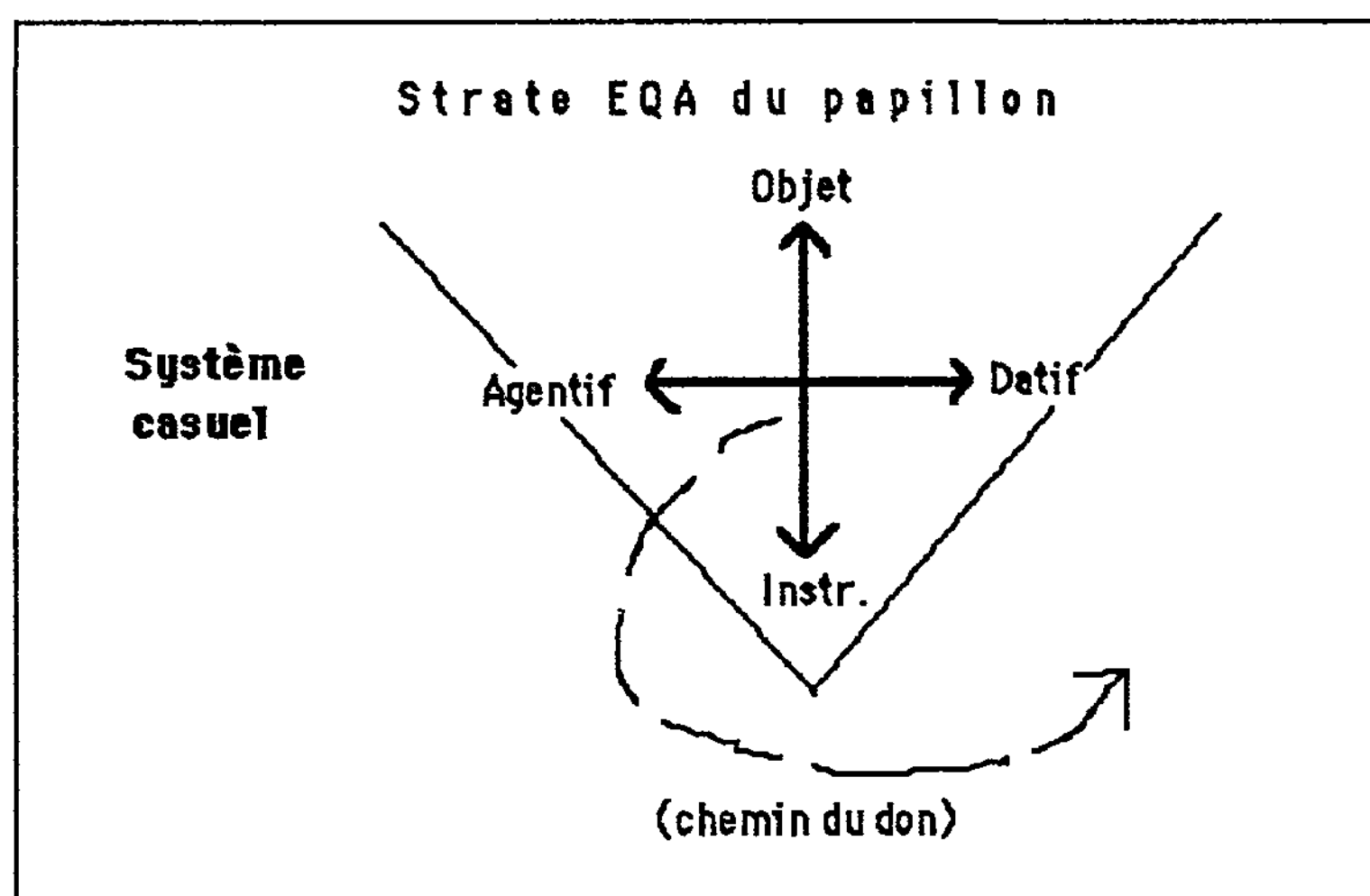
Ce scénario casuel de base correspond assez directement, on le voit, aux graphes dynamiques à chemins (*paths*) et à trajectoires du cognitivisme intuitif. Or, au lieu de figurativiser les cas de manière statiquement spatiale, sous forme de volumes¹⁴, nous nous rapprochons ici d’une phénoménologie temporelle. Si nous ajoutons l’objet, c’est-à-dire le Quid vu dans la perspective de l’agentif et de l’instrumental, et le datif, c’est-à-dire Alter vu dans la perspective de cet objet dont il va recevoir un état réalisé, qui va l’affecter, et si nous les interprétons de la même manière, comme possédant chacun un poids dynamique —

¹⁴Iconisation qui semble souvent maladroite — s’il s’agit par exemple de la gravitation ou d’une énergie abstraite —, et qui appelle alors une hypothèse ad hoc qui justifie les volumes par métaphoricité.



¹⁵On voit qu’un seul seuil du potentiel reste sans interprétation (celui qui sépare Quid et Alter). Le sens de cette absence est essentiel : ce qui sépare sémantiquement l’objet du datif est en effet le VERBE, qui fait que c’est un état de l’objet, et non pas l’objet lui-même, qui affecte la datif. Ex. : ‘Il lui vole son argent...’

— nous pouvons compléter le système casuel élémentaire¹⁵, qui constitue une vectorisation de la strate initiale du papillon :



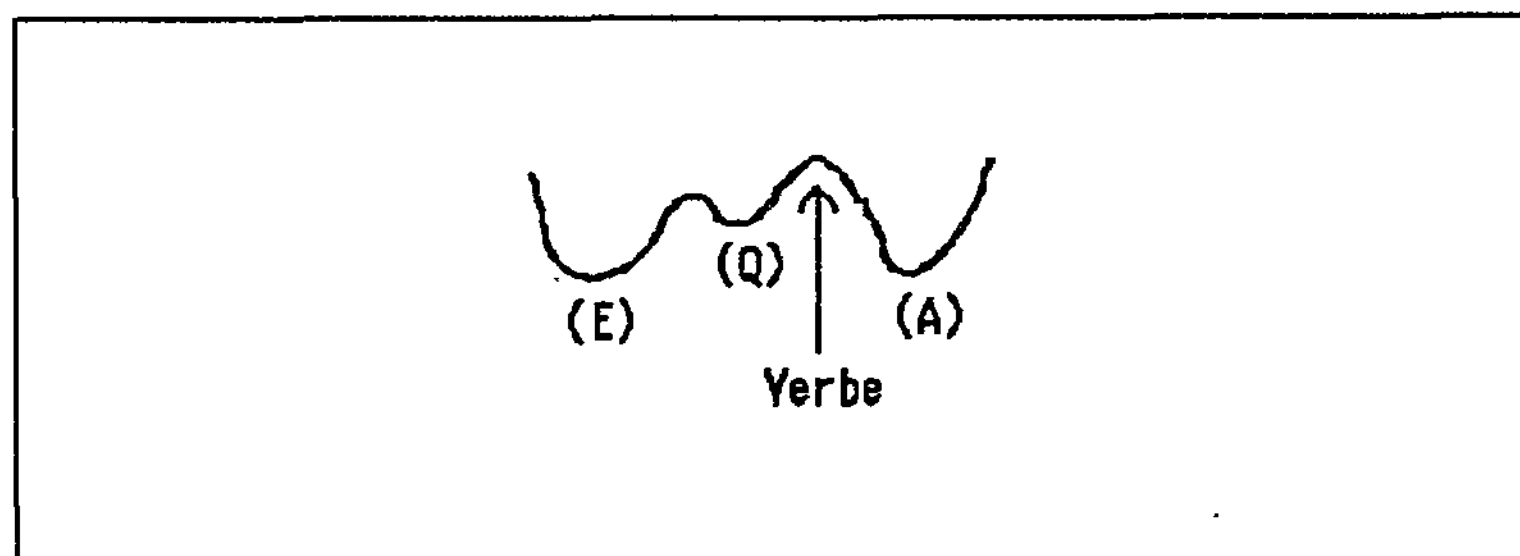
L'inertie de *l'objet* s'oppose naturellement à l'efficacité de *l'instrumental* ; *l'agentif* s'oppose au *datif* dans la perspective de la réalisation d'un état de *l'objet*, avant de le rejoindre en dernière instance, comme nous l'avons vu. Cette tension "horizontale" entre *causalité* agentive et *finalité* dative est une dimension inséparable du sens verbal ; elle est inhérente à tout *faire*. Les alliances naturelles que l'on observe entre, d'une part, *l'agentif* et *l'instrumental*, et, d'autre part, *l'objet* et le *datif*, s'opposent catégoriquement et dramatiquement, dans la mesure où la première, amenant le changement, la transformation (causale), "attaque" ou "agresse" la seconde, qui défend la simple permanence (principe-noyau de toute finalité), mais peut ou doit donc en dernière instance la servir (transformer en vue de conserver) : faire qch. "pour" qn. — éventuellement contre son gré¹⁶.

¹⁶Le datif, cas qui exprime l'instance qui finalise un acte, représente en effet à la fois ce qui s'oppose immédiatement à l'acte et ce qui lui donne un sens ; ainsi, le datif investi subjectalement est le prototype de l'autre éthique — celui qui détermine mon acte, mais qui s'y oppose, activement ou passivement, parce qu'il ne le comprend pas. On agit "contre" celui "pour" qui on agit. Solitude éthique de l'agent... De même dans l'énonciation, dire quelque chose à quelqu'un, c'est parler "contre" l'autre, dans son intérêt.

La structure casuelle est génériquement liée à la notion d'actant en narratologie (cf. le modèle actantiel de Greimas). On peut même penser que le programme narratif principal, qui met en scène la rencontre du protagoniste et du réel, constitue un *blow-up*, une transposition discursive agrandie et figurative, de la structure casuelle. La structure des cas, devenus des actants narratifs, assurerait ainsi l'intelligibilité naturelle de la narration. La narrativité serait en principe une expansion discursive de la casualité.

• N4. La sémantique du verbe de la phrase, celle qui ne concerne plus les circonstances de la réalisation d'un faire, mais le contenu spécifique de ce faire, est donc un sens *sui generis*, compatible avec la structure casuelle, mais indépendant d'elle. Ce sens verbal est préfiguré dans le

potentiel constitutif, à savoir par le seuil qui sépare le Quid et Alter (cf. l'objet et le datif) :



Le verbe “survole” ainsi l'ensemble dynamique de la phrase. Si la phrase effectue génériquement l'arrivée de Quid à Alter, par le détournement assertif de Ego, le sens verbal spécifie en revanche à chaque fois un mode d' “arrivage”. L'objet n'atteint sa finalité que par le détournement causal d'un faire ; mais le détournement est d'importance variable. La quantité ou densité verbale qui fait la différence entre la copule apparemment vide et le verbe transitif plein, et qui se manifeste par l'importance ou le poids de la structure casuelle dans la construction grammaticale qui entoure le verbe, correspond à la hauteur variable de ce seuil : et plus il s'élève, et plus devient importante la structure casuelle sur le versant Quid—Ego ; et plus se prolonge aussi *le temps de la réalisation*. Comparons les deux phrases suivantes :

(1) Cet objet est à moi.

(2) Il me construit une maison.

Dans (1), la copule relie l'objet et le datif, en faisant du premier son complément sujet. Le message pragmatique pourrait être : “Si tu le prends, cet objet, tu dois me le rendre”, ou simplement : “Donne-le-moi”. Prendre et rendre, donner, sont des actes qui esquissent le détournement par Ego. Chose remarquable, l'expression dative *à moi* fait parler la phrase (1) du point de vue de Alter. Les pronoms personnels¹⁷ permettent, autour du verbe, une *redistribution* complète des catégories de l'énonciation, et *sépare ainsi l'énoncé de l'énonciation*. Ego dans l'énonciation devient le complément d'objet indirect, datif, dans l'énoncé. Le travail supposé, dans l'énonciation, effectué par Alter, et qui n'est pas représenté dans l'énoncé, puisque le verbe “vide”, la copule, n'admet pas le cas agent, consiste simplement à procurer et à rendre, conditionnellement, ou alors à donner, ou à aider à retrouver, inconditionnellement, l'objet, c'est-à-dire à faire en sorte que l'objet arrive à sa destination dative (par exemple en prenant l'objet à un tiers pour ensuite le transmettre à l'autre). En revanche, la

¹⁷Nous pensons en effet que les syntagmes nominaux en général peuvent investir librement les positions préstructurées par le dynamisme casuel, précisément parce que le sémantisme pronominal, appartenant à la morphologie sémantique du verbe, a “libéré” les positions de la référence directe à l'espace pragmatique. Ce point est capital pour toute analyse phrastique.

¹⁸La référence à un tel 'il', à des entités nominales extérieures à la situation de communication, référence qui est capitale pour le langage humain, est possible dans la mesure où il s'agit d'un Quid (d'un quidam) hérité d'une phrase voisine ; en principe, c'est donc casuellement d'abord un objet ; ensuite, il est "délogé" et transposé en agent, quand le verbe qui le concerne exige un autre objet. La phrase intransitive 'Il construit' n'exprime plus un acte, et son sujet grammatical n'est plus un agent, mais reste casuellement objet. 'Construire', tout court, est un faire moins plein et plutôt un mode d'être, comparable à 'Être à moi'.

¹⁹La construction passive sans agentif — la bicyclette est réparée — est résultative.

phrase (2), dont le datif est également à la première personne — transposition d'un Ego d'énonciation — impose à l'agent explicite, à la troisième personne¹⁸, qui figure dans le complément sujet, un travail de construction matérielle, longue et compliquée, puisque le verbe est "plein". Les deux verbes introduisent donc un *sujet grammatical*, dont la source est n'importe lequel des cas. Le sujet grammatical n'est donc pas un cas, c'est un élément nominal issu de la sémantique du verbe, et que le verbe investit dans la structure casuelle selon un style déterminé de réalisation — d'où la différence aspectuelle que l'on observe entre les phrases équivalentes de diathèse différente¹⁹.

La sémantique du verbe introduit donc le "nominatif", le sujet grammatical, qui n'est pas un cas, mais une extension nominale inhérente à la mise en perspective de la temporalité et de la scène, pour ainsi dire, du contenu verbal, iconique comme tout le contenu lexical qui s'introduit à ce niveau. Ce que Tesnière appelait les circonstants : le *locatif* et le *modal* (les adverbes ou les syntagmes adverbiaux indiquant le lieu ou le moment et la manière d'un acte ou la valeur épistémique d'un énoncé), constitue une autre extension de ce sens verbal. Encore une fois, c'est le potentiel de base qui est investi. Le *locatif* est maximalelement présent dans les constructions impersonnelles au verbe intransitif, à dynamique faible, dont le sujet grammatical ne fait que projeter l'ombre du contenu verbal ; ici, le *locus* prédictatif devient littéralement le réceptacle de la figurativité verbale :

(3) Il pleut sur la ville.

La spatialité du circonstant locatif reprend Ego comme lieu assertif et lieu de réalisation (ce que le poète sait bien, qui enchaîne : "comme il pleure dans mon cœur"). Dans la phrase (3), le Quid nominatif, ombre (pluie) du verbe (pleuvoir), est en effet versé dans le chorème (la ville), qui reste proche de la première personne de l'énonciation. Le lieu remplace pour ainsi dire le sujet personnel. Si, par contre, la phrase manifeste à la fois un agentif et un locatif, le minimum de potentiel Ego est réellement investi deux fois, et l'agent est *situé* :

(4) Il bâtit sa demeure à Paris.

Ce que la structure casuelle fait se réaliser ici, à savoir le devenir-bâtisse de la demeure, a lieu à l'endroit où le sujet agentif *il* se trouve pendant l'acte (sinon, il faudrait dire : "il fait bâtir (par autrui) ..."). Le circonstant locatif exprime littéralement la spécification du domaine véhiculé par le sens verbal. D'ailleurs, la visibilité iconique de la scène située s'adresse à Alter dans l'énonciation²⁰ de la même manière que le fait la visibilité du locuteur lui-même, qui peut dramatiser devant lui par ses gestes, par tout le théâtre de sa motricité, à la fois ce qui se passe, le

²⁰C'est l'énonciataire qui est l'observateur de la scène phrastique : 'La bicyclette est réparée, tu vois...' — peut ajouter l'énonciateur.

lieu où cela se passe, et l'agent en question. C'est donc pour ainsi dire l'énonciataire qui "voit" la scène et qui est le sujet-spectateur spécifique du locatif.

Le circonstant *modal* exprime apparemment deux choses totalement distinctes : le mode de réalisation et la valeur épistémique de l'assertion. Le mode de réalisation (violemment, brusquement, lentement, facilement, difficilement) et la valeur épistémique (vraisemblablement, probablement, sans doute, pas du tout) ressortissent pourtant à la même attention portée à la résistance modalisante offerte par le seuil qui sépare l'objet de l'agent, et qui est modifié par l'instrumental : si la réalisation d'un acte est "facile/difficile" (pour un agent), elle est aussi bien "probable/improbable" (pour un évaluateur). Les deux caractéristiques modales — descriptive et évaluative — semblent souvent équivalentes du point de vue dynamique²¹. Mais elles ne le sont pas *temporellement*. Le circonstant modal descriptif est rétrospectif, alors que l'évaluatif est prospectif. Considérons contrastivement les phrases suivantes :

(5) Il pleut violemment à Paris.

(6) Il pleut certainement à Paris.

cf. (5') Il ne pleut pas violemment à Paris.

et (6') Il ne pleut certainement pas à Paris.

La scène invite dans (5) le parisien à se couvrir, s'il veut sortir ; la négation s'applique à l'adverbe descriptif et l'affaiblit, de sorte que dans (5'), il pleut toujours, mais de manière plus douce. En fait, il pleut *déjà* — la négation n'y peut rien. Quand la négation arrive, c'est trop tard — l'événement de la pluie est réalisé et envisagé rétrospectivement. Dans (6), en revanche, la pluie est attendue, ce pourquoi on ne peut *pas encore* la décrire ; la négation, qui est elle-même un circonstant modal épistémique, ne s'applique plus à l'adverbe, mais l'adverbe s'applique à elle, ce qu'exprime l'ordre des mots ; la négation annule ici la pluie, parce que cet événement, non réalisé pour l'évaluateur, est envisagé prospectivement : l'adverbe se réfère à l'attente épistémique, qui, elle, succombe à la négation. L'expérience du temps par rapport au présent est nécessairement celle d'un sujet sur place, il s'agit donc, constitutivement, de Ego ; c'est par conséquent l'énonciateur qui est le sujet-évaluateur du circonstant modal. Si l'énonciation est bien représentée ou inscrite dans les circonstants, elle l'est donc différemment dans le locatif et le modal.

A ce niveau de l'analyse (N4), on a vu que la phrase se lexicalise. Le fonctionnel dynamique, la structure casuelle, rencontre ici un phénomène nouveau et totalement distinct de tout ce que les niveaux précédents ont pu offrir, à savoir le sens lexical et sa base, la catégorie du mot²². Les classes de mots, qui, on le sait, ne se réduisent jamais aux fonctions syntaxiques,

²¹Ce pourquoi il est agrammatical de dire que qch. est violemment le cas : le verbe "vide" horizontalise le potentiel, affaiblit le dynamisme de la réalisation.

²²Nous ne discuterons pas ici la provenance cognitive de l'entité mot comme telle, c'est-à-dire l'origine et la structure du signe saussurien, arbitraire ; il nous semble certain que ce phénomène dépasse la problématique phrastique, puisque c'est une structuration qui opère sémiotiquement bien au-delà de la parole et du texte écrit, et qui est à l'œuvre, en deçà, dans le fonctionnement même de la perception, de la mémoire et de la pensée inférentielle (cf. les métaphores selon G. Lakoff).

²³Ce résumé peut paraître excessivement simplificateur; on pourrait se contenter terminologiquement de distinguer les "nominoides" et les "verboïdes", ou les deux types de Brøndal R et D, cf. [Actes du colloque, 1989].

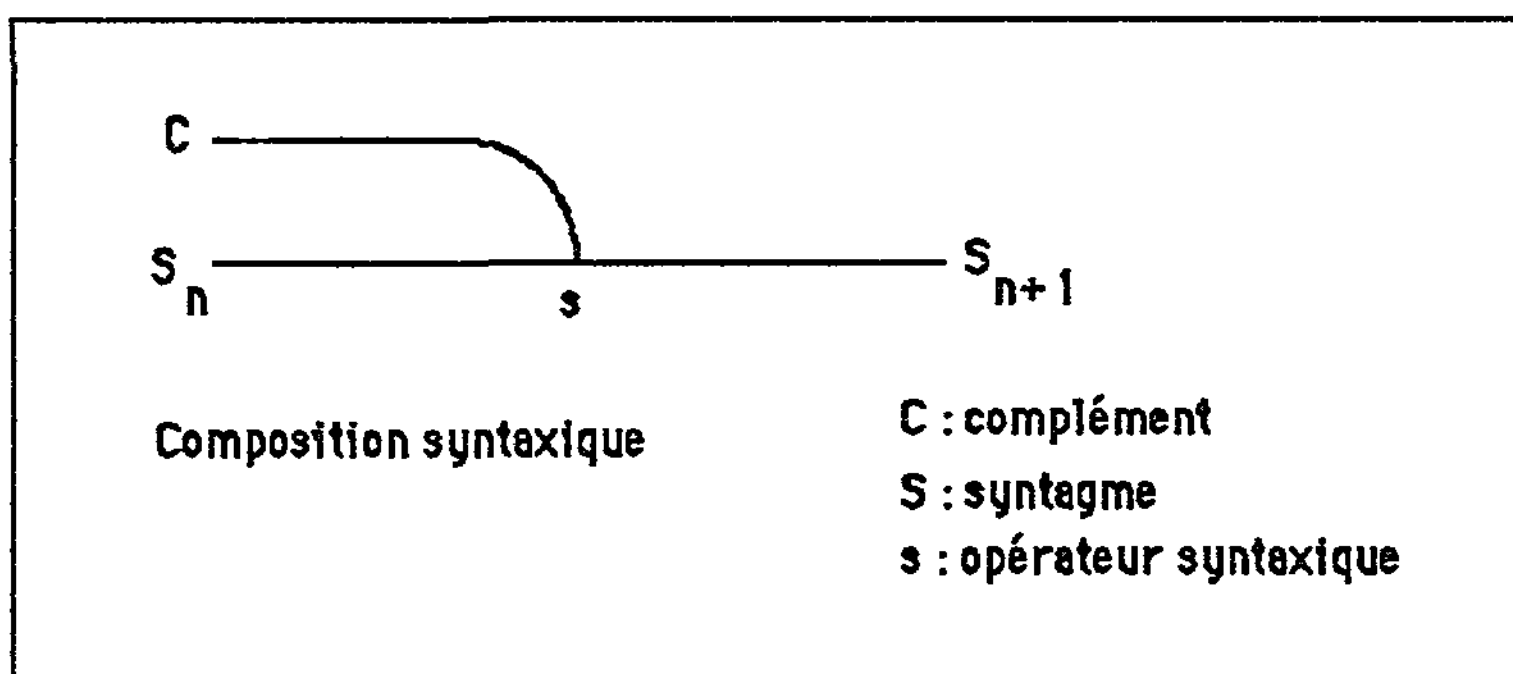
²⁴Pour être exact, l'instrumental est un nominal, et il investit pourtant un maximum, contrairement à ce que nous avançons ici.

Comme entité nominale, c'est en fait un objet appartenant à l'agent, approprié par lui, et donc lié à lui par un lien génitif, cf. *supra*; ce qui explique que son puits disparaît. Mais dynamiquement, c'est une pure force appliquée au seuil qui fait résister l'objet de l'acte. Sinon, ce serait un comitatif pur ('Il se promène avec son chien').

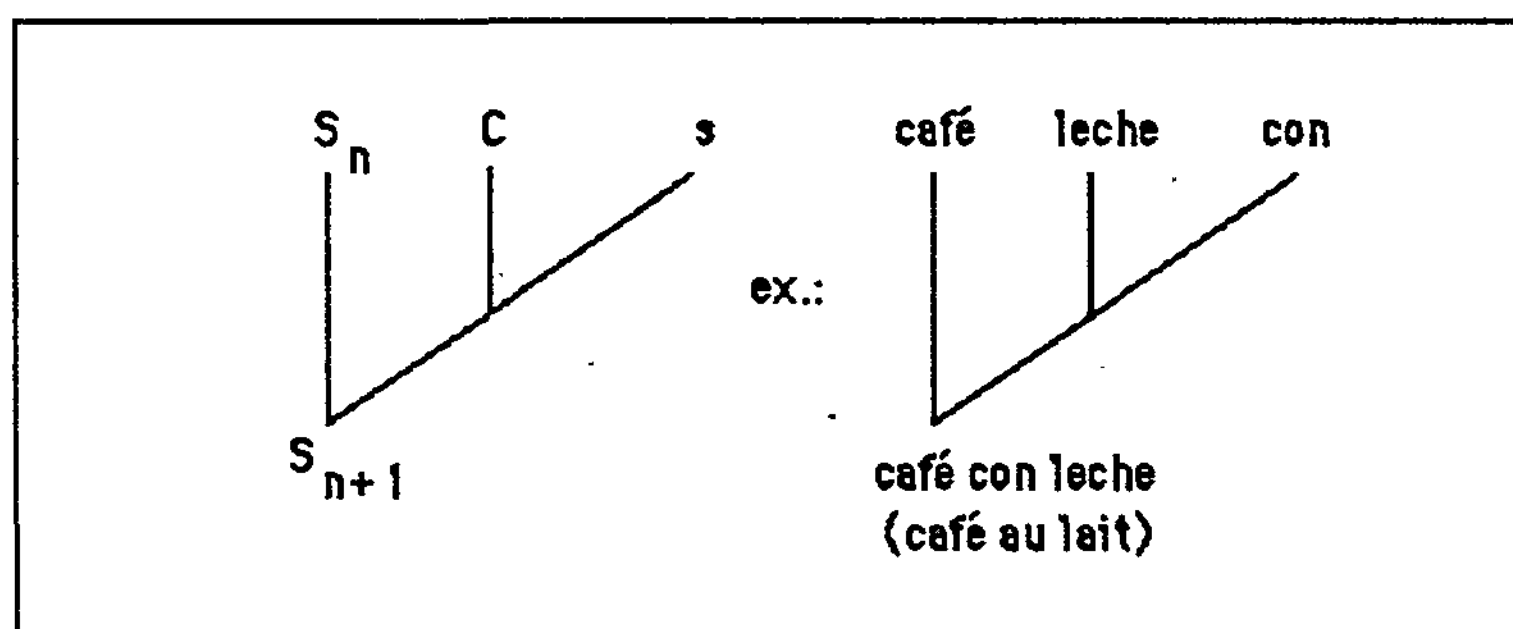
s'articulent cependant selon un principe assez simple, qui explique leur *compatibilité* avec le fonctionnel qui les reçoit. La grande division naturelle qui sépare les *nominaux* (substantifs, adjectifs, pronoms) et les *verbaux* (verbes avec adverbes, prépositions et conjonctions)²³ suit en effet l'articulation de la topologie dynamique que nous avons suivie à travers les étapes constitutives de la spécification (N1, N2, N3, N4). Dans l'espace interne de cette topologie, la distinction entre *minima* et *maxima* est capitale. Partout, la mise en relation, dans l'espace et dans le temps, des instances intelligibles (à partir de Ego, Quid, Alter) est étroitement liée aux maxima, aux seuils, alors que les instances mêmes, qui se trouvent mises en relation, sont identifiées par référence formelle aux minima, aux puits, du potentiel²⁴. Les nominaux sont simplement les classes qui s'investissent dans les minima; et les verbaux, celles qui s'investissent dans les maxima. Le résultat verbo-nominal que nous appelons une phrase est organisé par un entrelacs possédant deux propriétés formelles remarquables: il est hiérarchique, et il forme un ensemble fini (il n'est pas universellement récursif). La forme syntaxique de la phrase, comme construction grammaticale, exprime un ordre générique existant entre les compléments possibles, et exprime la finitude de ce possible. D'une part, on a vu que le verbe domine la "construction"; la cause de ce choix naturel est à chercher dans la structure dynamique de la communication (séparation entre Quid et Alter, détour par Ego, etc.), qui présente comme son pivot le maximum aliénant Alter. Toute la structure dépendancielle de la syntaxe est la conséquence de cette particularité de base. D'autre part, la série des spécifications est finie; ses produits, de niveau en niveau, forment par conséquent un ensemble dont on peut rendre compte de manière en principe exhaustive, et que la phrase constructible peut intégrer.

Les arbres phrastiques que l'intuition nous demande de considérer, construction pour construction, et que la linguistique actuelle semble en train de perdre tout espoir de théoriser — de linguistique générative en sémantique cognitive —, représentent effectivement une organisation qui *existe* et qui appartient à la réalité du langage. C'est une organisation relevant de N4, qui possède les propriétés formelles de hiérarchie et de finitude que nous venons de souligner, elle est donc *canonique*, et elle est indispensable à la *linéarisation* qui donne finalement à la phrase une manifestation dans le registre phénoménologique de la motricité expressive et phonatoire des humains.

Nous nous limiterons ici à évoquer la configuration générale de cette organisation canonique, que nous appelons, après [Tesnière, 1965 (1959)], la *stemmatique*. Elle introduit un motif formel — celui qui a fait croire à la récursivité universelle d'un "calcul" génératif — qui caractérise toute phénoménologie de la composition (par ex. en gastronomie): qch. est *modifié* par l'ajout et l'intégration de qch. Soit:



Ce schéma simple de capture peut être itéré en adoptant des chemins appropriés sur la topologie de contrôle (le *cusp*), de sorte que de nouveaux compléments C arrivent sur S et le modifient, ou de manière à ce que S devienne le complément C' d'un nouveau S', et ainsi de suite. Nous avons proposé [Brandt, 1973] la graphie compositionnelle suivante pour cette opération syntaxique — proche de la logique applicative de [Saumjan, 1971] —, qui se prête à la fois à la description formelle des constructions et à l'analyse de la structure canonique :



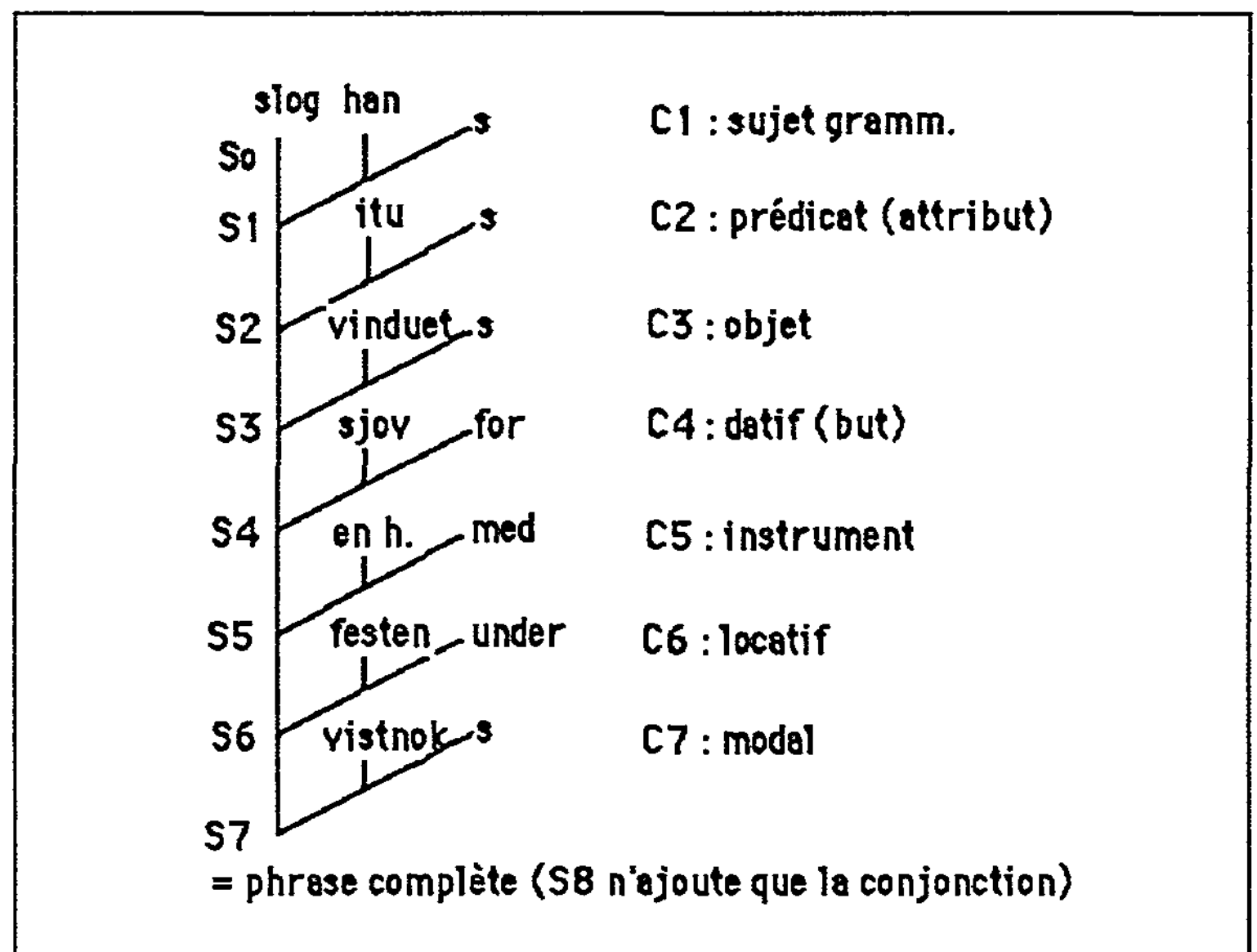
La phrase est à concevoir, dans cette perspective, comme un syntagme verbal, *sémantisé* par la structure casuelle et par sa spécification verbale, et de manière *canonique*, c'est-à-dire dans un ordre compositionnel qui reste constant, quelle que soit la construction. La linéarisation expressive, autre aspect décisif pour toute compréhension de la syntaxe, est en principe une projection de cette composition en deux dimensions sur une chaîne en une dimension — par itération d'une règle de base comme la suivante : $s(S, C) \rightarrow S^s C$, cf. l'exemple, suivie de règles sensibles à la construction, et de règles stylistiques — ; le *stemma* canalise ainsi le sens vers la linéarité expressive qui va finalement couvrir le Quid pragmatique, et permet inversement le déchiffrement non linéaire des chaînes phrastiques.

²⁵Les "bonnes" constructions laissent en réalité toujours un certain nombre de cases complémentielles vides ; celles-ci seraient plutôt difficiles à trouver dans la littérature.

La phrase danoise qui suit montre le fonctionnement de la complémentation canonique²⁵ de la phrase syntaxique :

Under festen slog han vist vinduet itu for sjov med en hammer.

Pendant la fête, il a sans doute brisé la fenêtre avec un marteau pour s'amuser.



Chaque case de complément peut évidemment enchâsser des constructions entières, nominales ou verbales. L'ordre S0 → S7 est invariable, et un Sn n'est rempli qu'une fois, si la construction ne le laisse vide. La première partie de la séquence (S0-S2) exprime la sémantique verbale directement ; la seconde partie (S3-S5) est casuelle (l'agentif a été subjectivé et apparaît en C1, alors que la construction passive le fait apparaître en C5, position²⁶ qui est ici investie par son instrumental) ; et la dernière partie de la séquence est circonstancielle. Cette régularité de la séquence "génétique" explique le caractère mécanique de la mise en parole et de son déchiffrement.

D'ailleurs, la morphologie lexicale et grammaticale — flexions, accords, prépositions, conjonctions —, c'est-à-dire la saillance des classes fermées en général, sur le stemma comme dans le flux lexical de la phrase-chaîne, permet de planifier et de détecter dans le *parsing* naturel la disposition globale de la phrase construite.

Comme cette analyse intégrale de la problématique phrastique est naturaliste — elle vise en effet "ce qui se passe réellement" dans la partie du cerveau accessible à la conscience —, il est légitime de se poser en

²⁶C5 reçoit l'instance "par laquelle" qch. se réalise, la cause ou le lieu de départ, la source ou l'origine d'une réalisation. Une sorte d'ablatif.

dernier lieu la question de savoir comment un tel processus est corrélé au savoir mémorisé. Quelle est la forme, en effet, du savoir mentalement représenté ? Il serait insuffisant de penser aux structures prédicatives de N2, bien trop vagues²⁷, et il serait contrefactuel²⁸ de croire que les structures stemmatiques de N4 puissent rendre la forme mentale du savoir. En revanche, la structure casuelle de N3 rend essentiellement compte des liens de causalité et d'intentionnalité qui s'établissent dans le savoir entre les catégories de la connaissance (on peut connaître par exemple la tendance d'un type d'événement à provoquer tel autre événement, par agentivité ou instrumentalité latérale). Notre conclusion est que *c'est par N3 que s'établit le contact entre le langage et le savoir*. Ce sont les structures dynamiques de ce niveau, mais sémantisées par toutes les autres structures linguistiques dans l'associativité émotionnelle de la mémoire, qui assurent la circulation entre la parole et la pensée. Les inférences métaphoriques lakoffiennes fonctionnent invariablement par projection de schématismes dynamiques, en superposant des structures casuelles. Le sujet de la pensée est un Ego qui voit mentalement se réaliser ou devenir possible en lui la réalisation ou la possibilité qu'il doit penser telle dans le monde indépendant de lui, monde dont le langage le fait être le théâtre.

27 L'Intelligence Artificielle simpliste du style de programmation Prolog semble pourtant montrer la force du préjugé selon lequel le savoir serait un réseau logique prédicatif.

28 Parce que des constructions différentes peuvent manifestement véhiculer le même savoir.

*Centre de Recherches Sémiotiques
Université d'Aarhus*

Fondation Danoise de la Recherche Scientifique

Bibliographie

[Actes du colloque]

1989, «Linguistique et sémiotique : actualité de Viggo Brøndal», Actes du colloque tenu à la société Royale à Copenhague les 16 et 17 octobre 1987, in *Travaux du Cercle Linguistique de Copenhague*, Vol. XXII.

BRANDT (P. A.)

1973, *L'Analyse phrastique*, Bruxelles-Paris, Éd. Aimav.

1991, "Pour une sémiotique de la promesse", *Qu'est-ce qu'une promesse ?*, *Poetica et Analytica*, Supplément 1, Aarhus University Press.

1992, *La Charpente modale du Sens : pour une sémio-linguistique morphogénétique et dynamique*, Aarhus-Amsterdam, Aarhus University Press-John Benjamins.

1994, *Dynamiques du Sens : études de sémiotique modale*, Aarhus University Press.

à paraître, "Éléments de sémiotique dynamique", *Hommage à René Thom*, M. Porte éd., Éditions de l'ENS de Fontenay-St Cloud.

SAUMJAN (S. K.)

1971, *Principles of Structural Linguistics*, La Haye-Paris.

TESNIERE (L.)

1965 (1959), *Éléments de syntaxe structurale*, Paris, Klincksieck.

La Sémiotique cognitive

Les sciences de l'esprit entre la nature et la culture

Pierre Ouellet

La pensée sémiotique des quinze ou vingt dernières années a connu d'importantes mutations, dont on peut repérer au moins deux moments clefs. Le premier est celui où l'on a renoncé à considérer les systèmes symboliques comme des systèmes clos et autonomes, n'obéissant qu'à leurs lois propres et à leur nécessité interne, dont il suffirait au chercheur d'établir la logique d'un point de vue formel ou opérationnel pour en comprendre les fondements et les aboutissants, réductibles à quelque structure élémentaire capable d'engendrer un nombre indéfini de phénomènes. Ce premier moment est contemporain de l'avènement des sciences cognitives, qui ont fait prendre conscience au sémioticien que les structures sémio-discursives ne sont pas des entités formelles indépendantes du fonctionnement de l'esprit humain et qu'il faut pour les comprendre les rapporter ultimement aux "facultés de l'âme" comme dirait Kant ou, dit plus modestement, aux différents "modules de notre activité mentale", comme le suggère Jerry Fodor. Les processus de signification ont bien sûr à voir avec les contraintes formelles des langues naturelles et les normes régissant les pratiques discursives — linguistes et sémiologues l'ont largement démontré —, mais ils s'appuient d'abord sur le fonctionnement de la conscience des sujets qui y participent, non seulement par leur faculté de parler, qui est l'un seulement des modules de notre esprit, mais par leur mémoire, leur imagination, leur pouvoir inférentiel, leur intentionnalité et, plus fondamentalement encore, leur perception, leur mode de présence corporelle au monde, leur sensorimotricité — facultés qui ont toutes un rapport plus ou moins étroit avec notre capacité de signifier, soit de produire et de comprendre des entités de nature symbolique.

L'intelligence artificielle nous a montré que le robot le plus simple ne peut comprendre et produire des signes qu'à la condition de posséder une mémoire et un pouvoir inférentiel qui lui tiennent lieu d'esprit et que toute tentative de complexification de telles machines rudimentaires exige

qu'on leur fournisse de surcroît des organes de perception, un substitut de conscience réfléchissante d'où elles tirent un simulacre d'intentionnalité ou d'identité à soi, de même qu'une image de leur présence plus ou moins mobile dans le monde spatio-temporel qui les environne. Il fallait donc, après leur avoir donné un esprit, tenter de leur donner un corps et, qui plus est, un *corps propre* au sens de la phénoménologie. Les neurosciences et la psychologie expérimentale peuvent venir au secours de l'IA dans ses tentatives de modélisation des opérations et des structures qui permettent le passage des "simulacres du corps" aux "simulacres de l'esprit", en établissant un lien causal entre les capteurs périphériques de tout système sensible à son environnement et le processeur central en quoi consiste la structure conceptuelle abstraite où les données de cet environnement font l'objet d'un calcul de type inférentiel fondé sur les entités formelles qui les symbolisent (v. [Churchland, 1990 (nouv. ed.)]). Une telle approche empirique ne peut toutefois, comme on l'a maintes fois souligné depuis l'ouvrage fameux d'Hubert Dreyfus sur les limites de l'IA [Dreyfus, 1984 (1979)], rendre compte du deuxième aspect de la corporéité dont j'ai parlé, celui de la *propriété* du corps, qui suppose une activité proprioceptive dont les racines plongent dans le "sentiment d'existence" au sens kantien — le *corpus sui* d'un Maine de Biran — et un type particulier d'intentionnalité, définie par la présence charnelle du sujet au monde.

C'est le deuxième moment clef de la pensée sémiotique des dernières années, où l'on a montré que l'ouverture du langage sur les autres facultés de l'esprit et l'ouverture de ce dernier sur son propre support physique, en ses dimensions organique et neuronale, ne peuvent suffire pour comprendre le fonctionnement des systèmes symboliques. Il faut aller plus loin et inclure l'expérience phénoménologique de soi, de l'espace et du temps, dans la perception et la motricité, notamment, pour saisir notre mode de présence au monde qui est le support ultime de toute intelligence et de toute capacité de symbolisation, dont le fondement réside dans l'imagination et l'intuition sensible, lieux et objets d'une esthétique transcendantale, plutôt que dans les structures conceptuelles abstraites de la logique propositionnelle ou les processus neuronaux a-symboliques immanents à notre système neuro-physiologique. C'est ce qui est en train de s'opérer dans la pensée post-structurale d'aujourd'hui, grâce au retour en force de la phénoménologie au sein des études sémiotiques et des sciences cognitives.

Un danger guette toutefois les recherches sémiolinguistiques dans ce reflux vers l'expérience vécue et la présence charnelle au monde. Si celles-ci permettent d'éviter tout réductionnisme bio-physique, à quoi nous confinent les hypothèses les plus fortes du cognitivisme classique,

elles n'en participent pas moins du gigantesque mouvement de "retour amont" où les sciences de l'esprit et les sciences de la culture — dont la sémiotique fait partie en tant que "psychologie sociale", dirait Saussure, "anthropologie de l'imaginaire", dirait Greimas — risquent de perdre de vue leur "aval", non plus leur source ou leur fondement, devenu l'obsession des sciences, mais leur visée et leur finalité, voire l'intentionnalité propre aux systèmes symboliques qu'elles sont censées étudier. L'enjeu, bien sûr, n'est pas de revenir à la fétichisation du langage et à l'autonomisation du discursif que le *linguistic turn* de la philosophie analytique et le nominalisme de la pensée structurale classique auront mis à l'honneur dans les dernières décennies, mais de redonner toute son épaisseur, sémiotique et historique, aux processus de symbolisation et de représentation mentale ou linguistique auxquels notre vécu phénoménologique donne lieu dans toute activité cognitive. Autrement dit, l'ouverture en amont que les sciences de la cognition et la phénoménologie de la perception ont permis de pratiquer dans l'édifice jusque là trop fermé sur lui-même du langage et du discours, ne doit pas masquer cette autre ouverture, tout aussi nécessaire, sur l'aval de notre activité cognitive et symbolique, qui vise ultimement non pas l'adéquation originare à un vécu de conscience perçu comme sa cause et sa raison d'être mais un ensemble indéterminé d'actions et d'interactions plus ou moins intentionnelles avec l'environnement naturel, humain et discursif extrêmement plastique, que l'on cherche à modifier, à transformer, selon des mécanismes qui donnent lieu à ce que nous appelons l'Histoire.

La question de la perception, par rapport au langage et au discours, marque bien la distance qui sépare ces deux types d'ouvertures : sur les conditions *a priori* de l'expérience et sur l'interprétation *a posteriori* des représentations symboliques qu'on en donne. On parle de *perception sensorielle* et de *perception sociale* : d'une part on décrit les mécanismes neuro-biologiques ou le vécu phénoménal d'un acte de perception immédiate et d'autre part on décrit la façon dont une communauté interprétative historiquement et socialement déterminée se représente mentalement un phénomène au sein d'une activité discursive donnée — comme cela se pratique depuis longtemps en psychologie sociale et en ethnométhodologie ou, plus récemment, en esthétique de la réception au sein de la critique littéraire. Dans les deux cas, diamétralement opposés, on parle de *perception* mais comme s'il s'agissait là d'une pure coïncidence due à la seule homophonie de termes qui renverraient en fait à deux notions irréductibles l'une à l'autre. Prises comme je viens de le faire, par l'un et l'autre bout du spectre qui sépare la sensation immédiate de l'interprétation qu'on peut donner de ses médiations discursives, les deux notions de perception évoquées semblent en effet n'avoir rien à faire l'une avec l'autre : leur mise en rapport paraît

une pure question de vocabulaire, et pousser plus loin l'investigation de leur lien donnerait sans doute l'impression que l'on obéit à un pur principe nominaliste. C'est ce que je vais pourtant tenter de faire, en déployant le spectre complet qui va de la "perception immédiate" à sa "représentation discursive" tout en marquant les passages où peut se nouer un véritable dialogue entre les sciences de la cognition et les sciences de la culture, les premières s'attachant aux conditions de l'expérience, psycho-physiques ou phénoménologiques, et les secondes visant ses finalités d'ordre historique, où l'on rencontre en chemin les dimensions à la fois esthétiques et épistémiques de toutes représentations symboliques.

Langage et perception

On s'est largement occupé depuis quelques années des rapports entre perception et langage, selon trois grands axes qui s'inscrivent chacun dans un domaine des recherches cognitives.

1) Le premier axe concerne le problème de la morphogenèse et de la sémiogenèse linguistiques à partir des structures et des processus de la perception visuelle. Trois disciplines, les neurosciences, la psychologie expérimentale et les théories morphodynamiques explorent ce lien génétique entre vision et langage, d'où ont émergé plusieurs hypothèses scientifiques, dont les plus intéressantes, de notre point de vue, sont les différentes formes de néo-localisme ou de spatialisme reposant sur le principe général d'une iconicité des langues naturelles par rapport aux lois de la perception ou à l'ordre de l'apparaître phénoménal¹.

¹On consultera à ce sujet [Petitot, 1992].

2) Le second axe ne concerne plus la genèse du langage mais sa reconnaissance par les locuteurs, soit la perception des formes graphiques et acoustiques des différentes unités qui composent les langues naturelles. Il s'agit là du vaste domaine de "la perception de la parole" qui retient l'attention des ingénieurs informaticiens, intéressés par ce sous-champ de la robotique qui regroupe les travaux sur les machines parlantes, et des psycholinguistes, intéressés notamment par les structures d'apprentissage du langage à partir des processus de discrimination, de "parsage" et de catégorisation des entités linguistiques. Ce problème est celui de la reconnaissance des formes appliquée aux symboles propres aux langues naturelles, envisagés sous leur aspect matériel, bien qu'y soit largement impliquée, on va le constater plus loin, la préappréhension de la dimension conceptuelle du langage².

²On peut lire sur ces questions les trois derniers chapitres de [Halle & Bresnan & Miller, eds, 1983].

3) Le troisième axe — sur lequel je vais m'attarder plus longuement parce qu'il me semble pouvoir faire la synthèse des deux autres et déboucher de la sorte sur le problème — que je veux poser à mon tour, de la "perception des univers du discours", concerne cette fois le vaste champ

de ce que les psychologues, depuis Kosslyn et Shepard, appellent l' "imagerie mentale" et que des linguistes sémioticiens, comme François Rastier récemment, nomment la "perception sémantique" [Rastier, 1991, chap. VIII]. L'hypothèse qui sous-tend ces travaux est bien résumée par Michel Denis dans *Image et cognition*, lorsqu'il définit l' "imagerie" comme "une modalité de la représentation mentale qui a pour caractéristique de conserver l'information perceptive sous une forme qui possède un degré élevé de similitude structurale avec la perception", tout en précisant que du percept à l'image mentale il y a bien sûr transformation de l'information sans que l'abstraction qui en résulte fasse toutefois "perdre à la représentation son isomorphisme structural à l'égard de la perception" [Denis, 1989, p. 9]. François Rastier reprend cette hypothèse, en ajoutant que ces "images mentales, rebaptisées «simulacres multimodaux», sont les corrélats psychologiques des signifiés linguistiques, et que la référenciation s'opère [conséquemment] par appariement entre images mentales et percepts d'objets" [Denis, 1989, p. 207]. On a là une chaîne sémiotique qui va du "signifié linguistique" à la "chose perçue" en passant par l' "image mentale" et le "percept d'ordre sensoriel", nous permettant d'emblée d'embrasser la quasi-totalité du processus sémio-génétique, à l'exclusion toutefois de la dimension plus proprement signifiante des symboles linguistiques, dont n'est retenu ici que le niveau propre aux "représentations lexicales", seules supposées par une telle approche de la perception sémantique.

Ce dernier axe de recherche reprend à l'envers ou à rebours les hypothèses du premier, concernant la sémiogenèse des langues naturelles, dans la mesure où il montre que l'impression référentielle résultant de la reconnaissance du sens des énoncés linguistiques dépend d'un isomorphisme structurel postulé entre l'imagerie, qui sert de support à cette reconnaissance, et la perception sensorielle, qui laisse en héritage à nos représentations mentales les formes les plus prégnantes et les plus saillantes qui la caractérisent. La chaîne est bouclée, qui nous fait passer de l'extrant du discours — la reconnaissance du sens — à son intrant le plus originaire — l'expérience perceptive dans laquelle prennent racine les formes et les entités sémantiques les plus primitives des langues naturelles. La théorie de la perception sémantique rejoint ainsi les hypothèses de base de la morphodynamique et d'une large part de la linguistique cognitive, fondée sur le lien étroit entre Gestalt, percept, image mentale et configuration sémantique — sans oublier, toutefois, les entités morphologiques et les structures syntaxiques qui elles aussi participent du procès, plus large, de la perception discursive, contrairement à ce que laisse entendre la sémantique interprétative d'un François Rastier, qui semble les négliger. Les spécialistes de la perception de la parole affirment en effet que la reconnaissance des unités morpho-phonémiques des

énoncés dépend étroitement de l'interface prosodico-accentuel, située entre l'acoustique et le sémantique : entre la perception sensorielle d'un symbole matériel et la perception mentale d'une représentation idéale. Cette structure intermédiaire, où la syntaxe et le morphème font l'objet d'un acte de perception catégorielle de la part d'un locuteur capable dès lors de discriminer et de découper les unités matérielles du discours selon une courbe prosodique sémantiquement reconnaissable, intervient directement dans la production ou l'actualisation de l'imagerie sous-jacente à la perception sémantique et, partant, dans l'appariement des images mentales aux percepts d'objets ou, plus précisément, à l'expérience perceptive d'une structure événementielle donnée, à laquelle renvoie ultimement tout énoncé.

Il faut faire le lien entre ces différents processus perceptifs en jeu dans l'activité linguistique et montrer l'unité intrinsèque des hypothèses sur lesquelles s'appuie leur élucidation, soit :

1) l'hypothèse du primat de la perception sur les processus conceptuels plus abstraits comme la catégorisation lexicale et morphologique (hypothèse que partagent les sciences expérimentales comme la neuropsychologie et les approches plus spéculatives comme la phénoménologie et la physique du sens) ;

2) l'hypothèse de l'isomorphisme percept-image en psychologie cognitive et de son corrolaire sémiolinguistique qui définit l'acte de référenciation comme l'actualisation d'une image mentale et son appariement aux percepts d'objets (cette dernière notion étant élargie à celle d'événement, qui inclut tout autant les propriétés que les individus) ;

3) l'hypothèse psycho-linguistique d'un niveau mi-idéal mi-matériel, où la perception de la parole articulée met en jeu une image mentale de nature sémantique et le percept d'une forme acoustique donnée, dans un processus où l'ordre linéaire des entités symboliques dans l'énoncé et la nature hiérarchique de leur lien morphosyntaxique font simultanément l'objet d'un double acte de reconnaissance corrélée de la part du sujet énonciateur ou coénonciateur.

Les travaux que je mène sur la *perception discursive* (v. notamment [Ouellet, 1992a, 1992b]) tentent de jeter un pont entre perception sensorielle, perception sémantique et perception de la parole, qui relèvent toutes trois d'une même problématique. Celle-ci concerne le mode d'émergence d'entités perceptibles ou aperceptibles (*Gestalten*, percepts, images mentales, sèmes figuratifs) à partir d'un espace substrat déjà hautement organisé, dont la structure topologico-dynamique contraint notre activité extéroceptive et intéroceptive sur laquelle reposent nos actes d'énonciation et de reconnaissance du discours. C'est sur l'exploration de ce phénomène transverse au langage, à la pensée et à la sensori-motricité

que je fonde l'avenir immédiat des théories sémiotiques, qui doivent saisir le lien entre *l'amont* du sens dans le percept, ancré dans l'organisation gestaltique du perceptible, et son *aval* dans les images mentales, dont le rôle est de remplir les unités de signification verbales d'un contenu intuitif ou figuratif qui servent de support aux représentations symboliques que nous produisons et reconnaissons au sein d'une histoire ou d'une culture données. Cela mène à une quatrième hypothèse, qu'il faut ajouter aux trois précédentes pour bien montrer que les sciences de l'esprit ne relèvent pas des seules sciences de la nature ou des seules sciences formelles, mais qu'elles sont d'emblée, de part en part, sciences de la culture et de l'interprétation, dont elles permettent notamment de renouveler les méthodes. Cette hypothèse pose que :

4) la représentation discursive de notre activité perceptive a une histoire, qu'on peut repérer à travers l'analyse des formes d'expression comme celles en jeu dans la littérature d'une société à une époque donnée — cette analyse permettant de dégager ce que j'appelle des *esthésies*, qui sont aux configurations de la connaissance sensible ce que les *épistémè* de Michel Foucault sont aux formes de la connaissance rationnelle.

Bien sûr, le fossé apparaît large et profond entre cette dernière hypothèse et les trois premières. Leurs visées sont apparemment fort différentes ; j'essaierai tout de même de montrer leur lien, si ténu soit-il, si incongru qu'il semble.

Perception discursive et imagerie mentale

Je voudrais revenir brièvement sur l'une des hypothèses qui fondent l'existence de l'imagerie mentale et, par le fait même, le lien entre l'intéroceptivité et l'extéroceptivité, d'où vient que l'on peut parler de "perception sémantique" à propos de l'appréhension du contenu intuitif ou figuratif de nos concepts ou de nos représentations mentales. Je ne veux pas m'attarder trop longuement sur le débat entre imagistes et propositionnalistes, qui a fait couler beaucoup d'encre chez les cognitivistes depuis plusieurs années³, ni trop me concentrer sur les subtilités du principe de "double codage" auquel la plupart des chercheurs semblent conclure au terme de cette très longue controverse⁴, mais j'aimerais m'arrêter sur l'une des expériences les plus déterminantes qu'ont menée les imagistes pour prouver l'existence d'un substrat spatio-temporel à nos représentations mentales, support que Kosslyn appelle *visual buffer* (traduit ici par "écran visuel") et qui présente toutes les propriétés d'un espace coordonné, capable de figurer analogiquement l'espace-temps que l'on perçoit et où l'on se déplace dans le monde phénoménal (v. [Kosslyn, 1980]). Il s'agit d'une expérience typique de

³Sur ce débat voir [Tye, 1991] et [Finke, 1989].

⁴On doit la première formulation de ce principe à [Paivio, 1979].

“balayage mental” (*mental scanning*) grâce à laquelle les chercheurs ont pu dégager une relation linéaire entre la longueur d’un “transport visuel” dans l’exploration d’une image mémorisée et la durée de ce même transport, corrélation qui atteste l’existence d’une véritable “métrique de l’image”, comme dit Michel Denis [Denis, 1989, p. 71], permettant de dégager les propriétés spatiales, topologico-dynamiques, du support où s’inscrivent les représentations mentales. L’expérience proprement dite consiste :

1) à présenter à des sujets la carte d’une île imaginaire sur laquelle figurent sept détails (un puits, une plage, etc.), disposés de telle sorte que les vingt et une distances qui les séparent soient toutes différentes ;

2) à demander aux sujets de mémoriser cette carte jusqu’à pouvoir la dessiner de mémoire sans erreur quant à la localisation des détails ;

3) à les inviter à former une image mentale de cette carte et à focaliser leur attention sur un des détails qui leur est nommé par l’expérimentateur ;

4) puis, au moment où un autre détail est nommé, à parcourir mentalement l’espace depuis l’élément de départ jusqu’à un nouvel élément et d’appuyer sur un bouton dès sa découverte.

Le résultat montre une corrélation entre la latence de vérification (soit la durée de l’exploration mentale) et la distance spatiale séparant les différents détails. Plus cette distance est grande, plus le temps de réponse est long — ce qui suggère qu’il y a un véritable processus d’exploration d’un support spatial des images, où la distance se traduit en durée, la colocalisation en un authentique balayage spatio-temporel⁵. Ces travaux sont relativement connus et je ne voudrais pas insister outre mesure sur leurs conséquences sinon pour souligner que l’“écran mental” comporte des limites et des contraintes de mise en œuvre parfaitement similaires à celles qui régissent la perception visuelle immédiate ; ainsi la taille apparente de l’objet imaginé aura des incidences sur la durée du balayage mental, tout comme le volume de l’objet perçu détermine partiellement le temps nécessaire à son repérage et à sa reconnaissance. Tous les résultats auxquels ces études aboutissent nous obligent à considérer que nos représentations visuelles ne se contentent pas d’enregistrer le contour apparent des objets perçus, soit leur apparence structurale schématique, mais qu’elles encodent aussi de manière analogique la colocalisation de ces objets dans un espace coordonné.

⁵Cette expérience fut menée par [Kosslyn & Ball & Reister, 1978]. Voir aussi [Denis, 1989, p. 70-74].

Or cette colocalisation, si elle dépend de la structure spatiale “apparente”, et iconiquement déterminée, du support mental des images, relève tout autant du processus même de balayage qui suppose une certaine temporalité, liée à la durée relative de l’acte, et une certaine motilité ou motricité mentale, liée à sa trajectoire et au mouvement

inhérent au tracé exploratoire. Ainsi la *structure* analogique inerte ou "statique" de l'image mentale par rapport au percept sensoriel dépend-elle ultimement d'un *processus* perceptif "dynamique", dont on peut décrire deux types : 1) les processus d'*exploration* mentale, où le "point de vue" se meut imaginativement dans un espace d'images fixes et 2) les processus de *rotation* mentale, qui font subir à l'image certaines transformations, la faisant imaginativement tourner dans l'espace autour d'un "point de vue" relativement stable. Si le premier type de motilité perceptive caractérise l'*acte noétique* du sujet intentionnel, comme dit la phénoménologie, le deuxième caractérise son *contenu noématique*, propre à l'état de choses visé (v. [Husserl, 1950 (1926), chap. III "Noèse et noème", p. 300 et sq.]). On peut, bien sûr, concevoir des processus où les deux types de motricité mentale, exploration et rotation, se conjuguent étroitement, comme c'est le cas le plus souvent dans la perception discursive des espaces sémantiques générés par la description littéraire, dont je vais parler à l'instant.

Nature de la vision, vision de la nature

C'est sans doute dans la description poétique du monde naturel que l'expérience perceptivo-motrice conditionnant l'élaboration discursive d'univers référentiels se manifeste de la manière la plus riche, la plus élégante et la plus évidente à la fois. Toute description suppose un sujet observateur qui ancre l'énonciation dans la perception : le "regard descripteur" (v. [Hamon, 1981] et [Bertrand, 1986]) parcourt un espace que la structure morpho-sémantique de l'énoncé descriptif "donne à voir" et qui dès lors devient le corrélat d'un double acte de vision et de motion imaginaires, dont la structure morpho-dynamique est isomorphe à celle de la sensori-motricité immédiate. La vue que nous avons de notre "champ de vision" mental (du *visual buffer*) et, *a fortiori*, des espaces sémantiques générés par la perception discursive (des formes d'expression et des formes de contenu de nos énoncés en langue naturelle), ne repose pas, comme on l'a constaté, sur la simple reconnaissance de formes ou de *Gestalten* aux contours fixes, stables, appréhendées globalement comme individus ou comme entités. Notre activité intéroceptive relève d'une structure et d'une dynamique intentionnelles complexes où prégnances subjectales (noétiques) et saillances objectales (noématiques) sont conjointement ou successivement "activées" à travers un double *motus*, où l'on reconnaît, d'une part, le "motif" ou la "motivation" qui met en branle la noèse perceptive vers son objet et, d'autre part, la "mouvance" ou la "mobilité" propre au contenu noématique, objet non tant d'une aperception immédiate que d'une découverte consécutive à une exploration et, dès lors, non plus forme individuelle reconnue d'emblée mais effet d'une (co)localisation dans un espace-temps sous-jacent.

⁶Voir "What Is a Concept, That a Person May Grasp It?" in [Jackendoff, 1992, p. 34 et sq.].

⁷Voir "Semantics of Spatial Expressions" in [Jackendoff, 1983, chap. 9., p. 161-187].

Il en découle que l'instance intéroceptive "va vers" l'instance intéroceptible qui "vient à" elle, en un double mouvement — noétique et noématique, exploratoire ou rotatoire, œil balayant l'espace ou chose s'animant sous le regard — dont le support spatio-temporel mental est un "chemin" (*a path*), au sens que Ray Jackendoff donne à cette notion considérée comme "catégorie ontologique" de base⁶. La notion de "chemin" est pour Jackendoff "un constituant conceptuel" ("*a conceptual part of speech*") qui, combiné à celui de "place", permet de rendre compte de tous les phénomènes de (co)localisation dans l'espace et de changement de lieux tels que la langue naturelle peut les décrire⁷. Le chemin est l'expression commune du mouvement et de la vision dans la mesure où il "schématise" en une seule figure la motion de l'œil parcourant l'espace jusqu'à son objet, sinon l'espace même de cet objet, et la mobilité de l'objet vu se déplaçant dans son propre espace, sinon dans le champ de vision de l'œil qui le vise. C'est le "trajet" du sujet à l'objet ou inversement qui, dans tout acte noético-noématique, symbolise le lien entre 1) le "regard" détaché de l'*ego percipio*, dont il s'émancipe pour voler de ses propres ailes sur l'étendue infiniment extensible de son champ de vision, et 2) la "chose regardée" détachée de la *res extensa*, dont elle se libère pour s'envoler dans tous les sens sous le regard qui la suit et la poursuit jusqu'à perte de vue. Une double *Epokhè* permet, dans la perception discursive, de mettre entre parenthèses l'existence égologique du sujet et celle, ontologique, de l'objet — ancrées toutes deux, par la deixis spatio-temporelle, dans un espace-temps stable et identique à lui-même —, grâce à une double émancipation de la noèse perceptive, devenue pur "regard", et du noème de la perception, devenu pur "apparaître", qui relèvent tous deux d'un même *mouvement* d'allées et de venues, d'une même expérience phénoménologique du "trajet", qu'il soit de l'œil ou de la chose vue. Deux citations de Lamartine illustrent à merveille ce double mouvement, exploratoire et rotatoire, où le trajet se dessine tantôt à partir de la noèse (1), tantôt à partir du noème (2) :

(1) "Je m'approche et mon *regard* tombe, à travers l'échancrure de la roche, sur le plus magnifique et le plus étrange horizon [...] : labyrinthe de jardins, de vergers, de palais, de ruisseaux, où l'*œil* se perdait, et ne quittait un enchantement que pour en retrouver un autre".

⁸[Lamartine, 1862, p. 202-203]. Je remercie Larry Marks d'avoir attiré mon attention sur ce texte, dont on trouvera une analyse détaillée dans [Ouellet, 1995b].

(2) "[...] la ville s'étendait à *perte de vue* dans un labyrinthe de jardins en fleur, jetait ses bras immenses çà et là dans la vaste plaine, [...] *semblait se perdre de temps en temps* sous la voûte des arbres, puis *reparaissait plus loin* en larges lacs de maisons, de faubourgs, de villages [...]"⁸.

Le corrélat noématique de la motricité, qu'elle soit pure motion ou support de la vision, comporte trois dimensions au moins : 1) il a son siège dans le "sujet" ou l' "instance" du mouvement (*ce qui bouge*), qu'il affecte directement, 2) il possède une cible dans l' "objet" ou la "destination" du mouvement (*vers quoi ça bouge*), qu'il vise intentionnellement, et 3) il

dessine un "trajet", qui lui confère un "sens", une direction (*le "bouger" en tant que tel*), dont le tracé dans l'espace et dans le temps constitue sa signification, ce "trajet" pouvant avoir son support dans l'objet perçu ou dans le sujet percevant. La "teneur" d'un acte moteur peut donc résider dans l'une ou l'autre de ces trois conditions, qui se trouvent parfois réunies comme dans "Le regard tombait d'abord sur la ville qui [...] s'étendait à perte de vue dans un labyrinthe de jardins en fleurs [...]" [Lamartine, 1862], où Lamartine synthétise en un même énoncé le double point de vue, noétique d'abord, noématique ensuite, sur le trajet qu'emprunte la perception motrice, où sujet et objet de la motion et de la vision conjuguées sont tour à tour affectés, tantôt comme cible, tantôt comme siège de la motricité, et cela grâce à leur trajectoire respective dans l'espace mental représenté.

Chemins d'écarts

La littérature contemporaine, tout comme la romantique, fait une large place à l'activité sensori-motrice qu'incarne dans la langue l'énonciation descriptive. Après Proust, Giono, Gracq et le Nouveau Roman, la littérature descriptive refait aujourd'hui surface chez de jeunes écrivains qui, sur les traces d'un Bonnefoy, d'un Marteau, d'un Trassard, explorent les limites et les profondeurs de la "vision motrice" que la langue et son pouvoir imageant permettent de mettre en scène et en mouvement. Claude Douguin fait partie de ces auteurs récemment révélés, qui consacrent la plus grande part de leur œuvre à cette expérience à la fois perceptive et énonciative de la description⁹. *Écarts*, l'un de ses derniers livres, rassemble des textes qui sont pour la plupart des "paysages", dont on a dit qu'ils sont "la topographie d'un imaginaire"¹⁰. Celui que j'ai choisi pour les besoins de l'analyse s'intitule *À Spasskoïe, peut-être* — mais ce pourrait être n'importe où, en effet. Il s'articule tout entier autour de la figure du "chemin" (appelée aussi "route", "sentier" ou "allée", parfois encore "ligne de pente", "ligne de champ", "sillon" ou "descente"), dont j'ai dit plus haut qu'elle confère schématiquement son "contenu intuitif" à l'acte noético-noématique de la vision motrice, qui est à la base de toute énonciation descriptive¹¹.

C'est par le chemin, littéralement, qu'on entre dans le paysage décrit et qu'on y est conduit, de place en place ; mais c'est aussi, de manière moins littérale, par ce même chemin élevé au rang d'*eidos*, de "forme (typique) de l'apparaître et du sentir"¹², plus que de *topos*, lieu commun de l'expérience immédiate, que nous avons doublement accès au monde sensible et à notre propre sensibilité dans le parcours perceptivo-cognitif que l'énonciation descriptive nous fait faire dans l'état de choses

⁹Cinq livres récents ont paru sur ce thème, tous aux éditions Champ vallon (Seyssel), dans la collection "Recueil" : «La Lumière des Villes» (1990), «Lettres de l'Avent» (1991), «Recours» (1991), «Écarts» (1994) et «La Forêt périlleuse» (1995).

¹⁰Voir le texte de la quatrième de couverture.

¹¹Sont placés en annexe dix extraits de ce texte où l'on trouve des occurrences de la figure du "chemin" (que j'ai soulignées). Les expressions renvoyant au champ de la motricité sont en italiques et celles qui dénotent le champ de la perception visuelle sont en caractères gras. Les références aux citations non contenues dans ces extraits renvoient aux pages du livre, les autres au numéro de l'extrait.

¹²"Ontological category", dirait, quant à lui, Ray Jackendoff.

représenté. Claude Dourguin écrit : "Les uns après les autres s'énumèrent les champs de seigle, de sarrasin" (p. 143), montrant comment les "apparences" défilent sous nos yeux telle une phrase ou une proposition qui trace un véritable parcours, "sillon" ou "ligne de pente" dans notre champ de vision, le long duquel va le "regard" spectateur ou descripteur, "énonçant une à une les parties d'un tout", dit la définition d' "énumérer", qui consiste aussi à "définir un concept par son extension en [nommant successivement] les individus ou les espèces qui en font partie"¹³. On se rappellera la phrase de Giono dans *Fragments d'un paradis*, où l'être de la langue prête sa figure, affine et linéaire — tel le mouvement d'un corps ou le tracé d'une route, qui va de voisinage en voisinage au sens topologique —, à l'apparaître même de la Nature :

"Il semblait que tout était en phrases claires et qu'une plus grande voix que celles jusqu'à présent connues parlait en se servant de formes et de couleurs et d'odeurs, en place de mots" [Giono, 1974 (1948), p. 892].

Le mouvement de la perception discursive trace une ligne idéale qui va de l'apparaître au paraître puis au disparaître dans une séquence qui s'apparente à la chaîne parlée, sinon au flux même de la conscience, où l'attention, dans tout acte intentionnel, voisine avec la rétention, qui la ramène en amont, et la protention, qui la conduit en aval.

¹⁴Pour cette notion, voir [Barbaras, 1994, p. 74]. Affirmant, à la suite de Merleau-Ponty, que "c'est le même qui se meut et qui perçoit", Barbaras définit l' "approche" comme cette "direction qui n'appartient ni à la seule perception ni au seul mouvement [mais] qui exprime clairement [leur] unité originare" (p. 66).

¹⁵Pour éviter d'utiliser le terme ambivalent de "représentation", on peut parler, avec Husserl, de "présentification" imaginative ('*Vergegenwärtigung*'), qui rend présent l'inactuel, par opposition à la "présentation" ('*Präsentation*' ou '*Gegenwärtigung*') de la chose même dans sa présence réelle (voir [Husserl, 1950 (1926), p. 345]).

L'incipit du texte de Claude Dourguin dit bien ce mouvement où, du côté noétique, le sujet percevant "a traversé (des forêts et des bois)", puis a "franchi (un ravin et d'autres)" et a enfin "rejoint (des étendues plantées d'orge)", alors que, du côté noématique, un village a d'abord été "laissé", puis des isbas paraissent "en dérive", du fait que "la route a [enfin] repris sa course (vers les horizons mauves)". La superficie (les "étendues") et les limites (les "horizons") du champ de vision mental où se déploie le paysage décrit, dont les parties vont par la suite "s'énumérer", sont ainsi subordonnées à une *approche*¹⁴, à un mouvement qui, littéralement, va vers par un *travers*, dans la mesure où l'appréhension de tout espace sensoriel et mental, *a fortiori* lorsqu'il fait l'objet d'une présentification verbale¹⁵, est le corrélat d'une exploration linéaire (*scanning* : "balayage") des points saillants qui le constituent. Si le sujet de l'acte noétique, que figure ici le "on", finit en conjonction avec son objet (on a "rejoint les étendues"), c'est qu'il a "traversé" et "franchi" ce qui dans l'ordre de l'apparaître y mène en même temps qu'il lui fait obstacle. C'est là le sens d' "à travers", qui donne accès aux choses *médiatement*, par un effet de retardement, où la force de l'acte est temporairement entravée en même temps que sans cesse guidée vers sa cible, en une sorte d' "effet tunnel" qui n'est pas sans rappeler la manière dont une phrase nous fait passer par le couloir obscur de ses constituants pour nous permettre d'atteindre au bout de sa chaîne à la lumière de son sens. De la même façon, si le corrélat

objectif de l'acte moteur, que figure "la route", peut métaphoriquement "reprendre sa course vers les horizons", avec lesquels il n'entre pas en conjonction mais qu'il peut enfin viser de loin, c'est que les corrélats de l'acte perceptif manifestés par le "village" et les "isbas" auront été d'une part "laissés" et d'autre part mis "en dérive". C'est en une séquence, encore une fois, qui fait apparaître les horizons sur le fond d'une disparition progressive du village et des isbas, que la phrase en même temps que la route organise selon un mouvement prorétentionnel, qui n'avance qu'en repoussant derrière, l'attention que le sujet de l'acte noétique porte aux différents éléments de l'espace imaginaire présentifié. L'effet tunnel reste identique, même si dans le premier énoncé le sujet de la noèse prend l'initiative du "trajet", dont la destination nous est présentée sous l'aspect terminatif (on a *atteint* les étendues), tandis que la seconde phrase met en évidence l'objet ou le contenu noématique des actes perceptifs et sensori-moteurs, soit les isbas en dérive, d'une part, et, d'autre part, la route qui reprend sa course, dont la destination se présente sous l'aspect inchoatif (la route *(re)commence* à courir vers l'horizon).

On a là une double traversée où le sujet et l'objet du mouvement donnent conjointement accès aux choses perçues ou à percevoir selon un parcours qui les fait tour à tour paraître et disparaître, suggérant que l'expérience esthétique en jeu dans l'énonciation descriptive réside moins dans le portrait d'*apparences fixes* correspondant à un monde d'objets empiriques relativement stable que dans les *formes mobiles de l'apparaître* où se manifeste la phénoménalité propre aux mondes décrits ou présentifiés. Le "chemin", qu'il soit parcouru par le sujet ou qu'il conduise ce dernier, est la figure schématique de ces formes de l'apparaître : celles-ci ne se donnent à voir qu'au long ou au bout d'une traversée du "champ de présence" qu'une exploration ou une *approche* seule, éventuellement projetée en une rotation de l'objet exploré, peut transformer en véritable "champ de vision", soit en une "étendue, illimitée, sans autre mémoire que celle des traversées, scarifiée par les routes lointaines", comme le dit si bien le texte de Claude Dourguin (Extr. VI). Toujours un chemin coupe l'espace mental à la recherche de son objet, faisant lever les images le long de son parcours et les laissant retomber derrière dans la poussière, selon un mouvement de balayage qui les fait tour à tour paraître et disparaître, jusqu'à une limite qu'elle vise, appelée "horizon", "lisière", pour montrer qu'elle ne s'atteint pas, reculant sans cesse sous le pas ou devant le regard. S'il y a une "structure d'horizon" qui organise l'expérience perceptive et celle, imaginative, de l'énonciation poétique¹⁶, il faudrait concevoir aussi une "structure de chemin" qui la recoupe et la redouble, en ce sens que l'orée ne se manifeste qu'à la double condition d'une *route* imaginaire qui y mène, perpendiculairement, visant eidétiquement son franchissement, son au-delà, et d'une *ligne de*

¹⁶Voir [Husserl, 1970 (1954), chap. 8, p. 35] : "Que l'horizon fait partie de la structure de l'expérience". Michel Collot [1989] a étendu la notion d'horizon aux présentifications imaginatives comme la poésie, lieu d'expression de l'imagerie mentale prise dans ses dimensions topologiques et dynamiques esthétiquement vécues.

pente ou d'une *ligne de champ*, imaginaire aussi, qui le longe ou l'épouse, parallèlement, donnant à voir et à suivre, eidétiquement encore, son fictif ou illusoire tracé.

Ce qui étonne d'abord, dans le texte de Claude Dourguin, c'est le petit nombre d'occurrences d'événements perceptifs explicitement exprimés : à peine une douzaine dans l'ensemble du texte, dont la moitié se trouve dans nos dix courts extraits, soit dans le voisinage de la figure du "chemin" qui, comme on l'a vu, sert d'attracteur sémantique au champ lexical de la vision autant qu'à celui de la motion. Les événements moteurs, quant à eux, abondent : près d'une cinquantaine d'expressions du mouvement parsèment les seuls passages qui s'organisent autour d'une "route", d'une "allée", d'un "sentier", comme si la modalité sensorielle propre à l'expérience perceptive de l'énonciation descriptive était moins la vision proprement dite que la motricité, support ultime de l'exploration d'un "champ de présence" même fictivement présentifiée comme c'est le cas dans la littérature. On devrait parler d'un "pas" plutôt que d'un "regard" descripteur, le déplacement dans l'espace étant l'une des conditions de son appréhension, du moins dans l'esthésie ou le mode de sensibilité de l'époque contemporaine.

Parmi les quelques événements perceptifs repérables dans le texte, deux, seulement, mettent en présence, dans une structure transitive canonique (SVO), un sujet animé relativement stable, source de l'acte noétique, et un objet inanimé apparemment fixe, qui constitue son contenu noématique : "on aperçoit un bois d'aunes, de peupliers" (Extr. II) et "on aperçoit un grand sorbier" (p. 147). Mais outre le fait *qu'apercevoir* dénote, par opposition à *voir* ou *observer*, par exemple, une vision supposée imparfaite, le plus souvent transitoire ou de courte durée, due notamment à une position désavantageuse du sujet percevant¹⁷, les deux énoncés sont précédés d'une proposition impliquant le mouvement ou son arrêt momentané : "la voiture *s'est engagée dans une descente*, on aperçoit un bois [...]", "De ce coin de salle où l'on se tient maintenant, on aperçoit un [...] sorbier". Toutes les autres expressions visuelles ont pour objet un événement moteur dont le siège réside le plus souvent dans l'objet perçu (et parfois dans la figure du chemin) auquel on attribue une activité motrice ou un mouvement "rotatoire" autonome : "on le voit [le *chemin*], large *sillon* vide, *aller par les champs au-devant de l'horizon*" (Extr. V), "on les reconnaît [les demeures] par les *allées* [...] qui obscurcissent le ciel" (Extr. VIII), "le regard [...] voit la plaine rejoindre le ciel", "on perçoit des *allées et venues*" (p. 145)¹⁸. D'autres expressions comme "le regard se fixe" (Extr. II) et "le regard les suit (les vautours qui *planent*)" (Extr. X), attribuent un prédicat lié au mouvement à l'événement perceptif lui-même, rendu autonome par rapport au sujet de l'acte noétique (le "on")

¹⁷Pour une typologie des actes de perception visuelle lexicalisés, on consultera [Ouellet, 1993a].

¹⁸Un seul événement : "il arrive qu'on aperçoive quelqu'un, allongé dans la fraîcheur [...]" (p. 145), a pour objet un être animé, paradoxalement présenté au repos, dans un arrêt du "mouvement", autour duquel toutefois la scène continue de s'organiser, ne serait-ce pour le nier.

qui laisse s'émanciper la noèse dans son approche exploratoire sinon dans l'appréhension de l'objet visé, lui-même pouvant être en mouvement, rotatoire, comme dans le deuxième exemple cité. Il arrive aussi qu'un œil, organe de la vision, soit prêté à la chose vue, objet d'un mouvement que la figure du "chemin" lui imprime : "les grandes *lignes de champs* repoussent l'horizon sous l'œil énorme du ciel" (Extr. IX). Immédiatement après ce dernier passage, on trouve l'énoncé suivant : "[...] vent, poussière, neige par *rafales* et *tournoiements* en quelques heures font disparaître le paysage où rien ne se distingue plus" (p. 149), dans lequel l'événement perceptif est empêché, du côté noétique, par l'événement moteur qui caractérise son contenu noématique, présentifié ici dans une structure morphosyntaxique privilégiant le factitif (faire disparaître) et le pronominal (se distinguer) dont le double rôle est de mettre en relief le caractère agentif du corrélat noématique et de constituer le sujet de l'acte comme patient, voire comme victime, dans la mesure où la vision (du paysage) est littéralement subie par le sujet percevant comme non-vision ou aveuglement. Toutes ces expressions renvoyant directement ou non au champ sémantique de la vision contribuent à placer l'activité perceptive sous la dépendance de la motion, propre à l'œil bien sûr, comme dans "le regard suit", mais spécifique aussi au champ de vision lui-même ou à l'espace perçu dans lequel le "chemin", l' "allée", la "ligne de champ" jouent un rôle prépondérant.

Cette dépendance est renforcée par le nombre élevé d'événements moteurs dans l'ensemble du texte et dans nos dix extraits, dont je n'analyserai ici, faute de place, que ceux qui ont trait à la figure du "chemin", où se marque le passage de la vision à la motion en une même directionnalité, une même approche, un même "rayon de monde" dirait Merleau-Ponty, sous l'angle changeant duquel le réel n'apparaît jamais que par esquisses successives, sous l'un puis l'autre de ses aspects, comme à travers la vitre avant, arrière ou de côté d'une voiture en marche, justement. On rencontre plus d'une vingtaine d'occurrences de "chemin" dans nos extraits et la quasi-totalité en position de sujet de l'énoncé (d'un verbe de mouvement, toujours, à quelques exceptions près) et d'agent de la scène événementielle décrite. Parmi celles qui s'inscrivent dans une structure transitive, huit ont pour objet un inanimé, appartenant le plus souvent au champ sémantique de l'espace¹⁹ :

- (1) la *route* — *laisser* → le village (Extr. I)
- (2) la *route* — *faire dériver* → les isbas (Extr. I)
- (3) le tourbillon du *chemin* — *ébouillir* → la campagne (Extr. I)
- (4) le *chemin* — *suivre* → la *ligne de pente* (Extr. III)
- (5) la *route* — *scarifier* → l'étendue (Extr. VI)
- (6) la *route* — *faire reculer* → l'horizon (Extr. VII)
- (7) les *allées* — *obscurcir* → le ciel (Extr. VIII)
- (8) les *lignes de champs* — *repousser* → l'horizon (Extr. IX)

¹⁹On trouvera dans [Ouellet, 1992a] et dans [Ouellet & El Zaïm & Bouchard, 1994] un exposé de la méthodologie utilisée ici, que le manque d'espace m'empêche de présenter en détail, sinon pour préciser qu'elle fait appel à la lemmatisation du lexique et à une schématisation de la morphosyntaxe destinées à mettre en évidence, sous forme d'esquisses, les relations prédicatives et primitives qui structurent et dynamisent les formes d'énoncés.

tandis que trois seulement possèdent pour objet un sujet anthropomorphe, source première de l'acte noétique, devenue la "cible", l' "agi" ou le "patient" d'un acte de mouvement, qui possède ainsi son moteur dans le "trajet" plutôt que dans le "sujet" proprement dit²⁰ :

²⁰On remarquera par ailleurs que le sujet de l'acte objectivé dans la structure syntaxique n'apparaît explicitement qu'en (10) — (9) et (11) l'ayant éludé.

(9) un *chemin* — *conduire* → x (Extr. II)

(10) une *allée* — *mener* → le visiteur (Extr. III)

(11) l'autre *allée* — *mener* → x (Extr. IX).

Les autres occurrences où le "chemin" est en position de sujet d'un verbe de mouvement sont présentifiées dans des structures non transitives, qui peuvent prendre la forme de synapsies verbales²¹, comme dans (12) et (14), ou de verbes pronominaux, comme dans (13) et (16) :

²¹De verbes composés, dont l'objet est indétachable du sens même du procès.

(12) la *route* → *reprendre son cours* (Extr. I)

(13) le *chemin* → *s'assagir* (Extr. III)

(14) les *sentiers* → *aller leur mesure* (Extr. IV)

(15) le (*chemin*) → *aller* (par les champs) (Extr. V)

(16) la *route* → *s'étirer* (Extr. VII).

Enfin, le "chemin" est aussi sujet d'un verbe d'état possédant un attribut à valeur spatiale :

(17) le *chemin* → être proche (Extr. X)

de même qu'il est, dans un seul cas, objet d'un acte de mouvement dont le sujet est aussi une occurrence de "chemin" (voir (4)), alors que dans deux autres cas il occupe la place d'un circonstant de lieu qui permet de situer un acte de mouvement :

(18) une *prairie* — *glisser* → jusqu'au *chemin* (Extr. V)

(19) les *troikas* — *voler* → sur le *chemin* (Extr. X).

Des locutions déverbales (des nominalisations) sont aussi associées au "chemin" ou à son corollaire, le mouvement, comme dans :

(20) des *traversées* (Extr. VI)

(21) l'*errance* proche, le *vagabondage* (Extr. VII)

(22) une *pause* du *chemin* (Extr. IX).

— ce à quoi il faut ajouter toutes les expressions renvoyant aux mouvements du sujet noétique en tant que tel (figuré le plus souvent par le "on", fortement désubjectivant par son caractère indéfini), et qui suggèrent toutes la présence corrélatrice d'un chemin ou d'une route : *traverser*, *franchir*, *rejoindre* (Extr. I), *prendre de ce côté* (Extr. II), *quitter* (Extr. VIII). Beaucoup plus nombreuses, toutefois, sont les occurrences de verbes de mouvement dont le sujet est un inanimé, qui renvoie le plus

souvent au champ de la spatialité et des éléments naturels comme dans : les collines *se haussent* (Extr. I); une prairie *dégringole* et *glisse*; le ciel *grandit* (Extr. V) ; la plaine *rejoint* le ciel (Extr. X).

Les données de cette analyse rudimentaire convergent toutes vers une même configuration esthétique, dans laquelle le "chemin" occupe une position centrale, dans la mesure où il synthétise les processus exploratoires et rotatoires, grâce à l'objectivation spatiale 1) du mouvement propre à la noèse perceptive (nécessairement motrice : dirigée vers une cible que son "trajet" réel ou imaginaire permet seul d'atteindre et même de viser) et 2) de la motilité inhérente au corrélat noématique de la perception (nécessairement mouvant : tournant sous tous les angles où la noèse tente eidétiquement de le saisir dans ses différents aspects). Le chemin est cet autre nom du "rayon de monde" que Merleau-Ponty voit indéfiniment tourner dans la sphère du champ de vision, qui est aussi champ de présence où objet et sujet pris dans le même mouvement, exploratoire d'une part, rotatoire de l'autre, se rencontrent et se séparent selon les multiples trajets qui tantôt repoussent l'horizon, tantôt scarifient l'étendue, se haussent, se fixent, dérivent ou s'étirent, selon toutes les possibilités eidétiques que le texte de Claude Dourguin énumère à merveille.

L'abondance des expressions dénotant le mouvement dans un texte descriptif où, paradoxalement, la vision n'est qu'à peine représentée, l'ergatification des expressions spatiales renvoyant au corrélat noématique de la perception et la responsabilité que le noème en acquiert dans l'ensemble des actes sensori-moteurs, bref la subordination de la noèse perceptive au noème de la motricité, dont le "chemin" semble le motif ou la figure exemplaire, donnent à la présentification imaginative et discursive de la Nature une certaine configuration dont on peut dire qu'elle marque un changement esthétique important dans l'histoire récente de la sensibilité poétique. Le monde décrit n'est plus mis à distance par le "regard descripteur" mais parcouru et pénétré de toutes parts selon différentes *approches*, qui ne projettent plus l'objet perçu ou imaginé sur l'écran de la représentation mimétique mais conduisent littéralement le sujet et l'objet l'un vers l'autre, aveuglément, selon divers chemins ou trajets qui se recoupent et qui les ciment, les rapprochent et les repoussent, obéissant aux différents mouvements qui donnent aux actes intentionnels leur force formatrice ou leur énergie motrice, bref leur motivation profonde dans l'épaisseur du corps propre comme dans la chair du monde, animées toutes deux par un même *motus*. L'esthésie contemporaine, révélatrice d'un acheminement commun à la vision et à la motion, qui accomplissent l'une et l'autre une traversée des apparences vers la mouvance propre à l'apparaître corrélatrice aux mouvements du sentir, met en scène et en jeu des processus d'exploration et de rotation mentales où le

²²On trouvera d'autres analyses, qui étayent cette hypothèse, in [Ouellet, 1995a].

trajet prime sur le sujet et l'objet, brouillant ainsi l'image stable qu'on avait d'eux dans l'égologie et l'ontologie classiques, où ils étaient face à face, placés en vis-à-vis, alors que nos présentifications imaginatives, dont témoigne exemplairement la poésie, tendent plutôt à les entrelacer, lacis de chemins et de lignes de pente, d'allées et de descentes, où l'œil ne se reconnaît plus sans le pas qui l'accompagne bien au-delà de la vue²².

Conclusion

Nous sommes parti du problème de la morphogenèse du langage, dont la base réside dans la structure et les processus de l'expérience perceptive, tels que les sciences empiriques (la neurobiologie et la psychologie expérimentale), les sciences formelles (la morphodynamique ou la physique du sens) et les recherches plus spéculatives ou introspectives (comme la phénoménologie de la perception), nous ont permis de les mettre en lumière, confirmant ou étayant certaines hypothèses linguistiques formulées depuis Humbolt et réactualisées aujourd'hui par la grammaire cognitive. Puis nous avons abouti, en fin de parcours, à cette notion d'*esthésie*, où je vois se condenser certaines figures ou schèmes gestaltiques relatifs à la manière dont nous nous représentons symboliquement, notamment à travers des énoncés linguistiques à contenu figuratif, comme c'est le cas de la littérature descriptive, nos expériences perceptives ou sensori-motrices, relayées au niveau de l'imaginaire par les processus d'imagerie mentale, dont l'*exploration* et la *rotation* sont deux des types les plus prégnants. Notre point de départ repose sur les conditions *a priori* de l'expérience symbolique, de nature épistémique ou esthétique, qui plonge invariablement ses racines dans les formes de l'intuition que sont l'espace et le temps en jeu dans toute perception et dans tout mouvement, fondement de notre présence au monde. Notre point d'arrivée, quant à lui, consiste dans la visée ou la finalité *a posteriori* des systèmes symboliques, qui est de mettre en jeu, dans l'univers des représentations mentales mises en discours au sein d'une collectivité socio-historique, les données mêmes de cette expérience sensori-motrice de base pour les explorer et les transformer imaginativement, c'est-à-dire pour les interpréter dans le cadre de l'ensemble des formes symboliques où s'exprime notre connaissance sensible à une période donnée de notre Histoire.

La médiation entre ces deux pôles réside dans le double passage de l'*extéroceptivité* à l'*intéroceptivité* — où l'image mentale hérite, en les transformant, de la dynamique ou de la structure du percept — et de l'*imagerie cognitive* ainsi obtenue aux structures et à la dynamique propres aux *représentations lexicales et morphosyntaxiques* des langues naturelles — telles qu'elles s'organisent à partir et autour d'un sujet

cognitif qui n'est pas seulement énonciateur mais, plus fondamentalement encore, percepteur ou observateur (source et support d'un ensemble d'actes sensori-moteurs). Le texte de Claude Dourguin exprime, par le choix fait de telles ou telles représentations lexicales et morphologiques, un point de vue particulier sur l'expérience perceptive du monde décrit : ce point de vue induit une activité d'imagerie mentale chez le lecteur, qui revit imaginativement cette expérience selon une orientation et un parcours cognitif particuliers au sein de l'espace mental qu'une telle activité intéroceptive génère. Ce lieu mental, comme on l'a vu, possède des propriétés formelles (topologico-dynamiques) similaires à celles du monde phénoménal, mais il est aussi le lieu où peuvent se déplacer et se transformer les données relatives à l'espace et au temps en jeu dans la motilité visuelle ou la vision motrice sensibles au sein de tout processus d'exploration ou de rotation. Ces transformations, rendues possibles par le jeu des ressemblances et des dissemblances entre percepts et images, dénotent quant à elles des changements plus ou moins radicaux dans les formes de la sensibilité d'une société, telles qu'elles s'expriment au sein des systèmes symboliques au rang desquels la littérature joue sans doute un rôle privilégié. Les configurations discursives que prennent ces formes plus ou moins changeantes de la sensibilité sont précisément ce que j'appelle des *esthésies*, qui permettent de repérer, à travers des formes d'expression de la langue naturelle, les formes de contenu de notre expérience sensori-motrice, telle que l'histoire même des formes symboliques dans lesquelles nous vivons nous contraint à la penser et à l'imaginer.

C'est là ouvrir la sémiologie par les deux bouts, soit par le pôle qui la lie à notre présence muette au monde en tant que corps mouvant et percevant, et par cet autre qui l'attache aux mouvances de l'histoire grâce auxquelles on peut parler d'une vie de l'imaginaire et d'une dynamique propre aux formes symboliques qui la prennent en charge. L'étude du phénomène de la perception discursive montre ainsi que les sciences de la cognition et les sciences de la culture — qui relèvent toutes deux des sciences de l'esprit — ont en réalité le même objet, dont la sémiotique s'est effectivement occupé depuis les Stoïciens, la Grammaire spéculative, les Empiristes anglais ou les Sensualistes français, jusqu'à sa relance dans le structuralisme contemporain — soit le lien tantôt serré tantôt ténu entre l'univers complexe des sensations qui nourrit notre activité symbolique et celui des représentations mentales ou sémiotiques que cette activité génère en rapport à la fois avec la structure anthropologique et historique de notre imaginaire et avec la dynamique psycho-physique ou phénoménologique de notre expérience sensori-motrice²⁴.

Université du Québec à Montréal

²⁴Les recherches dont ce texte rend compte sont subventionnées par le CRSHC. Je tiens à remercier Larry Marks, Hervé Bouchard et Adel El Zaïm, chercheurs au sein de l'équipe REPER (REprésentation de la PERception), pour les discussions qui ont mené à la rédaction de cet article.

Bibliographie

BARBARAS (R.)

1994, *La Perception : essai sur le sensible*, Paris, Hatier.

BERTRAND (D.)

1986, *L'Espace et le sens*, Paris, Amsterdam, Hadès-Benamins (Actes sémiotiques).

CHURCHLAND (P.)

1990, *Matter and Consciousness*, Cambridge (MA), MIT Press (nouv. éd.).

COLLOT (M.)

1989, *La Poésie moderne et la structure d'horizon*, Paris, PUF (Écriture).

DENIS (M.)

1989, *Image et cognition*, Paris, PUF (Psychologie d'aujourd'hui).

DOURGUIN (C.)

1994, *Écarts*, Seyssel, Champ vallon (Recueil).

DREYFUS (H.)

1984 (1979), *Intelligence artificielle : mythes et limites*, trad. par R.-M. Vassallo-Villaneau et D. Andler, Paris, Flammarion.

FINKE (R. A.)

1989, *Principles of Mental Imagery*, Cambridge (MA), MIT Press.

GIONO (J.)

1974 (1948), *Fragments d'un paradis*, in *Œuvres romanesques complètes*, t. III, Paris, Gallimard (La Pléiade).

HALLE (M.) & BRESNAN (J.) & MILLER (G. A.), eds.

1983, *Linguistic Theory and Psychological Reality*, Cambridge (MA), MIT Press.

HAMON (P.)

1981, *Introduction à l'analyse du descriptif*, Paris, Hachette.

HUSSERL (E.)

1950 (1926), *Idées directrices pour une phénoménologie*, trad. par P. Ricœur, Paris, Gallimard.

1970 (1954), *Expérience et jugement*, trad. par D. Souche, Paris, PUF (Épiméthée).

JACKENDOFF (R.)

1983, *Semantics and Cognition*, Cambridge (MA), MIT Press.

1992, *Languages of the Mind : Essays on Mental Representation*, Cambridge (MA), MIT Press.

KOSSLYN (S.)

1980, *Image and Mind*, Cambridge (MA), Harvard University Press.

KOSSLYN (S.) & BALL (T. M.) & REISTER (B. J.)

1978, "Visual Images Preserve Metric Spatial Information : Evidence from Studies

of Image Scanning", *Journal of Experimental Psychology : Human Perception and Performance*, n° 4, p. 47-60.

LAMARTINE (A. de)

1862, *Voyage en Orient*, in *Œuvres complètes*, Paris, [sans mention d'éditeur].

MERLEAU-PONTY (M.)

1964, *Le Visible et l'invisible*, Paris, Gallimard.

OUELLET (P.)

1992a, *Voir et savoir : la perception des univers du discours*, Montréal, Balzac, (L'Univers des discours).

1992b, *Signification et sensation*, Presses de l'Université de Limoges (Nouveaux actes sémiotiques).

1992c, "La Dimension cognitive du discours littéraire : perception discursive et imagerie mentale", *TLE*, "Épistémocritique et cognition", n°10, p. 59-68.

1993a, "Quantité et qualité dans la représentation des verbes de «perception»", p. 175-197, in *La Quantité et ses modulations qualitatives*, J. Fontanille ed., Presses de l'Université de Limoges.

1993b, "Le Don des formes : schématisation et actes perceptifs", *Protée*, vol. 21, n°1, p. 15-24.

1995a, "L'Image mue : vision et motion dans le langage poétique", *Protée*, vol. 23, n°1 (sous presse).

1995b, "L'Œil en creux : la perception discursive du paysage dans les *Voyages en Orient* de Lamartine, Nerval et Flaubert", *Eutopias*, Université de Valence (Espagne), à paraître.

OUELLET (P.) & EL ZAIM (A.) & BOUCHARD (H.)

1994, "La Représentation des verbes de perception : le cas de *paraître*", *Praxilingue, Cahiers de praxématique*, "La catégorisation verbale", n°22, p. 135-156.

PAIVIO (A.)

1979, *Imagery and Verbal Processes*, Hillsdale (NJ), Erlbaum.

PETTITOT (J.)

1992, *Physique du Sens*, Paris, Éditions du CNRS.

RASTIER (F.)

1991, *Sémantique et recherches cognitives*, Paris, PUF (Formes sémiotiques).

TYE (M.)

1991, *The Imagery Debate*, Cambridge (MA), MIT Press.

Annexe

Extrait I

On *a traversé* des forêts d'ormes et de chênes, des bois mêlés — trembles, cerisiers sauvages et sapins —, *franchi* un ravin, puis d'autres, *rejoint* des étendues plantées d'orge. Un village *laissé*, quelques isbas *en dérive*, les palissades vieilles autour des chenevières, la *route a repris sa course* vers les horizons mauves. Les uns après les autres s'énumèrent les champs de seigle, de sarrasin, les grandes plaines où *se haussent* quelques collines, de loin en loin des massifs de bouleaux, des buissons ; partout la campagne déserte, vite *ébouffée* par le vent et ce tourbillon de poussière du grand *chemin* (p. 143).

Extrait II

Plus tard — peut-être s'est-on endormi ? — le regard *se fixe*, l'attention s'éveille : la voiture *s'est engagée* dans une *descente*, on *aperçoit* un bois d'aunes, de peupliers ; au carrefour un *chemin* y *conduit* : il faut *prendre de ce côté* (p. 143).

Extrait III

Le *chemin s'assagit*, suit une molle *ligne de pente*, bientôt une *allée* de vieux tilleuls *mène* le visiteur sous la provende lumineuse de ses feuilles (p. 144).

Extrait IV

Au jardin soigneusement sablés, les *sentiers* clairs *vont leur mesure* au long des parterres et des buissons (p. 145).

Extrait V

De l'autre côté de la barrière une prairie *dégringole* la *pente*, *glisse* jusqu'au *chemin*. On *le voit*, large *sillon* vide, *aller* par les champs au-devant de l'horizon, et le ciel, alors, *grandit* (p. 146).

Extrait VI

[...] — l'étendue, illimité, sans autre mémoire que celle des *traversées*, *scarifiée* par les *routes* lointaines (p. 146).

Extrait VII

Là-bas, mais c'est à deux *pas*, la *route s'étire*, *fait reculer* l'horizon sans cesse au milieu des champs de seigle fleuris de bluets, là-bas l'*errance* proche, le *vagabondage* muet [...] (p. 147).

Extrait VIII

Elles [les demeures] *furent quittées* il y a un, cinq ou dix ans, de loin on les reconnaît aux *allées* compactes de sapins qui obscurcissent le ciel (p. 148).

Extrait IX

L'une puis l'autre *allée* de sable de proche en proche bordée de sapins *mènent* à la lisière du domaine. Les grandes *lignes de champs repoussent* l'horizon sous l'œil énorme du ciel : ici, une *pause* du *chemin*, l'étendue un instant domestiquée (p. 149).

Extrait X

Deux vautours *planent* au-dessus des seigles, le regard les *suit*, voit la plaine *rejoindre* le ciel. Derrière soi, la maison *abandonnée à sa dérive* végétative *laisse venir* à elle les songes lointains, cernée par les tentations et les peurs de l'immensité : le *chemin* est proche sur lequel *volent* les troïkas (p. 149).

Les Fondements sémiotiques et philosophiques d'une grammaire morphodynamique

Wolfgang Wildgen

Le dessein de cet article est de donner un résumé bref des présuppositions philosophiques et sémiotiques d'une grammaire morphodynamique. Il est impossible de discuter en détail le problème philosophique du réalisme, les controverses épistémologiques sur le rôle des mathématiques dans les sciences et les essais multiples pour une définition du signe linguistique. J'ai choisi quelques questions fondamentales auxquelles une théorie du langage et une grammaire doivent apporter une réponse. Les réponses seront données sous forme de postulats commentés de façon à indiquer qu'elles ne sont guère définitives. Elles suffisent pourtant à montrer au lecteur la *Weltanschauung* qui motive l'innovation radicale d'une *grammaire morphodynamique*. Pour l'élaboration d'une grammaire prototype de ce genre je renvoie à *l'Introduction à la description morphodynamique des langues*¹, dont la publication est actuellement en préparation.

¹Cette contribution correspond en partie au chapitre 2 de ce livre.

1. Signification et réalité

La question fondamentale d'une sémiotique qui sert à fonder la grammaire morphodynamique consiste à définir la relation sémiotique, c'est-à-dire la relation entre le signe lui-même et une ou plusieurs entités/processus auxquels le signe renvoie, qu'il signifie. La question fondamentale dans toute la science du langage est donc la suivante : quelle est la signification d'une entité linguistique (mot, phrase, texte) ?

Cette question en amène immédiatement deux autres, l'une ontologique, l'autre méthodologique :

a) Est-ce qu'une entité, que nous appelons "signification", existe ? et quel serait alors son statut ontologique ?

b) Comment peut-on accéder à la connaissance, à la description et à la modélisation théorique de cette entité ? Quel est l'apport des recherches en sémantique déjà existantes ?

Les réponses données à ces questions conduiront à une notion plus précise et nouvelle de la *signification* et pourront, dans la mesure où elles sont valables, lui donner plus de *réalité*. Elles devront permettre d'améliorer nos modèles et théories en sémantique et en linguistique théorique. Le concept même de réalité n'est pas fixé dans ce cas ; il s'agira de préciser le type de réalité correspondant au phénomène de la signification.

On peut déjà prévoir que cette question va opposer :

(a) la **réalité scientifique** ; surtout celle établie par les lois fondamentales de la physique et, avec moins de certitude, celles de la chimie, de la physiologie et de la biologie. Celles-ci dépendent elles-mêmes de façon compliquée des méthodologies scientifiques et surtout des langages scientifiques (langages mathématiques et techniques) ;

(b) la **réalité phénoménale** ; celle dont nous nous rendons compte dans la perception normale que nous avons de nous-mêmes et de notre environnement ;

(c) la **réalité organique** de l'homme pour lequel un mot, une phrase ont une signification. Un sous-domaine de cette réalité est constitué par l'esprit de l'homme parlant et elle est en partie sub-consciente (donc elle ne rentre pas dans (a)). L'esprit (*mind*) ne peut pas être définitivement capté par une description scientifique, car il est une partie constituante du sujet, de l'individu humain qui met en œuvre le travail scientifique.

Cette liste de strates ontologiques est préliminaire. On pourrait ajouter d'autres sources de conflits telles que :

(d) la **réalité socio-historique** ;

(e) la **réalité symbolique**.

On peut aussi essayer de réduire les trois strates fondamentales, soit en considérant la réalité scientifique comme une spécialisation de la réalité phénoménale, soit au contraire en voyant dans la réalité phénoménale un type de strate préscientifique pouvant être entièrement remplacé par la réalité scientifique, au fur et à mesure que les sciences progressent. Enfin, on pourrait considérer l'aporie du *mind-body problem* qui apparaît dans (c) comme non pertinente et adopter un monisme à la base de (b), ce qui nous donne un monisme phénoménologique radical, ou bien à la base de (a), ce qui mène à un monisme réductiviste (physicaliste) radical.

Toutes ces positions sont défendues au cours des discussions philosophiques et épistémologiques. Provisoirement, nous partons des

trois types d'ontologie (a), (b), (c) qui ne sont pas indépendants et pour lesquels tous les essais de réduction semblent avoir échoué. Comme notre dessein est de développer une théorie de la signification, nous avons accès à une réalité symbolique que nous espérons pouvoir expliquer ; c'est pourquoi nous ne la présupposons pas. Les trois niveaux de réalité fondamentaux contiennent déjà des éléments du social :

1) Le problème de l'intersubjectivité et de sa constitution dans un monde qui est d'abord phénoménal a été traité par [Husserl, 1931], mais il reste le problème central pour toute approche phénoménologique.

2) Le problème du consensus de la communauté scientifique devient fondamental pour la notion de vérité scientifique qui ne veut plus se servir du critère aristotélicien de l'adéquation du mot à la chose.

3) L'évolution des animaux sociétaires et le développement des sociétés (animales et humaines) montrent que notre base organique est adaptée à la vie sociale de façon évolutive et qu'il y a des bases héritées du social dans l'individu.

Ces arguments ne font que rendre plausibles les trois niveaux de base. La modélisation à laquelle notre analyse arrivera devra corroborer cette décision de base. Il faudra donc se demander en dernier lieu si les aspects symboliques et sociaux pourraient être expliqués et intégrés de façon satisfaisante. En sémantique, cela revient à montrer que l'indépendance du symbolique observé surtout dans le langage littéraire (dans la poésie, par exemple) et dans la conventionnalité des systèmes linguistiques peut être expliquée de façon satisfaisante. Si nous appliquons la méthode de l'explication par des ontologies fondamentales (voir le postulat [1]), c'est que nous pensons y trouver une **généralité** qui, à partir de principes très fondamentaux, peut expliquer (générer) une multitude de phénomènes assez compliqués. Notre choix ontologique est donc guidé par une vision méthodologique qui espère pouvoir trouver des lois fondamentales en sémantique. La sémantique appliquant cette stratégie est appelée "réaliste", car nous essayons de la fonder dans un réel accessible par nos méthodes scientifiques et de l'expliquer à l'aide des principes et des lois déjà trouvés dans le domaine des sciences naturelles. Ce programme correspond à la stratégie que Thom appelle la "réduction de l'arbitraire". Nous pouvons dès lors formuler le postulat de base de toute sémiotique réaliste.

Postulat réaliste [1] —————

Nous considérons que le phénomène du signe renvoie à une réalité (en dehors du 'corps' du signe) qui nous est connue et accessible sous trois aspects fondamentaux (non réductibles les uns aux autres) :

- a) notre propre corps,*
- b) le monde phénoménal,*
- c) le monde, tel que les sciences le font apparaître (surtout par leurs lois générales).*

2. Quelques postulats pour le fondement d'une sémiotique morphodynamique

Une sémantique *réaliste* n'est pas empiriste dans la mesure où un type d'observations facilement accessibles serait choisi comme point de départ. Elle n'est pas *behavioriste*. Nous devons faire une série de présuppositions qui ne sont ni immédiatement évidentes, ni primitives au sens d'un empirisme. La seule chose que nous pouvons faire pour éclaircir ces présuppositions sera de les mettre au jour de façon systématique. Les postulats et corollaires de ce travail ont pour but de présenter les présuppositions ; ils ne peuvent guère les prouver.

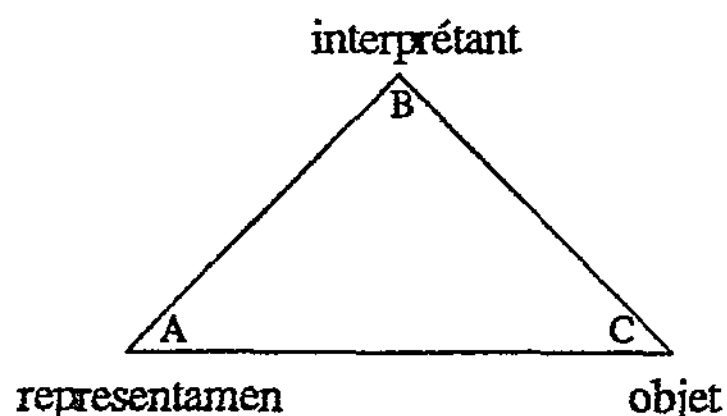
Postulat de Peirce [2]

L'unité fondamentale du signe est donnée par une triade A, B, C et les relations entre les éléments de cette triade: (A, B), (B, C) et (A, C).

— La forme du signe (le *representamen* selon Peirce) est normalement un type de processus (de production verbale ou d'écriture) et son résultat, qui n'est stable que dans le cas de l'écriture.

— La forme interne (la représentation) du signifié (*l'interprétant* selon Peirce²), est normalement un processus interne, mental. Peirce admet en plus des interprétants abstraits qui sont eux-mêmes des signes — ce qui donne lieu à une itération infinie dans la position de l'interprétant.

— La forme externe (*l'objet* ou le processus qui sont externes au système cognitif). Il s'agit surtout du monde phénoménal accessible à la personne qui parle, écrit etc., par exemple du contexte d'un locuteur, du monde dans lequel il agit, qu'il perçoit et dont il forme une représentation interne à l'aide de la mémoire et de l'imagination.



Graphe de la triade de Peirce

Cette triade de Peirce peut être théorisée sous des aspects partiels :

— si l'on se focalise sur le système porteur, locuteur/auditeur, on peut mettre en relief la dyade A-B et considérer C comme l'arrière-plan. C'est la position de F. de Saussure ;

²Peirce inscrit sa sémiotique dans le cadre de la pragmatique qu'il a développée. L'interprétant peut être lui-même un signe, qui de nouveau se décompose de la façon décrite, et ainsi de suite. Nous acceptons la possibilité d'une telle boucle réflexive, mais nous croyons que ces boucles sont strictement limitées et qu'elles ont une source finale qui a un caractère cognitif. Cette source (ce but final) peut être appelée la représentation cognitive d'un mot, d'une phrase, d'un texte.

— le réductionniste (*behavioriste*) considère *C* comme "stimulus", *A* comme "réaction" et *B* comme "black box" ;

— on peut choisir *B* comme centre organisateur d'un modèle et dire que *A* n'est que la réalisation d'une représentation interne (par la bouche ou la main) ou le résultat de cette réalisation et que *C* n'est qu'une conséquence externe d'une réalisation motrice dans l'acte de pointer ou dans la mise en acte d'une intention cognitive (position cognitiviste) ;

— enfin, on peut choisir un point de vue abstrait qui postule des structures abstraites (logico-mathématiques) sous-jacentes à *A*, *B* et *C* (position platoniste).

Chacun de ces choix mène à des techniques de modélisation, à des conceptualisations et même à des stratégies de recherche empirique différentes. Le postulat [2] rejette toutes ces solutions comme étant inadéquates ; elles essayent de nier le caractère foncièrement *hétérogène* de la triade sémiotique et de la réduire à l'un de ses sous-aspects ou à un aspect abstrait, extérieur à chacun des trois aspects (voir la position platoniste).

Postulat génétique [3]³

La structure hétérogène et irréductible de la triade peircéenne ne peut être ramenée à une base unique que par un renvoi théorique à la genèse (phylogenèse, genèse socio-historique, ontogenèse) du système des signes.

Le postulat [3] ne contredit pas le postulat [2], si celui-ci vaut pour une triade déjà constituée (une sémiotique synchronique, déployée).

La première conséquence du postulat génétique est tout d'abord de compliquer la théorie. En effet, une explication des phénomènes sémantiques nécessite, en dernier lieu, un retour en arrière dans le temps — ce qui ne facilite guère l'étude empirique des fondements de la sémantique et la rend, à la limite, impossible. On peut assimiler les étapes suivantes dans la régression théorique au cadre d'une sémiotique génétique.

a) Dans l'ontogenèse, la triade *A-B-C* se déploie chez le nouveau-né dans un réseau quasi morphogénétique. Le nouveau-né doit d'abord savoir faire la distinction entre le corps (par exemple de la mère) *C* et le signe (indépendant du corps de son énonciateur) *A* et enfin la "signification" *B*, cette entité d'abord impalpable mais que son esprit est apte à reconstituer (voir les travaux de [Mottron, 1983, 1987]).

b) Dès que l'enfant a *compris* le fonctionnement triadique fondamental, il a accès au système linguistique qui est porté par ses parents. Il est donc confronté à un système qui a connu une morphogenèse

³Le postulat génétique a ses sources surtout dans l'épistémologie du 17^e et du 18^e siècles. Déjà Condillac proposait une construction "génétique" quand il donnait à sa statue successivement des facultés perceptives, motrices et cognitives. C'est surtout Piaget qui, au vingtième siècle, s'est fait l'avocat d'une épistémologie génétique, voir [Piaget, 1972].

historique qui, à son tour, exploite la faculté de langage héritée. Nous savons qu'il y a un seuil de complexité marqué entre les capacités sémiotiques des primates d'aujourd'hui et l'homme, quoique les substrats neuronaux soient comparables.

c) Enfin, les capacités sémiotiques des animaux (mammifères, oiseaux) renvoient à des comportements informatifs chez les animaux primitifs. Elles peuvent être causées par des adaptations à l'environnement, surtout à la niche écologique de l'animal. René Thom a poursuivi cette ligne de pensée jusqu'à la prégnance⁴ de la lumière.

⁴Cf. par ex. [Thom, 1988, p. 15-52] et [Thom, 1990, p. 53-65]
Cette position théorique a été élaborée vers une théorie de la prégnance.

Ce type d'explication de l'hétérogénéité du sémiotique par une régression génétique a la même limite, le même noyau que toutes les sciences, c'est-à-dire qu'il renvoie à la création du monde (le Big Bang dans une des cosmologies actuelles). De ce fait, il n'y a pas d'explication *sui generis* pour le monde des signes ; ceux-ci font plutôt partie du mouvement général qui a mené au monde où nous vivons aujourd'hui. C'est justement cette régression à la singularité créatrice qui montre que les positions de Peirce et de Thom se complètent de façon immédiate. Il n'y a pas de réduction possible du sémiotique, sinon celle qui renvoie au début de tout ce qui existe.

Une conséquence du postulat génétique peut être formulée comme corollaire :

Corollaire aux postulats [2] et [3]

Comme la triade sémiotique émane de façon naturelle d'un processus génétique général, elle en partage les lois générales. C'est-à-dire qu'une étude théorique en sémiotique devra trouver des lois analogues aux lois de la physique (lois du mouvement, lois de la conservation de l'énergie, etc.).

Cela ouvre un domaine de recherche que Thom appelle *sémiophysique* :

"Mais qu'est donc l'objet de cette Sémiophysique ? La Sémiophysique concerne d'abord la recherche des formes signifiantes ; elle vise à constituer une théorie générale de l'intelligibilité" [Thom, 1988, p. 12].

Cette conception fut élaborée par Jean Petitot. Elle "cherche à jeter les bases d'une «physique mathématique» des structures et de leurs effets symboliques dans le champ de la représentation", en utilisant la théorie des catastrophes comme "premier exemple d'*analysis situs* structurale". Elle cherche également à fonder le structuralisme même, non pas "sur l'évidence fallacieuse, imposée par le positivisme logique, que les mathématiques sont un langage", mais sur "*une topologie dynamique de places et de connexions*" [Petitot, 1985b, p. 19].

La conséquence méthodologique de ce corollaire fait que, au lieu de chercher des *règles* (donc des normes du comportement social), il faut

chercher des lois fondamentales qui puissent régir la genèse du langage et des comportements langagiers dans tous les stades génétiques. On peut conclure que :

a) les lois cherchées doivent être valables pour tous les stades de la genèse. Elles garantissent ainsi la continuité et la stabilité (relative) du langage ;

b) les mêmes lois opèrent dans toute la triade sémiotique (voir postulat [2]), car elles n'appartiennent pas spécifiquement à un domaine ontologique.

Comme ces lois ne sont ni des lois socio-historiques (qui sont elles-mêmes un produit de la genèse), ni des normes imposées par le consensus social, elles opèrent de façon locale. C'est dire que si certaines conditions générales sont remplies, le développement pour les instances locales suit les lois générales. Ainsi la loi darwinienne du *mieux adapté* s'applique indépendamment dans les situations spécifiques et forme à long terme un réseau de genres animaux. Les conséquences macro-morphologiques sont donc à distinguer des lois qui mènent à la formation de ces structures. Cette observation vaut pour la sémiotique, la sémantique et aussi pour la grammaire (voir [Wildgen, à paraître]). On peut distinguer différents types de lois :

— les lois de l'équilibre dynamique, qui ont comme seul paramètre temporel celui de la stabilisation rapide dans un attracteur ou de la déstabilisation dans une singularité (catastrophe). Ces systèmes *oublient* pour ainsi dire tout ce qui mène à l'équilibre et ne conservent que l'information sur la configuration de l'équilibre et à la limite du chemin qui y mène⁵.

— Les lois de la micro-évolution d'un nouvel état d'ordre (surtout dans le contexte d'une fluctuation stochastique). Prigogine et l'Ecole de Bruxelles ont trouvé des systèmes d'autocatalyse qui peuvent décrire la formation d'un état ordonné à partir d'une distribution stochastique et de certaines fonctions globales qui sélectionnent un nombre d'états favorisés.

— Les lois coopératives de la "synergétique" de Haken, qui expliquent l'asservissement d'un système avec des variations complexes sous un ou plusieurs modes. Contrairement aux systèmes de Prigogine, ces modes font déjà partie du système ; l'ordre n'apparaît pas vraiment.

René Thom a critiqué la conception de Prigogine sur le fait qu'elle engendre l'ordre du bruit ; le hasard devient quasiment une force d'ordre ; ce qui semble une *contradictio in adjecto*. Si Prigogine trouve ses meilleurs exemples en chimie (la réaction dite *de Bruxelles*), et Haken dans la physique du laser, cela montre déjà le caractère plutôt

⁵La classification des chemins dans les catastrophes élémentaires va déjà au delà de celles-ci et appartient à une analyse des bifurcations.

métaphorique de leurs modélisations en biologie et dans les sciences humaines. On peut dire en fait que, dans tous les systèmes qui sont hétérogènes (à plusieurs niveaux) et qui sont plutôt caractérisés par des morphologies héritées, et non par des morphologies spontanées, on trouve des phénomènes qui sont au delà de la capacité de modélisation des systèmes dynamiques que nous venons d'énumérer. Derrière toutes les variantes de systèmes dynamiques, on trouve pourtant un principe valable pour tous les systèmes non chaotiques : la possibilité de réduire un système très complexe (avec beaucoup de degrés de liberté, de constituants, de connexions) à un système central. En dynamique qualitative, le théorème du *splitting*, qui élimine tous les composants stables d'un système dynamique multi-dimensionnel, ainsi que le *center manifold theorem*, en dynamique stochastique, sont les bases mathématiques d'une modélisation simple dans un domaine dont la complexité n'est même pas connue. Ces théorèmes permettent une modélisation simple, car ils autorisent une réduction radicale de la complexité du problème.

Le corollaire aux postulats [2] et [3] spécifie aussi le rôle des mathématiques dans la théorisation. Il s'agit d'utiliser les conceptualisations contenues dans les meilleures théories trouvées dans les sciences (naturelles). La théorie des catastrophes doit donc sa pertinence au simple fait que la théorie des systèmes dynamiques — qui sert à formuler la plupart des lois de la dynamique (classique) et qui se montre très efficace dans les recherches récentes en physique et en chimie (biologie) — connaît un noyau fondamental : les systèmes dynamiques autonomes. Pour ces systèmes, on peut, dans le cas très simple de l'existence d'une fonction potentiel, classer les points décisifs, les singularités en nombre fini. La théorie des catastrophes est donc d'une part le résultat d'un choix épistémologique (théorie des systèmes dynamiques — systèmes autonomes) et d'autre part d'une décision méthodologique (choix d'un cas très simple et fondamental).

3. Quelques postulats pour le fondement d'une grammaire morphodynamique

L'idée d'appliquer de façon très générale, et dans un cadre théorique moins strict, les résultats de la classification des systèmes dynamiques simples est due à René Thom qui a appelé le champ de ces applications *théorie des catastrophes*. En sémiotique, Thom utilise la réduction drastique d'un système dynamique naturel à ses invariants topologiques, les singularités de Thom et Mather, comme modèle de la sémiosis. Son postulat fondamental est le suivant :

Postulat de Thom [4]

La complexité des processus réels (voir le 'postulat réaliste') peut être dramatiquement réduite par une considération des singularités à évolution stable de ces processus. Cette classification constitue la dynamique profonde des phrases dans toute langue naturelle⁶.

⁶Les effets translocaux existent et sont même très importants pour l'efficacité du langage.

Pour rendre le postulat plausible, il faut considérer de plus près la triade sémiotique : A-B-C (voir le postulat de Peirce) avec l'intention de trouver un invariant qui puisse relier de façon stable les trois termes de la triade :

- A : la forme du signe (la catégorie langagière),
- B : la forme du signifié (la catégorie cognitive),
- C : la forme du référent (la catégorie ontologique).

Si nous considérons une opération de transformation (un groupe de transformations) qui nous donne les formes prototypiques pour chaque domaine, par exemple les formes prototypiques des phrases simples pour A, les classifications percepto-mnésiques pour B ou encore les grands types d'objets (processus) tels que la science nous les fait voir, nous pouvons nous demander s'il est possible de trouver des invariants dans les ensembles prototypes A*, B*, C*.

Le problème fondamental d'une sémantique mathématique consiste donc à trouver des applications pouvant conserver un noyau commun très pauvre et *ignorer* les détails. Une solution très générale est apportée par la topologie différentielle, qui permet la classification d'un système dynamique (donné par une équation différentielle) en éliminant les constituants stables (les éléments quadratiques) et en minimisant le système de façon à obtenir un système irréductible. La relation appelée "difféomorphisme" permet d'oublier la métrique des échelles et de les traiter comme des rubans en caoutchouc. Une conception simplifiée (qui élimine le *feed back*) pourrait partir de C (le monde ambiant) et considérer des pièces (régions autour d'un point choisi) de C. En appliquant le postulat de Thom, nous choisissons les régions autour d'une instabilité (d'une catastrophe) dans C. Dans le cas général, cette instabilité peut être très compliquée. Nous devons donc trouver une convention de sélection pour choisir les régions les plus générales, les plus fréquentes, les plus répétitives. Il s'agit des régions qui sont *structurellement stables*. Ces régions restent identiques même dans le cas de déformations de paramètres, de variations de taille et de changements de forme.

Pour être plus exact, nous ajoutons un corollaire au postulat de Thom :

Corollaire (au postulat de Thom)

Les régions de C considérées sont des voisinages d'une instabilité ponctuelle qui sont structurellement stables. Le système dynamique est un système gradient, c'est-à-dire que ses attracteurs (régimes stables) sont atteints par le chemin le plus court (sans fluctuations).

Par ce corollaire, le choix des régions de C est restreint à des *points* catastrophiques et leur environnement (leur déploiement) structurellement stable (avec une dynamique de gradients). Les attracteurs plus compliqués sont exclus. Cette restriction permet d'aboutir à une classification finie (non continue).

Pour arriver à une classification finie (qui est nécessaire pour pouvoir projeter la classification dans C sur celle dans A (les signes langagiers)), il faut introduire un dernier postulat que Thom a formulé et qui est motivé par les travaux de Tesnière.

Postulat de Tesnière-Thom [5]

La classification dans C sélectionne les régions qui ont :

- *ou bien un paramètre de base (corang 1) et un maximum de trois régions stables en conflit en un point unique du déploiement,*
- *ou bien deux paramètres de base (corang 2) et un maximum de quatre régions stables en conflit en un point unique du déploiement.*

Le postulat de Tesnière renvoie au fait que, dans les langues, on ne trouve normalement que trois actants (centraux) qui, avec un noyau verbal, constituent la phrase simple. La diathèse causative peut en ajouter un quatrième (voir [Tesnière, 1988 (1959), p. 109 s. et p. 260 s.]). L'explication de Thom renvoie à la règle des phases de Gibbs, donc à une loi valable dans C . Cette règle affirme qu'il existe une relation stable entre le nombre de composants du système Co , le nombre de degrés de liberté Li (de paramètres de contrôle qu'on fait varier indépendamment) et le nombre de régions (phases) en conflit Ph :

$$\bullet \text{ } Ph + Li = Co + 2.$$

Il manque une loi comparable pour B (le cerveau ; le cortex, les systèmes sensoriels et moteurs). Une telle loi pourrait être trouvée pour les cartes de la perception catégorielle (voir [Petitot, 1985a]). Nous allons considérer maintenant le domaine cognitif comme déterminé par C et A , soit un système transitoire qui relie C et A . Dans ce cas, le postulat de Tesnière-Thom ouvre la voie à la recherche des invariants de la relation sémiotique entre :

- C : régions de l'espace-temps (monde),
- A : phrases simples (structures actantielles).

Le niveau cognitif s'intercale entre les deux d'une façon qui demande à être élucidée⁷.

Nous allons montrer les conséquences des postulats énoncés en utilisant un exemple très simple, souvent utilisé par René Thom, la phrase : "Le chat mange la souris".

⁷Le corollaire au postulat de Thom est problématique pour le domaine B (cerveau). Certes [Zeeman, 1965] a présenté une description de la neurodynamique où il utilise des dynamiques gradients (la pensée est un attracteur ponctuel de l'activité neuronale), mais la physiologie du cerveau suggère plutôt une dynamique cyclique, une dynamique avec forte diffusion et peut-être même chaotique. Il est pourtant possible qu'une macro-dynamique existe qui remplisse approximativement les conditions des postulats 4 et 5. Il faudra alors se demander où réside le noyau de l'invariance, si par exemple le langage est constitutif pour la classification cognitive d'un certain niveau de complexité.

Nous pouvons distinguer les trois niveaux :

- A* : la phrase noyau : "Le chat mange la souris",
- B* : le schéma percepto-mnémonique de la capture : capture/prédation,
- C* : une expérience zoologique consistant à présenter une souris à un chat affamé et un protocole de l'événement typique.

On trouve dans ce cas les invariants suivants :

1. nombre maximal d'entités (*A* : groupes d'animaux ; *B* : entités mémorisées ; *C* : entités de l'expérience zoologique) = 2 ;
2. asymétries en faveur d'une entité *A* : sujet ; *B* : agent ; *C* : prédateur (le chat) ;
3. existence d'une instabilité dynamique (à évolution stable) :
A : le verbe (d'action, par exemple "manger")
B : l'image conceptuelle de l'action,
C : l'accident dans l'espace-temps à haute énergie.

Dans l'application du postulat de Tesnière-Thom aux structures langagières, nous pouvons distinguer plusieurs niveaux que nous placerons sur une échelle allant d'une position forte à une position superfaible. Dans les trois cas, des sous-ensembles de *A*, *B*, *C*, que nous appelons prototypiques, sont considérés. Un sous-ensemble prototypique est centré et définit un bassin autour de ce centre. Une région de *A* est représentée par son prototype *A** et surtout par le centre du prototype *a**. Dans ce qui suit nous ne traiterons que la version faible et superfaible.

4. Trois versions du lien entre ontologie, représentation cognitive et structure linguistique

a) Version forte

Soient *a**, *b**, *c** les centres des prototypes *A**, *B**, *C**, il existe des applications (difféomorphes et structurellement stables) entre *a**, *b** et *c**, de sorte que leur invariant soit l'un des schémas déduit de l'application du postulat de Tesnière-Thom.

b) Version faible

Au lieu de considérer directement les centres *a**, *b** et *c** dont la modélisation et le contrôle expérimental sont difficiles, nous considérons des représentations schématiques des centres prototypiques *a**, *b**, *c**. Comme *b**, le prototype neuronal ne nous est accessible (pour le moment) que par l'intermédiaire de *a** ou de *c**, nous ne considérons que les représentations de *a** et *c**.

Pour la version faible, nous exigerons que les représentations a^* et c^* puissent être traduites l'une dans l'autre de façon explicite.

Pour la plus grande partie des exemples, la version faible est trop directe. Il faut des filtres supplémentaires car tout élément du scénario dans c^* n'apparaît pas nécessairement dans a^* — et a^* contient des constituants obligatoires qui ne renvoient guère au schéma de c^* .

c) Version superfaible

Nous supposons deux composantes supplémentaires :

(a) un filtre qui peut éliminer des éléments du scénario (c^*) dans la réalisation langagière. Ce filtre est réalisé surtout par les techniques langagières d'incorporation d'un type d'actant dans le lexème verbal et de contextualisation ; l'actant peut être deviné par l'auditeur ou bien par l'intermédiaire d'une régularité (l'ellipse) ou par un effet du contexte.

(b) Un enrichissement par "attribution" (voir la psychologie "attributionnelle" développée par [Heider, 1958]). Les schématisations de c^* sont enrichies :

— par des interprétations des variables de contrôle et de base. Ainsi on peut passer d'une interprétation localiste à une interprétation possessive, qualitative ou actantielle ;

— par des interprétations hétérogènes. Un processus dans l'espace peut avoir un effet qualitatif. Ainsi, le chat qui mange la souris passe d'un état de faim à un état de satiété ;

— par des macro-scénarios et des réseaux génériques (genre-espèce, tout-partie) ;

— par des processus métaphoriques.

5. Conclusion

Les grammaires traditionnelles portaient de la forme du signe et des règles qu'on pouvait directement en déduire ; les grammaires structuralistes cherchaient les lois générales de *ce système où tout se tient* sans référence explicite à des entités/processus extra-linguistiques. La grammaire morphodynamique est un défi à la philosophie structuraliste étant donné qu'elle cherche explicitement à relier, à renvoyer les formes du signe à une réalité extra-linguistique (cognitive, psycho-physique, physique). Elle utilise la dynamique qualitative pour former des "précatégories" sur lesquelles les catégories proprement linguistiques peuvent être à leur tour fondées.

Pour la grammaire morphodynamique trois phénomènes sont primordiaux.

a) L'analyse part toujours de continuums segmentés par des instabilités. Dans les premières étapes de cette segmentation d'un continuum, la théorie des catastrophes sert de classification générale.

b) Elle utilise le matériel des langues spécifiques, mais recherche surtout les principes d'une grammaire générale. Les analyses spécifiques servent à illustrer ces principes. Une étude comparative et historique devrait ensuite compléter notre modèle, mais elle reste encore à faire⁸.

Ainsi la grammaire morphodynamique sert à démontrer que l'approche morphodynamique est féconde et qu'elle rend possible une théorie intégrale et complète du langage. Les grammaires spécifiques (par exemple du français, de l'allemand) pourront être écrites à partir des principes développés dans cet article (voir [Wildgen, à paraître]).

⁸La même méthode peut être appliquée à la phonétique/phonologie [voir Petitot, 1985a]. Comme elle catégorise la substance phonique, elle n'est sémiotique que secondairement (par la faculté de rendre possible une certaine "quantité" de distinctions lexicales).

*Sprach- und Literaturwissenschaften
Universität Bremen*

Bibliographie

HAKEN (H.)

1983, *Synergetics : An Introduction*, 3^e éd., Berlin, Springer.

HEIDER (F.)

1958, *The Psychology of Interpersonal Relations*, New York, Wiley.

HUSSERL (E.)

1931, *Méditations cartésiennes*, Paris, Cohn.

MOTTRON (L.)

1983, *Contraintes communes à l'acquisition, la théorisation et la pathologie de la deixis*, thèse de doctorat, Paris-Sorbonne.1987, *Morphogenese und Bedeutungsstruktur* (traduit par W. Wildgen), p. 233-243, in WILDGEN (W.) & MOTTRON (L.).

PETITOT (J.)

1985a, *Les Catastrophes de la parole : de R. Jakobson à R. Thom*, Paris, Maloine.1985b, *Morphogenèse du Sens*, Paris, PUF.

PIAGET (J.)

1972, *L'Épistémologie génétique*, Paris, PUF.

PRIGOGINE (I.)

1980, *Physique, Temps et Devenir*, Paris, Masson.

TESNIERE (L.)

1988 (1959), *Éléments de syntaxe structurale*, Paris, Klincksieck.

THOM (R.)

1988, *Esquisse d'une Sémiophysique : physique aristotélicienne et théorie des catastrophes*, Paris, InterÉditions.1990, *Apologie du Logos*, Paris, Hachette.

WILDGEN (W.)

1994, *Process, Image and Meaning : A Realistic Model of the Meanings of Sentences and Narrative Texts*, Amsterdam, Benjamins.à paraître, 1995, *Introduction à la description morphodynamique des langues*, Aarhus, Semiotica et Analytica.

WILDGEN (W.) & MOTTRON (L.)

1987, *Dynamische Sprachtheorie : Sprachbeschreibung und Spracherklärung nach den Prinzipien der Selbstorganisation und der Morphogenese*, Bochum, Studienverlag Brockmeyer.

ZEEMAN (E. C.)

1965, "Topology of the Brain", *Mathematics and Computer Science in Biology and Medicine* (Medical Research Council), p. 240-256.

Approche morphodynamique de deux régimes fondamentaux du langage

David Piotrowski

Si la question des propriétés essentielles (définitoires) du langage ne fait pas l'objet d'un large consensus au sein de la communauté linguistique, certains *régimes* que manifestent les langues naturelles se trouvent néanmoins reconnus, et avec insistance, comme fondamentaux dans de nombreuses perspectives théoriques, actuelles ou plus anciennes.

Parmi ces régimes, qui contribuent à circonscrire une région de phénomènes empiriques, on peut recenser les trois suivants : la propriété *d'intégration* qui concerne la nature holistique des constructions langagières, la *dualité* du signe linguistique où les deux faces de l'expression et du contenu se déterminent en une unité synthétique, et le différentiel *d'acceptabilité* qui renvoie au fait qu'il existe un impossible en langue. On nommera ces trois régimes les "fondamentaux" de la linguistique.

Les "fondamentaux" suscitent chacun de nombreuses polémiques. De surcroît, lorsqu'ils sont admis comme tels, leur expression n'est pas toujours limpide, et, plus grave, leur thématization s'avère des plus variable : ils peuvent occuper des positions théoriques fort hétérogènes et, au plan de la méthodologie descriptive, remplir des fonctions plus ou moins centrales.

Chacun des fondamentaux véhiculant son lot de controverses, on imagine combien la tentative de leur traitement conjoint au sein d'un système unitaire peut s'avérer périlleuse. Car pour ce faire, et pour chacun d'eux, des décisions théoriques et épistémologiques des plus délicates sont à prendre sans que rien ne garantisse jamais leur justesse. En quelque sorte, c'est élever les problèmes à la puissance trois. Mais inversement, si l'on accepte un principe de cohésion interne en langue, on dispose d'un critère d'investigation rigoureux : on conviendra que les qualifications particulières des fondamentaux qui les établissent en un système organique cohérent sont rationnellement motivées.

Dans ce sens, nous porterons notre attention sur le différentiel d'acceptabilité et la nature duale du signe linguistique. Après avoir rappelé

succinctement à travers quelques citations (dont il serait aisé de multiplier le nombre) les principales conceptions de ces deux régimes ainsi que les difficultés qu'elles soulèvent, nous proposerons, en nous inscrivant dans une perspective morphodynamique, un modèle rendant compte de leurs articulations internes et mutuelles au sein d'un système dynamique unitaire. Ce modèle pourra alors contribuer au débat sur la détermination théorique, fonctionnelle et épistémologique des "fondamentaux" de la linguistique.

Le différentiel d'acceptabilité

Le différentiel d'acceptabilité occupe une place centrale dans la problématique linguistique contemporaine. Un parcours même sommaire de quelques ouvrages ou revues de linguistique suffit pour le constater. Très grossièrement, d'abord, le différentiel apparaît principalement à deux titres. Au plan d'une épistémologie popérienne, il constitue un critère de réfutation des analyses et des constructions théoriques : les propriétés assignées à telle ou telle unité ne sont valides que pour autant qu'elles n'induisent pas de calcul aboutissant à des configurations inacceptables. Parallèlement, au plan de la pratique descriptive, le différentiel opère comme indicateur : les valeurs d'acceptabilité produites par les procédures variationnelles signalent la présence de qualités fonctionnelles spécifiques et contribuent à orienter les analyses. On rappellera aussi que le différentiel n'est pas sans dimension historique : il fonde, on le sait, l'activité grammaticale [Milner, 1989].

Mais si le fait du différentiel semble par trop avéré pour être dénié, s'il s'agit de toute évidence d'une donnée massive, apparaissant concrètement dans l'exercice du langage et s'imposant à l'observation, la question de son statut linguistique reste ouverte. Sur ce point, deux perspectives majeures s'affrontent. La première défend une conception normative, tandis que pour la seconde le différentiel est constitutif de la réalité langagière.

Dans la conception normative, le jugement d'acceptabilité n'est pas associé aux contraintes intrinsèques d'un système linguistique mais exprime des échelles diverses : sociale, esthétique ou autre, dont les divisions sont projetées sur les données langagières. La légitimité en langue relèverait alors d'une axiologie.

Suivant le point de vue adverse, l'acceptabilité d'un énoncé est une caractéristique propre à l'ordre langagier ; elle participe, en relation avec son pôle converse, de la réalité linguistique. Plus précisément, une langue pourra être définie comme le "*répondant objectif*" d'un système de répartitions différentielles [Milner, 1989, p. 59].

Plutôt que de reprendre le débat frontal entre ces deux points de vue, revenons sur ses origines et essayons de le contourner ou de le dépasser. Comme l'analyse J.- C. Milner, la conception de la norme prend appui sur une disjonction entre la possibilité en langue et la possibilité matérielle :

“l'usage par un sujet d'une forme jugée impossible en langue est toujours *matériellement* possible : rien de physique ni de moral n'empêche Martine de proférer *ne servent de rien*, rien de physique ni de moral n'empêche Philaminte de proférer *ne servent pas de rien*” [Milner, 1989, p. 83].

Conséquemment, puisque les énoncés impossibles en langue sont passibles d'une existence concrète, le principe de leur prohibition est naturellement porté à prendre la figure de la norme :

“la distinction entre le *quid facti* et le *quid juris* ne désigne rien d'autre que la disjonction entre le possible matériel et le possible juridique, [et] la notion de norme n'a d'autre justification de départ que de capter ce phénomène caractéristique” [Milner, 1989, p. 83].

Pour rendre compte de la disjonction entre ce qui est légitime et ce qui est attestable une conception systémique du différentiel aura pour sa part recours au schéma de *la règle* qui formule explicitement ses articulations internes : “la notion de règle (...) instaure une démarcation entre ce qui suit la règle et ne la suit pas (...)”, elle permet d'expliquer “de manière détaillée” les constructions inacceptables, et se conforme au “fait que les deux versants du différentiel puissent se trouver réalisés dans l'observation : la notion de règle (...) ne dit pas que ce qui ne suit pas la règle soit irréalisable” [Milner, 1989, p. 87].

Mais si les points de vue de la norme et de la règle assignent des statuts distincts au différentiel, ils en partagent la même analyse. En particulier, pour ces deux optiques, le différentiel se formule sous une forme prédicative :

“[la proposition empirique en linguistique] a la forme d'un jugement d'attribution dont le sujet est une donnée de langue et dont le prédicat est du type «se rencontrer/ne pas se rencontrer», «se dire/ne pas se dire», etc.” [Milner, 1989, p. 51].

Or, pour autant que la répartition différentielle relève pleinement de l'ordre linguistique, son expression prédicative implique l'articulation d'un double niveau d'objectivité qui assure à l'ensemble des énoncés observables, en deçà de leur qualification dans le système de la langue, une forme de réalité linguistique *matérielle*. Car si les énoncés impossibles en langue étaient effectivement dénués de toute forme d'existence linguistique, il n'y aurait aucun sens à leur appliquer le prédicat différentiel et ce dernier ne serait jamais qu'une illusion, une formule sans corrélat objectif.

Or, comme le souligne J.- C. Milner, ce double niveau d'objectivité qui confère à l'objet "langage" un statut singulier n'est pas sans occasionner quelques complications épistémologiques :

"la grammaire se propose de décrire des propriétés ; par définition, elle les tient pour objectives, mais il se trouve que ces propriétés censément objectives concourent à définir une répartition du possible et de l'impossible qui ne se confond pas entièrement avec l'attesté et l'inattesté. La grammaire est donc amenée nécessairement à supposer une objectivité de langue qui soit distincte de l'objectivité généralement définie pour les objets" [Milner, 1989, p. 89].

De plus, un certain nombre d'observations et d'appréciations semblent aller à l'encontre de la conception prédicative du différentiel qui suppose d'assigner aux configurations matériellement possibles une sorte d'existence linguistique indépendante du système des contraintes en langue.

Ainsi, on constatera que, si l'on fait abstraction des perturbations liées à la "performance" et si l'on se situe à un niveau homogène du fonctionnement langagier, la dimension matérielle entretient des rapports de réalisation très étroits avec la dimension proprement linguistique : "(...) à l'intérieur d'un système de répartition [du différentiel] unique ou même simplement cohérent, ne sont possibles matériellement que les formes possibles linguistiquement", aussi "le vocabulaire de la loi naturelle se révèle plus approprié" [Milner, 1989, p. 122 et p. 123]. Corrélativement, une fois atteint un certain niveau de "dégradation" des formations phrastiques, il semble plus pertinent, comme le propose R. Martin, de parler de "suites" que d'énoncés ou de phrases [Martin, ed., 1978, p. 9].

Sans que l'on puisse remettre en question la latitude de construire des enchaînements incorrects, on reconnaîtra donc que les configurations matérielles produites dans l'exercice du langage sont naturellement réalisées en adéquation aux distributions du possible en langue ; ce qui témoigne d'une forme de connexion entre la dimension matérielle et le possible en langue, d'une sorte d'articulation entre la dimension matérielle et la législation linguistique.

Aussi, tout porte à refuser aux configurations matérielles le statut d'entité linguistique lorsqu'elles sont dégagées des conditions de structure dont elles s'investissent et qui déterminent l'objectivité linguistique, i.e. l'existence en langue. Le pôle de l'inacceptabilité absolue apparaît donc comme le lieu de l'inexistence en langue : un énoncé totalement inacceptable est nié dans *toute* sa réalité linguistique, c'est une enveloppe matérielle dépourvue de qualification en langue.

Dès lors qu'il n'est plus question d'allouer à la dimension matérielle *seule* un quelconque statut linguistique, la conception du jugement différentiel comme forme prédicative — conception qui, répétons-le, impose d'attribuer à la dimension matérielle une forme autonome d'existence linguistique — ne peut plus être maintenue.

Ce point de vue, s'il permet de dépasser les conceptions de la norme et de la règle ainsi que les difficultés épistémologiques qu'elles entraînent, reste néanmoins hautement problématique. Il requiert, pour prétendre à quelque validité, l'élucidation de deux points singulièrement obscurs, à savoir, d'une part, le principe de cette connexion "débrayable" entre la dimension matérielle et la législation linguistique, et, d'autre part, la forme et la fonction linguistique du différentiel d'acceptabilité.

La structure du signe

Les difficultés relatives au premier point présentent quelque parenté avec le problème que soulève le second des "fondamentaux", à savoir le rapport de consubstantialité qu'entretiennent les deux faces du signe linguistique. Dans les deux cas il s'agit de comprendre le mode et la possibilité d'une fusion de deux ordres hétérogènes et primitivement extérieurs l'un à l'autre.

Peu de linguistes sans doute contesteront, tout du moins dans ses grandes lignes, la nature duale du signe linguistique, ou, plus généralement l'articulation des phénomènes linguistiques sur les deux plans de l'expression et du contenu.

Ainsi, outre, bien entendu, Saussure, et parmi de nombreux autres, on pourra citer Hjelmslev :

"Expression et contenu sont solidaires et se présupposent nécessairement l'un l'autre. Une expression n'est expression que parce qu'elle est l'expression d'un contenu, et un contenu n'est contenu que parce qu'il est contenu d'une expression. Aussi est-il impossible, à moins qu'on les isole artificiellement, qu'il existe un contenu sans expression ou une expression sans contenu" [Hjelmslev, 1968, p. 66-67].

Et parmi les linguistes contemporains, Hagège :

"la propriété cardinale du signe est la même qui fait tout le mystère des langues en tant qu'organismes investissant la substance phonique par l'intention de signifier, ou faisant, de la matérialité des sons, surgir le sens : signifiant et signifié sont absolument indissociables et ne peuvent d'aucune manière être conçus séparément. (...) Ainsi, pour un même signe, le signifiant, quelle que soit sa forme, et le signifié, quel que soit son domaine, sont, par l'effet des structures de la langue qui les contient, deux faces constitutivement solidaires d'une même réalité" [Hagège, 1985, p. 131].

Mais si le fait empirique semble clair et "(...) aussi élémentaire que les résumés des manuels d'écoliers" [Hagège, 1985, p. 131], son explication reste des plus problématique. On connaît d'ailleurs les termes du problème que pose la dualité du signe : soit les deux faces, le signifiant et le signifié, sont constituées préalablement à leur fusion, et en ce cas leur consubstantialité est foncièrement inexplicable et contingente, soit leur

fusion procède d'un processus de constitution réciproque : le signifié préside à l'élaboration du signifiant et réciproquement, de telle manière que "[telle] tranche acoustique n'existerait pas sans l'idée correspondante et *vice versa*" [Benveniste, 1966, p. 54]. Mais dans ce cas, priorité doit être donnée à l'une des deux faces. En effet, il semble pour le moins difficile de concevoir un processus où chaque terme commande la formation de celui qui lui est apparié sans qu'aucun ne soit donné préalablement.

Par ailleurs, les descriptions du principe de cette fusion constituante entre matières sonore et sémantique tiennent plus de la métaphore et de l'axiomatique que de l'analyse. Par exemple, Benveniste :

"Il y a entre [le signifiant et le signifié] (...) symbiose si étroite que le concept «bœuf» est comme l'âme de l'image acoustique *böf*" [Benveniste, 1966, p. 51].

"Le signifiant et le signifié, la représentation mentale et l'image acoustique, sont (...) les deux faces d'une même notion et se composent ensemble comme l'incorporant et l'incorporé" [Benveniste, 1966, p. 52].

Hjelmslev, quant à lui, axiomatise cette jonction problématique sous la forme d'une relation de dépendance réciproque qu'il fixe au sommet de l'architecture linguistique. Enfin, dans le cadre d'une approche cognitiviste, Langacker, qui situe parmi les relations fondamentales la *symbolisation*, "where a correspondence is established between a semantic structure and a phonological structure" et déclare renouer avec "the spirit of classic Saussurean diagram" [Langacker, 1987, p. 73-74 et p. 11], ne retient ni le principe d'un signifié spécifiquement linguistique, ni la forme d'une authentique consubstantialité des deux faces du signe. Les grandeurs de l'expression et du contenu sont constituées indépendamment l'une de l'autre et leur connexion présente la forme d'un appel mutuel par déclenchement d'une procédure intériorisée : Langacker définit d'abord la structure *d'unité* comme "a structure that a speaker has mastered thoroughly (...), he can employ it in a largely automatic fashion (...). Psychologists would speak of a "habit", or say that "automatization" has occurred" [Langacker, 1987, p. 57], et analysent l'articulation forme/concept comme une structure *d'unité* :

"the semantic structure automatically calls the phonological structure to mind, and conversely, since the symbolic relation between them has unit status" [Langacker, 1987, p. 58].

Si le principe d'une connexion étroite entre les deux faces du signe semble bien admis par tous, le mode de leur constituance mutuelle *interne au signe* reste donc inexploré. Mais bien qu'il fasse l'objet d'affirmations très assurées, le principe d'une consubstantialité radicale entre signifiant et signifié reste discutable, et, tout particulièrement, on ne peut ignorer les difficultés conceptuelles qu'il soulève. De surcroît, si certains faits

empiriques attestent son bien-fondé, par exemple les mots tabous, en revanche d'autres faits tout aussi certains incitent à modérer la vision d'un signifiant et d'un signifié mutuellement incorporés. Ainsi, il semble effectivement difficile de ne pas préférer le mot pour convoquer l'idée, mais des constructions phonologiques satisfaisant les règles d'une langue et ne véhiculant aucun contenu sont tout à fait réalisables.

Partant, une conception "faible" de la structure du signe linguistique sera sans doute plus appropriée. Et tout particulièrement une conception qui évite la problématique symétrie mentionnée plus haut, et que les faits empiriques contredisent.

On choisira donc d'établir le pôle phonologique en position d'antériorité dans l'ordre de la constitution du signe et, reprenant les conceptions de Langacker sur le rapport procédural entre grandeurs phonologiques et sémantiques — conceptions en écho à différents développements contemporains de la linguistique — nous examinerons comment ce schéma primitif du rapport son/sens peut s'articuler avec le mode du différentiel d'acceptabilité pour produire un signifié spécifiquement linguistique dont l'objectivité sera intrinsèquement corrélée à son signifiant.

Eléments d'une architecture dynamique

Avant d'esquisser la forme d'une architecture susceptible de rassembler les deux fondamentaux et d'explicitier leurs régimes internes, rappelons sommairement les principales articulations du modèle morphodynamique. Nous nous restreindrons aux seuls éléments nécessaires à notre investigation, et renvoyons pour une présentation exhaustive et détaillée aux travaux de R. Thom et de J. Petitot sur lesquels nous prenons appui.

Dans sa version *élémentaire*, la théorie morphodynamique étudie les champs $\sigma : W \rightarrow F$ où W est un espace de contrôle et F un espace de systèmes dynamiques définis par des fonctions potentiel $f_w : M \rightarrow R$ (M étant une variété différentiable compacte et R la droite numérique réelle). La donnée d'une topologie et d'un groupe d'automorphismes sur F permet d'y introduire les notions de stabilité et de type qualitatif. La géométrie des stabilisés d'une dynamique instable $f \in F$ s'exprime alors par un "déploiement universel" : un champ contrôlé par un espace externe W "universel" et une catégorisation de W suivant une stratification de frontières K dont les points correspondent aux éléments instables f_w de F .

Le modèle morphodynamique permet fondamentalement de comprendre l'articulation d'un espace substrat "externe" W contrôlant un espace de formes où sont définis des états "internes" qualitatifs (c'est-à-dire des valeurs relatives aux attracteurs des dynamiques, celles-ci n'étant

considérées qu'en fonction de leur type qualitatif) qui se manifestent par une "trace" (un système de frontières) sur *W* et y instaurent, sur le mode de rapports d'opposition, des identités relationnelles structurales.

Une modélisation morphodynamique suppose donc de spécifier les espaces respectivement interne et externe. Par exemple, dans le cas de la théorie phonologique, *W* sera un espace de *stimuli* de nature acoustique "(...) jouant le rôle d'entrée pour une «boîte noire» perceptive *S*. Dans cette «boîte noire» inobservable, un processus dynamique global *X* (...) définit des états internes *A*, *B*, *C*, ... (des «images acoustiques») et se trouve contrôlé par *W*" [Petitot, 1985b, p. 105].

Or, pour une approche morphodynamique du signifié linguistique, aucun espace externe n'est *a priori* disponible : l'espace du sens n'est pas une donnée tangible, mais une hypothèse de conceptualisation. Aucun paramètre de contrôle n'est d'emblée susceptible de tenir le rôle d'entrée d'un champ de dynamiques et d'instancier en retour un réseau différentiateur exprimant l'articulation des stabilisés de l'espace interne et distribuant des identités structurales.

Toutefois, si l'on admet l'existence d'un espace général de sens, et si l'on accepte le principe d'une dissymétrie entre les pôles de l'expression et du contenu, on disposera avec le signifiant d'une sorte d'amorce, d'une *impulsion*, localisée dans l'espace sémantique et qui va y déclencher, non pas l'activation d'une valeur sémantique bien définie, i.e. d'un signifié, mais une propagation d'activations sémantiques.

L'idée qu'une forme de l'expression définit un point d'entrée dans un espace dynamique de sens est partagée aujourd'hui par différents courants théoriques. Par exemple, pour Langacker "the entity designated by a symbolic unit can be thought of as a point of access to a network (...). A lexical unit provides a conventionally determined *point of access* to knowledge systems" [Langacker, 1987, p. 163-164] et "a complex predicate (...) is more accurately viewed as a set of routines, which are interrelated in various ways facilitating their coactivation" [Langacker, 1987, p. 162]. Sous un angle différent, A. Culioli dénonce cette "illusion que les mots désignent un sens, (...) que les mots pointent un sens, sont des pointeurs sémantiques" [Culioli, 1990, p. 83] et poursuit :

"(...) il y a foisonnement, c'est-à-dire que vous avez tout un ensemble de propriétés qui s'organisent les unes par rapport aux autres, qui sont physiques, culturelles, anthropologiques, et qui font qu'en fin de compte un terme ne renvoie pas à *un* sens, mais renvoie (...) à un *domaine notionnel*, c'est-à-dire à tout un ensemble de virtualités" [Culioli, 1990, p. 86].

Aussi, l'espace du sens peut être envisagé comme un espace de contrôle *restreint* : les grandeurs de la substance de contenu associées aux signifiants (paramètres de contrôle) ne remplissent pas l'espace du sens mais y localisent différents accès. L'espace du sens n'est donc praticable

comme espace de contrôle qu'en certains lieux où se déclenchent des procédures d'activation et de diffusion plus ou moins généralisées.

Soulignons que cette conception n'engage aucune caractérisation linguistique : les associations entre formes de l'expression et les activations sémantiques relèvent d'un mécanisme d'ordre psycho-cognitif, et les valeurs du contenu ainsi activées ne présentent aucune délimitation précise.

L'espace sémantique étant considéré dans sa forme discrétisée comme l'espace externe d'un champ de dynamiques («la boîte noire» de "perception" du sens), on s'interrogera maintenant sur la fonction des dynamiques qu'il commande et sur les modalités d'établissement d'un authentique signifié linguistique au regard des contraintes du différentiel d'acceptabilité.

Rappelons que, fondamentalement, la théorie morphodynamique est un modèle de la catégorisation d'un substrat par un réseau de seuils différentiateurs, "trace des instabilités et des conflits des états internes [que le substrat] contrôle" [Petitot, 1985b, p. 106]. Le système de frontières instancié réalise les rapports d'oppositions suivant lesquels s'établissent des identités structurales : des identités de position entretenant des rapports de détermination réciproque.

La compréhension structurale saussurienne des valeurs sémantiques que schématise une approche morphodynamique suggère alors d'interpréter l'inacceptabilité en langue comme l'effacement des frontières qui instituent les identités, i.e. comme la disparition des régimes structurants du langage. En effet, on a vu que le pôle négatif du différentiel correspond à un néant de langue : une construction violant la légalité linguistique est dépourvue d'existence en langue. Partant il est logique d'exprimer cet anéantissement par l'annihilation des régimes aux fondements de l'objectivité en langue.

Par ailleurs, et pour autant qu'on lui alloue une réalité linguistique — et il semble difficile de ne pas le faire — on sait que le différentiel ne peut fonctionner comme prédicat puisqu'il est amené à qualifier des configurations sans corrélat objectif en langue. Aussi, il doit nécessairement opérer à un niveau supérieur des réalisations du système linguistique — tout en y restant attaché pour conserver son statut. En d'autres termes, le différentiel d'acceptabilité doit porter sur des grandeurs indépendantes de la légalité linguistique et se connecter d'une certaine façon à cette dernière.

Or on admettra bien que le différentiel d'acceptabilité est l'expression de la légalité linguistique : le jugement différentiel formule en effet le fait même de l'existence ou de l'inexistence en langue. Et comme, en tant que réalité linguistique, il est censé manifester sa propre fonction dans le système de la langue, il est donc logique de lui conférer le rôle de *principe*

de formation des identités linguistiques. C'est du moins une hypothèse *naturelle* qu'il conviendra d'évaluer au regard des autres paramètres de notre investigation.

La configuration qui progressivement se dessine est la suivante : d'une part, les signifiants, constitués de façon autonome, sont appariés à des entrées dans l'espace du sens mais ne véhiculent pas de contenu de nature linguistique, et d'autre part, le différentiel porte sur des entités "matérielles" indépendantes de l'ordre linguistique et contribue à les y inscrire. Cette configuration se prête alors naturellement aux identifications suivantes : les signifiants sont les formes matérielles et le différentiel qui porte sur ces dernières instaure une démarcation oppositive dans l'espace sémantique auquel les signifiants donnent accès. L'espace du sens corrélé aux entrées se trouve ainsi géométriquement façonné et des régions de contenu établies en qualité de signifié. Avant d'avancer dans la caractérisation de ce schéma, illustrons-le déjà avec un exemple simple et quelque peu métaphorique.

Considérons une champ restreint de signification autour du verbe *couper*. Ce champ outre l'entrée relative au signifiant *couper* est accessible par les points d'accès que sont les signifiants des lexèmes synonymes de *couper*, par exemple : *enlever*, *ôter*, *diviser*, *scinder*...

Ces différents lexèmes peuvent entrer dans un certain nombre de constructions phrastiques et se trouveront catégorisés au regard d'un différentiel d'acceptabilité. Par exemple, si l'on convoque le schéma phrastique *V en trois (morceaux, parties...)*, le champ de signification sera catégorisé en deux zones, l'une rassemblant les entrées *couper*, *diviser* et *scinder* et l'autre *enlever* et *ôter*. Alors que le contenu de *couper*, i.e. la substance de contenu associée à l'entrée *couper*, est préalablement indéterminé (les activations de *couper* vont diffuser sur l'ensemble du champ sémantique et produire une valeur indifférenciée), la réalisation de *couper* dans une énoncé construit sur le schéma *V en trois (morceaux, parties...)* va produire, par l'instance de catégorisation du différentiel d'acceptabilité, une frontière déterminant une opposition de sens entre les deux zones de contenu : *couper* ne se confond plus avec *ôter* et *enlever*, et un signifié encore vague regroupant les substances que dominent les trois entrées *couper*, *scinder* et *diviser* est ainsi produit.

Il faudra donner maintenant une description moins simpliste et plus précise d'un tel système de régulation, et tout particulièrement caractériser le principe et le fonctionnement de la différenciation par jugement d'acceptabilité.

Ce fonctionnement est pour l'essentiel le suivant. Soient *A* et *B* deux lexèmes. Si la construction *AB* est jugée acceptable et *AB'* inacceptable, le régime du différentiel d'acceptabilité, d'une part, détermine l'émergence

d'une opposition potentielle entre B et B' lorsque AB est produit (B' restant alors virtuel), et d'autre part, contrevient au principe de cette opposition lorsque AB' est produit, annihilant alors toute existence linguistique.

En d'autres termes, les clauses " AB est acceptable" et " AB' est inacceptable" doivent être schématisées dans une double fonction : d'une part, en tant que *repère* d'un régime de différenciation repère situant le lieu d'une opposition actualisable, et, d'autre part, en tant qu'instruction déterminant le caractère acceptable ou inacceptable de l'activation d'un des signifiants, c'est-à-dire, dans le premier cas, l'actualisation effective d'un région de signification en opposition à une région adverse, ou au contraire, dans le second cas, la suspension de tout processus de différenciation.

Pour conclure

En choisissant de conférer au différentiel d'acceptabilité le statut de régime structurant, il semble que l'on puisse lever les difficultés conceptuelles qu'occasionne sa formulation prédicative et appréhender plus clairement le phénomène de consubstantialité des deux faces du signe. Tout particulièrement il est possible d'éclairer les rapports du "possible matériel" au "possible linguistique" sans faire appel à un problématique double niveau d'objectivité : les jugements différentiels portent sans incohérence sur la dimension autonome des signifiants car leur fonction n'est pas *constatative* mais *productive* ; ils déterminent dans l'espace du sens auquel les signifiants donnent accès des régimes d'opposition qui président à l'élaboration des signifiés. On explique alors aussi la raison d'une consubstantialité unilatérale entre les deux faces du signe : la production des signifiés passant nécessairement par le niveau du signifiant qui, seul, permet l'expression des clauses d'acceptabilité, la convocation du signifié s'avère de ce fait tributaire de la profération du signifiant.

Pour autant qu'elle présente quelque validité, la configuration morphodynamique, bien trop sommaire, ici proposée pour une détermination du différentiel d'acceptabilité et de la constituance du signe reste toutefois encore à parfaire. Tout particulièrement, une exigence de praticabilité pose la question d'un modèle dynamique effectif, et la question du troisième des "fondamentaux", l'intégration, est encore à traiter.

CNRS — INaLF
(27 rue Damesme, 75013 Paris)

Bibliographie

BENVENISTE (E.)

1966, *Problèmes de linguistique générale*, t. 1, Paris, Gallimard.

CULIOLI (A.)

1990, *Pour une linguistique de l'énonciation*, Paris, Ophrys.

HAGEGE (C.)

1985, *L'Homme de paroles*, Paris, Fayard.

HJELMSLEV (L.)

1968, *Prolégomènes à une théorie du langage*, Paris, Éditions de Minuit.

LANGACKER (R.W.)

1987, *Foundations of Cognitive Grammar*, vol. 1, Stanford University Press.

MARTIN (R.), ed.

1978, *La Notion de recevabilité en linguistique*, Paris, Klincksieck.

MILNER (J.- C.)

1989, *Introduction à une science du langage*, Paris, Le Seuil.

PETTITOT (J.)

1985a, *Les Catastrophes de la parole : de R. Jakobson à R. Thom*, Paris, Maloine.

1985b, *Morphogenèse du Sens*, Paris, PUF.

Dynamical Constituency

An Epistemological Analysis

Jean Petitot

Introduction

One of the main problems of Cognitive Grammar (CG) — and of its European precursors such as Lucien Tesnière, Hans Jacob Seiler or Bernard Pottier — is to find adequate *mathematical* tools for describing the topological and dynamical information — in fact the *morphological* information — which is provided by perception, iconically encoded in cognitive image-schemas and processed by the semantics of natural languages.

Algebraic and logical formalisms are dramatically not convenient for such a task. Indeed, as far as they are discrete and symbolic, they cannot do justice to the continuous, spatio-temporal, morphological and Gestaltic aspects of the image-schemas. What we need for exploring what David Touretsky called the “iconic way” is another sort of formalism, of a topological, dynamical and physical nature.

Following René Thom’s propositions in catastrophe theory and morphodynamics, I have tried to develop such a research program in semio-linguistics since 1975. It fits very well with the new trends in cognitive linguistics. We agree for example with Ray Jackendoff that there does exist a conceptual structure, that is a deep cognitive level of representation at which perception, action and language become compatible. This conceptual structure transforms the physical world into a projected world — a morphologically structured phenomenological world — endowed with a qualitative ontology.

The consequence of the fact that natural language is rooted in this morphological basis is for instance, in what regards syntactic structures, *the localist hypothesis*. It says that the possible relations of position (static and dynamic) between local domains in space-time play the role of archetypes — of basic schemes — for syntactic structures. I have explained it in details in my book *Morphogenèse du Sens*¹. It is the same idea that we find in Ray Jackendoff when he says, following Gruber, that :

¹[Petitot, 1985].
Concerning the localist hypothesis, see also [Petitot, 1979, 1989b, 1989d]

"In any semantic field of [EVENTS] and [STATES] the principal event-, state-, path-, and place- functions are a subset of those used for the analysis of spatial location and motion" [Jackendoff, 1983].

We agree also with Leonard Talmy about the fact that *closed* grammatical classes (like prepositions, modal auxiliaries, grammatical relations, etc.) specify *semantic* contents of a very special and deep sort. As it is very well explained in "Relation of Grammar to Cognition" [Talmy, 1978] and in "How Language Structures Space" [Talmy, 1983], there exists a complex positional-topological information and a sophisticated morphological schematization which are specified e.g. by the closed class of prepositions :

"Grammatically specified structuring appears to be similar, in certain of its characteristics and function, to the structuring in other cognitive domains, notably that of visual perception".

So, we need for cognitive linguistics a good "cognitive topology" in Lakoff's sense, and for this a good mathematical theory of the figurative schemes which organize semantic structures is required. This "topology" must be in fact also a dynamics. It must also include more geometry than the pure mathematical topology, for instance differentiable structures, stratified decompositions of spaces, convexity properties, transversality properties, or even metrical structures. All this belongs to what I call "*cognitive morphodynamics*".

The question is : does there exist some adequate and non-trivial mathematics for cognitive morphodynamics ? The response is : yes ! These mathematics do exist. They are deep and sophisticated : qualitative theory of dynamical systems, bifurcation theory, singularity theory, etc. In fact, they are the mathematics used in connectionist modeling. But they have already been used in the morphodynamical approach long before the connectionist wave.

Now, the development of dynamical models for cognitive processing raises fundamental issues, some of which have already been tackled by connectionist models implementing dynamical devices. One of the most difficult challenges is the following : have dynamical models implemented in connectionist networks the capacity for adequately modeling in a purely dynamical way *constituency* (constituent structures) and *compositionality* which, classically, are modeled by symbolic ones ? At the linguistic level, the difficulty is for instance to model grammatical relations and semantic roles (in the sense of case grammars). In a nutshell, it can be formulated in the following way : if terms of sentences are modeled by attractors of some underlying dynamics, what is the dynamical status of a "syntax" relating these attractors ? *What can be a dynamical theory of syntactic constituency ?*

In the last years, a great deal of works has been devoted to this question. In 1991 and 1992 two Conferences, organized by Daniel Andler,

Elie Bienenstock and Bernard Laks ([COMPCOG I, 1991] and [COMPCOG II, 1992]), have been held in France at the Royaumont Abbey on this subject. Another Conference on *Dynamic Representations in Cognition* has been organized by Robert Port and Tim van Gelder at Indiana University in 1992. It led to a reference book, *Mind as Motion*, to appear in 1995 at the MIT Press.

The particular status of dynamical explanations, contrasting with the deductive-nomological ones, has been stressed by Tim van Gelder [1992]. The epistemological problems raised by a dynamical interpretation of basic theoretical concepts such as "structure", "constituency" and "syntax", have been analyzed with care by Daniel Andler [1990] and Yves Marie Visetti [1990]. They are not exclusively linguistic. They arise also in the realm of perception.

What can be a dynamical theory of syntactic constituency ? The problem is difficult for at least two reasons.

a) Weak CN vs. Strong CN²

To do syntax — deep universal and formal syntax and not English or French morpho-syntax — we need at least to distinguish :

(i) between two syntactic (categorical) types : things or objects (terms) and relations, and

(ii) between two types of relations : static and dynamic³.

Actually, a process described by a verb is a dynamical transformation of actantial relations. But, if we represent terms by activity patterns which are attractors of dynamical systems implemented in CN networks, the challenge becomes : how can we represent these two differences "term vs. relation" and "static vs. dynamic" ? It is clear that syntactic relations between attractors cannot be reduced to their mere linear superpositions. We call *weak CN* a CN which models semantic entities of *different* syntactic types by attractors of the *same* dynamical type, without taking into account the difference of their grammatical categories (this is a category mistake). To work out a response which can take up the challenge we need to strengthen weak CN into a *strong CN* having the capacity to model different grammatical categories by mathematical entities of *different* types.

b) Elementary vs. Non-elementary CN Syntax

One could think that it would be trivial to elaborate a strong CN. One would have only :

(i) to represent activity patterns (attractors) coding the terms by units belonging to some higher level layer (what is called in the literature a "localisation" through "grand-mother" neurons) and,

²In the sequel, CN abbreviates "connectionism" and "connectionist", CNC "connectionist cognitivism", CL "classic", and CLC "classical cognitivism".

³We must distinguish three meanings of "dynamical" : a physical phenomenon is dynamical when the causality of its temporal evolution is concerned ; a relation is dynamical when it is temporal ; a mathematical model is dynamical when it proceeds from the theory of dynamical systems (global analysis).

(ii) to represent the relations by connections between these units.

We call such a solution an *elementary* one. It does not work for it projects analogically the neuronal implementation into the functional architecture. Static and dynamic relations between terms must be modeled by *dynamical relationships between attractors*, and these relations are of a completely different nature than the underlying connections they are implemented in. We call *non-elementary* a dynamical model which does justice to this principle.

The main problem to be solved is therefore the following one :

Main problem : under the initial hypothesis that contents can be modeled by attractors, can an "attractor syntax" be worked out in the framework of a strong non-elementary CN, that is, of the theory of dynamical systems ?

In several papers [Petitot, 1979, 1989b, 1989d, 1991b, 1993b, 1994, forthcoming 1995], we have shown how this problem could be solved using bifurcation theory. We will not recall here this technical stuff. Our purpose will be to focus on more epistemological points. We will in fact analyze the debate which opposed Jerry Fodor, Zenon Pylyshyn and Brian McLaughlin to Paul Smolensky. In early 1988, Jerry Fodor and Zenon Pylyshyn published an important paper in *Cognition*, "Connectionism and Cognitive Architecture : A Critical Analysis" [Fodor & Pylyshyn, 1988]. It was a radical criticism of the thesis held by Paul Smolensky in his paper "On the Proper Treatment of Connectionism" appeared in a special issue of *The Behavioral and Brain Sciences* [Smolensky, 1988]. A little later, in his talk "Why Connectionism is such a Bad Thing ?", held at the Ecole Normale Supérieure in Paris, J. Fodor enhanced his criticisms.

As the debate is quite fundamental to CNC, we want to evaluate carefully its arguments. Our own arguments will seek to show :

(i) that Fodor and Pylyshyn's criticism is essentially valid when it concerns weak CN ; but that,

(ii) it is not at all acceptable for a more elaborate form of CN which makes use of the full mathematical power of the theory of dynamical systems ("strong CN").

I. Connectionism and the theory of dynamical systems⁴

1. The CNC's main thesis and its precursors

It is well known that one of the basic ideas of CNC is that the automatisms of competence described at the symbolic level by discrete

⁴The core of the paper develops some previous reflections (1988) published in [Petitot, 1991a]. They were presented at the Workshop "Connectionism and Language" held at the International Center for Semiotic and Cognitive Studies at the University of San Marino in October 1989. My thanks are due to Franson Manjali for some parts of the English translation. I thank also Gertrudis van de Vijver and Barry Smith for their relevant suggestions.

and sequential structures (symbols, symbolic expressions, rules, inferences, etc.) are formal and macroscopic "kinematical" structures which emerge from underlying microscopic "dynamics" of performance in a qualitative, stable and invariant way, and which have to be described at a subsymbolic level⁵. Classical symbolic cognitivism (CLC) is analytic and constructivist. It favours logical automatisms, conscious rules, calculus, and deductive inference. Subsymbolic cognitivism (CNC), on the contrary, is synthetic and associationist. It favours networks dynamics, intuitive performances, equilibrium positions and induction. For it, the macro entities endowed with a meaning are, at the microdynamical underlying level, global and complex distributed activation patterns of elementary meaningless local units. These scattered units are interconnected and process the data in parallel.

This key idea, which has been popularized under the acronym PDP (Parallel Distributed Processing) enhances in fact an idea proposed some twenty-five years ago by René Thom and Christopher Zeeman, according to which a meaning can be modeled, at the "macro" level, by the *topology* of an *attractor* of an underlying "micro" dynamics. These authors introduced the principle — deep but somewhat shocking at that time — that the *mathematical bifurcation theory of the attractors of non-linear dynamical systems should take over from formal logic in cognitive linguistics*. Syntactic and semantic structures were therefore considered as analogous to processes we meet in physics under the name of *critical phenomena*, especially with the thermodynamical phenomena of *phase-transitions*. Following Thom's proposal, if we consider the Liapounov functions on the basins of the attractors, we are led to gradient like dynamics, that is, to dynamics which minimize potential functions H ⁶. It was therefore the bifurcation theory of the minima of such potential functions — that is the *mathematical theory of unfoldings of singularities of differentiable maps* — which became the basic tool for modeling semantic and syntactic structures (cf. [Thom, 1972, 1980], [Zeeman, 1977] and [Petitot, 1985a, 1985b]).

2. CN networks and dynamical systems

Instead of being describable by formal languages where the semantics is simple and the syntax complex, the subsymbolic systems possess, on the contrary, a simple syntax and a complex semantics. The instantaneous global state of a network S of N elementary interconnected units (formal neurons) u_i is given by their activation vector : $x = (x_i)_{i=1, \dots, N}$, where x_i takes its values from a set of values V (finite or continuous). The main cases are $V = \{0, 1\}$, $\{-1, 1\}$, $[0, 1]$. Let $M = \{ (x_i)_{i=1, \dots, N} \}$ be the configuration space of S . It is called the *internal space* of S .

⁵We use the opposition between "competence" and "performance" in its classical Chomskyan sense.

⁶If X is a dynamical system (that is, a smooth vector field) on a differentiable manifold E , the attractors of X are the asymptotic and structurally stable limits of the trajectories of X . They are not necessarily punctual (that is the case for quasi-periodic movements on torus attractors) and their topology may be very complex (that is the case for the so called "strange" attractors). If A is an attractor of X , the dynamics is complex (ergodic, chaotic) on A . But on the complementary subset $B(A) - A$ of A in its basin $B(A)$, there always exists a real valued continuous positive function L , called a Liapounov function, which is strictly decreasing along the trajectories and which vanishes on A . So, on $B(A) - A$, the dynamics is of simple type. It minimizes a sort of energy function.

The units u_i are interconnected by connections characterized by their *synaptic weights* w_{ij} . They determine the computational program of the network. The $w_{ij} > 0$ correspond to excitatory connections, and the $w_{ij} < 0$ to inhibitory ones. In general $w_{ii} = 0$. The network computes in the following way. Each neuron receives afferent signals from its presynaptic neurons, changes its state according to a local transition law, and sends an efferent signal to its postsynaptic neurons. In general, the input of the neuron u_i is defined as the weighted sum of the afferent signals :

$$h_i = \sum_{j=1}^{j=N} w_{ij} x_j .$$

The simplest hypothesis is that u_i is a threshold automaton, the states of which are controlled by a local transition of the form :

$$x_i(t+1) = g(h_i(t) - T_i) ,$$

where T_i is a threshold, and g a gain function. Typically,

- g = the Heaviside function if $V = \{0, 1\}$;
- g = the sign function if $V = \{-1, 1\}$;
- g = a sigmoid function = e.g. $1/(1+e^{-x})$ if $V = [0, 1]$.

The weights w_{ij} and the threshold T_i vary in a control space W which is called the *external space* of S .

The local transition laws of an elementary unit from one state to another, considered as a function of the information that this unit receives from its immediate neighbors, define therefore an endomorphism $T : M \rightarrow M$. T associates the next state $T(x)$ with the instantaneous global state $x \in M$ of S . In general, it encodes a considerable amount of information. It is the *iteration* of T which defines the *internal dynamics* of the network S . The stable asymptotic states of T (its attractors) are the *internal states* of S . If x is an input of S (an instantaneous initial state), its trajectory $(T^k(x))_{k \in \mathbb{N}}$ will tend, in general, towards an attractor A which will be the output (the response) of S to the input x . The key dynamical phenomenon is therefore the asymptotic capture of the (global and instantaneous) initial state of the net by an attractor. CN nets compute in a way radically different from that of Turing-von Neumann machines. They are dynamical computers which bifurcate from attractors to attractors.

On varying the local transition laws — for instance by changing the weights of the connexions of S — T_w can be modified. The parameters w of the external space act therefore as controls on the internal dynamics T_w . This fundamental fact yields a basis for a theory of *learning*. The learning

problem is in fact a typical *inverse problem*. The direct problem is, given the synaptic weights w_{ij} , to find the attractors A_k of the internal dynamic T_w . The inverse problem is on the contrary, given attractors A_k , to find specific weights so that the corresponding attractors include the A_k . It is a difficult one, but it can be solved using some well known algorithms, such as the *backpropagation* one. These algorithms define *external slow dynamics* in the external space W .

The internal dynamics yielded by general nets can be tremendously complex. In general, attractors will have an *internal structure*, a non-trivial topology. This fact is fundamental, since one can think that the topology encodes the internal intrinsic semantics⁷ of the associated representations. But if the internal dynamics T is gradient (see above), then this semantics reduces to a trivial one, in as much the way as, in CLC, one represents a content by a symbol. This will be the case if T consists in minimizing an "energy" function $H : M \rightarrow \mathbf{R}$ ($\mathbf{R}$ = the real line). We get that way the elementary catastrophic models introduced for syntax and semantics twenty-five years ago ([Thom 1969, 1972, 1980], [Petitot, 1979, 1982a, 1982b, 1989b, 1989d], [Wildgen, 1982]).

This is the case when the connections are *symmetric*. Hopfield has shown that, for $V = \{-1, +1\}$ and g = the sign function, the equations of the net are exactly those of a system of *interacting spins*. The value E of the energy H minimized by the dynamics is given in that case by the formula :

$$E = -\frac{1}{2} \sum_{i \neq j} w_{ij} x_i x_j + \sum_i T_i x_i .$$

As far as the synaptic weights correspond to coupling constants and mix > 0 and < 0 values, we see that these nets — which exemplify *the simplest* cases of neural nets — correspond to *the most complex* cases of spin systems, namely that of spin *glasses* (see e.g. [Mézard et al., 1987]).

In general, the set M_{min} of minima of the energy H has a very large cardinal, and to use a statistical approach becomes therefore necessary. One introduces a "computational" temperature T and starts from an initial configuration x_0 with a high T value. This allows the system to access all the basins of E . One selects then randomly a neighboring configuration x_1 . If $\Delta E < 0$ the transition from x_0 to x_1 has probability 1. If $\Delta E > 0$, it has the probability :

$$\frac{1}{1 + \exp\left(\frac{\Delta E}{T}\right)}$$

(this enables the system to overcome the thresholds). One iterates this process until the system reaches a local minimum. One lowers then T and starts again.

⁷The internal intrinsic semantics must not be confused with an "external" denotative one. Using the Fodorian opposition between "narrow" and "large" contents, we can say that it concerns the narrow content of the representation.

More precisely, let us consider distributions of probabilities $P(x)$ over M . For a given mean energy C , that is for $\sum_{x \in M} P(x) H(x) = C$, the distribution $G_T(x)$ which maximizes the entropy is the Gibbs distribution :

$$G_T(x) = \frac{1}{Z_T} \exp\left[-\frac{H(x)}{T}\right],$$

where Z_T is the partition function :

$$Z_T = \sum_{x \in M} \exp\left[-\frac{H(x)}{T}\right].$$

When the computational temperature $T \rightarrow 0$, G_T concentrates on the minima of H , i.e. on M_{min} . This is a well known physical result. One can derive from it algorithms for minimizing H . One of the best known is “simulated annealing” (cf. [Azencott, 1988]). It is based on the construction of sequences of random variables X_n and of temperatures T_n such that :

$$\begin{aligned} \lim_{n \rightarrow \infty} P(X_n = x) &= G_T(x) \\ \lim_{n \rightarrow \infty} P(X_n \in M_{min}) &= 1. \end{aligned}$$

As it is well explained by R. Azencott [1988], we are given :

(i) An exploration matrix $Q = (q_{x,y})_{x,y \in M}$ satisfying the condition that, for all $x, y \in M$, there exists a chain (x_k) linking up x and y and such that $q_{x_k x_{k+1}} > 0$ for all k . If we call $V_x = \{y \mid q_{x,y} > 0\}$ the set of neighbors of x , this condition means that two configurations x and y are always connectable by a chain of neighboring configurations.

(ii) A sequence T_n of “temperatures” such that $T_n \rightarrow 0$ (cooling schedule), the decreasing of T_n being sufficiently “slow”. This later condition is expressed by the fact that $\lim_{n \rightarrow \infty} T_n \log(n) = R$ where R is a sufficiently large constant.

(iii) A Markov chain of random variables X_n on M such, that for $y \neq x$

$$P_n(x, y) = P(X_{n+1} = y \mid X_n = x) = q_{x,y} \exp\left[-\frac{(H(y) - H(x))^+}{T_n}\right]$$

(where $a^+ = a$ if $a \geq 0$ and $= 0$ else), and

$$P_n(x, x) = 1 - \sum_{y \neq x} P_n(x, y).$$

A configuration x is a *local* minimum of H ($x \in M_{locmin}$) if $H(x) \leq H(y)$ for all $y \in V_x$. Its depth $D(x)$ is then defined as the minimal height of the thresholds (the saddles) which limit its basin of attraction.

A theorem due to Hajek says that :

$$\lim_{n \rightarrow \infty} P(X_n \in M_{min}) = 1$$

$$\text{iff } \sum_{n=1}^{n=\infty} \exp\left(\frac{D}{T_n}\right) = \infty, \text{ where } D = \sup\{D(x) \mid x \in M_{locmin} - M_{min}\}.$$

When the synaptic weights become *asymmetric* there exists no longer an energy function and the dynamics will be in general extremely complex. Steve Renals et Richard Rohwer have considered systems :

$$x_i(t+1) = g\left(r \sum_{j=1}^{j=N} w_{ij} x_j(t)\right)$$

where r is the slope of the sigmoid function g . They have retrieved that way many classical routes towards chaos, and in particular, for $r \in [12, 14]$, the period doubling subharmonic cascade of Coullet-Feigenbaum-Tresser (cf. [Renals & Rohwer, 1990]).

H. Sompolinsky, M. Samuelides and B. Tirozzi have also investigated such systems when N becomes very large and when the asymmetric synaptic weights w_{ij} are random variables with mean value = 0 and variance = w^2/N (cf. [Sompolinsky *et al.*, 1988], [Doyon *et al.*, 1993] and [Tirozzi & Tsodkys, 1991]). For the critical value $rw = 1$, they present a *phase transition* from a convergent regime to a *chaotic* one.

Many similar results show that it is henceforth possible to give a rigorous mathematical and physical status to the deep Thomian idea that the mental contents can be modeled by complex attractors of complex dynamical systems implemented in neural nets.

3. Harmony theory

In his harmony theory, Paul Smolensky [1986] applied these ideas to many subsymbolic processes and stressed the importance of going beyond the (von Neumann) computational conception towards a dynamical conception of information processing. His aim was to link up the higher

levels of cognition with the lower levels of perception. For example, in the interpretation of a visual scene, we can suppose that the cognitive system undergoes the following operations.

(i) Let $(r_i)_{i \in I}$ be a set of representational features which "constitute the cognitive system's representation of possible states of the environment with which it deals". A representational state is thus a vector r of values of the r_i (± 1 , for example). The cognitive system interprets its environment using "knowledge atoms". Each of these atoms α is characterized by a knowledge vector k_α which attributes values to each feature r_i . It is or it is not activated (we introduce an activation variable $a_\alpha = 0/1$). The atoms α encode sub-patterns of values of the features occurring in the environment. Their frequency is encoded in their force σ_α .

(ii) Let (r, a) be the state of the cognitive system and let K be the knowledge base defined by k_α and σ_α . We define H_k by :

$$H_k(r, a) = \sum_{\alpha} \sigma_{\alpha} a_{\alpha} h_k(r, k_{\alpha})$$

with $h_k(r, k_{\alpha}) = \frac{r \cdot k_{\alpha}}{|k_{\alpha}|} - k$.

(iii) We now apply to this particular case the general thermodynamical method sketched above. This allows us to interpret the visual scene (i.e. the vector r) by completion, that is, by optimizing the *global coherence* (the consistency) of the (local) partial interpretations α : for a given representational state r and for an activation state a of the cognitive system, H is a sum including a term for each knowledge atom α , each term being weighted by the force σ_α of r and each weight σ_α multiplying the self-consistency between r and α (i.e. the similarity between r and k_α).

(iv) This inference and decision process is then identified with the result of a parallel and distributed stochastic process driven by H_k .

4. Some epistemological issues concerning the "morphological turn"

4.1. The morphodynamical agenda

The development of a morphodynamical explanation of performance raises a number of hard epistemological issues [Petitot, 1979, 1982a, 1985a, 1987, 1988]. The most important feature of the "morphological turn" consists in a radical questioning of the logico-combinatorial formalist point of view on perception and language (and more generally on cognition).

In a formalist conception which reduces syntax to a formal *description* of competence automatisms (see for example the Chomskyan transformational and generative conception), all the levels of description are of the *same* formal type. Consequently, in regressing from surface structures towards deep structures, one ends up with abstract primitive ("atomic", "kernel") structures which are *of the same formal type* as the surface ones (for example syntactic trees). It is thus impossible, on the one hand, to establish their link with the structures of perception, and on the other hand, to understand their genesis and their emergence in terms of the underlying dynamical mechanisms of performance. Now, the structures of perception, as well as the dynamical mechanisms of performance, impose certain *universal* constraints on grammatical structures. If one does not take them into account, one is committed to interpret them as genetic constraints.

That is why, since 1975, we have argued against the CL Chomskyan "evidence" that our ignorance of the physical foundations of mental structures forces us to an abstract characterisation of them. The conclusion of such a "bad" syllogism is that structural properties of language, which cannot be derived from such an abstract characterisation, must be explained in innatist terms. It seemed to us, on the contrary [Petitot, 1979], that the "good" syllogism was rather the following :

(a) we do not know yet the physical (neurophysiological) bases of language ;

(b) but we can nonetheless hypothesize them and thus assume the existence of dynamical processes underlying performance, processes from which emerge the formal and abstract kinematical structures of competence ;

(c) formal grammars formalize only certain aspects of these emerging formal structures ;

(d) but there exist other aspects, linked with perception, which impose additional cognitive constraints on the "humanly accessible" grammars.

It is such a strategy which was provided by Thom's topological syntax, where a meaning is coded by the topology of an attractor, and where the syntagmatic trees are trees of bifurcations of attractors into sub-attractors.

4. 2. The PTC agenda

Very similar questions have been recently tackled by P. Smolensky. Adopting a dynamical point of view in semantics and an emergentist one in syntax (emergence of formal, discrete, sequential structurally stable structures), he has very clearly presented their bases, their characteristics, and their epistemological consequences. Let us recall briefly the essential points of his article "On the Proper Treatment of Connectionism" [Smolensky, 1988] :

(i) The CN level is neither the conceptual and symbolic CL level nor the neuronal one. It does not concern the implementation of cognitive algorithms in massively parallel machines, but rather the structure, the architecture and the dynamical behaviour of the cognitive processes themselves.

(ii) We cannot model the performances of *intuitive* knowledge on the assumption that the intuitive processor applies, simply in an unconscious way, sequential programs of formal rules. The processes at work are not adequately modeled in the framework of the CL symbolic paradigm, where symbols denote external objects (denotative semantics) and are operated upon syntactically by application of rules.

(iii) In a dynamical CN model, the units possessing a semantics are complex patterns of activity distributed over many elementary units. This conception of semantics is characteristic of the CN approach.

(iv) Whereas in the symbolic formal models, all the processing levels are of the same type, in the subsymbolic semantic ones there exists a semantic shift :

“Unlike symbolic explanations subsymbolic explanations rely crucially on a semantic shift that accompanies from the conceptual to the subconceptual levels”.

(v) In order to get a unified conception of cognition, it is necessary to combine the CN and CL approaches. The rules consciously applied by the conscious rule interpreter will then be interpreted as structurally stable emergent regularities :

“patterns of activity that are stable for relatively long periods of time (of the order of 100 ms.) determine the contents of consciousness”.

(vi) As far as the linguistic rules are concerned, this presupposes in particular the possibility of representing subsymbolically and subconceptually in CN dynamical models the propositional structures of language. This is certainly very difficult, but it is necessary. Now, according to Smolensky, a constituent structure can be obtained for the patterns of activity possessing a conceptual semantics, when considering them as *superpositions of subpatterns* (constituents). As we will see, this is the Achille’s heel of the PTC agenda and Fodor and Pylyshyn’s criticism focuses on this very point.

(vii) The mathematical universe of CN models is not that of formal languages and of Turing machines. It is that of dynamical systems, that is, of the qualitative theory of differential equations (global analysis). The subsymbolic computation is *continuous, geometrical and differential*. Inference here is a statistical inference which optimizes the fit with the input (harmony theory) :

"macro-inference is not a process of fixing a symbolic production but rather of qualitative state change in a dynamical system, such as phase transition".

(viii) The CN dynamical systems are suited for the elaboration of a theory of schemata, of prototypicality and of categorisation. (This point also has been already well elaborated in the Thomian morphodynamical approach : see [Petitot, 1983a, 1985b, 1989a].)

II. Fodor and Pylyshyn's (F-P) arguments against connectionism

1. The general structure of the F-P arguments

Let us begin with a summary of Fodor and Pylyshyn's main points :

A1. Both classical (CL) and connectionist (CN) cognitivists are *representationalist*. They both admit mental states encoding the properties of higher level cognitive processes and, therefore, CLC/CNC debate is internal to cognitivism. It concerns a precise issue : "the architecture of representational states and processes"⁸. The CN paradigm is thus committed to show that it can provide a good theory of the cognitive architecture, that is, of "processes which operate on the representational states of an organism".

⁸In this section, all quotations are from [Fodor & Pylyshyn, 1988].

A2. Now, there exists a fundamental difference between the CL and CN paradigms. CL cognitivists assign semantic content to expressions, and admit between semantically evaluable entities, not only causal relations but also, *structural relations*. They consider it to be characteristic and essential :

(i) that mental representations share a combinatorial syntax and semantics, and

(ii) that mental processes *are dependent on this structure* ("structure sensitive") : operations operate on the mental representations as a function of their combinatorial structure, that is, of their syntactic form.

On the contrary, according to the authors, CNC would assign semantic content only to holistic entities, without internal combinatorial structure (labeled nodes symbolizing activity patterns). Further, they would only admit *causal* relations between the semantically evaluated entities. In a nutshell, only CL cognitivists

"are committed to a symbol-level of representation, or to a "language of thought", i.e. to representational states that have combinatorial syntactic and semantic structure".

Contrary to CNC, CLC insists upon the fact that computational operations act on the syntactic structure of complex symbols and that, as far as the syntactic relations are parallel to the semantic ones,

"it may be possible to construct a syntactically driven machine whose state transitions satisfy *semantical* criteria of coherence".

Such is "the foundational hypothesis of Classical cognitive science".

A3. According to the authors, the limits of CNC are quite dramatic since mental representations *must* possess an internal syntactico-semantic combinatorial structure, if they are to explain four fundamental aspects of cognition.

(i) *Productivity and Generativity*. As all natural languages, the "language" of thought is able to generate an indefinite number of expressions from finite means. Consequently, there must exist rules of generation, and this presupposes an internal structure of the expressions.

(ii) *Systematicity*. Even if we challenge the productivity and the generativity of cognitive capacities, we cannot reasonably challenge their systematicity, that is, the intrinsic links relating the comprehension and the production of certain expressions with those of certain other expressions. Systematicity is explainable only if there exists an internal structure of expressions providing well-formedness rules and structural relations between different expressions.

(iii) *Compositionality*. There exist *semantic* transformations (a "covariance") between systematically related expressions (like "John loves Mary" and "Mary loves John", or like, "being a brown cow", "being brown" and "being a cow", etc.). The principle of compositionality according to which the semantic properties of constituents are independent of the context can be understood only if there exists a syntactico-semantic constituent-structure.

(iv) *Inferential coherence*. The relations of logical similarity between different inferences presuppose the same conditions.

A4. In a nutshell, if we accept an internal structure of representations, then we are legitimate to speak of representations sharing the same structure, of similar structures, or of structures which are related to each other in different ways. But, according to the authors, an essential feature of CNC would be to reject such a structure. For CNC, cognitive systems are systems :

"that can exhibit intelligent behaviour without storing, retrieving, or otherwise operating on structured symbolic expressions".

Of course, the labels which identify semantically evaluable holistic entities, will have, in general, a constituent-structure, but the processual dynamics of the system is *not* determined causally

“by the structure –including the constituent-structure– of the symbol arrays that the machines transform”.

CN graphs are *not* structural descriptions of mental representations, but specifications of purely causal relations.

“The intended interpretation of the links as causal connections is intrinsic to the theory”.

“A network diagram is not a specification of the internal structure of a complex mental representation. Rather, it's a specification of a pattern of causal dependencies among the states of activation of nodes”.

On the other hand, the fact that mental representations are distributed over micro-features extracted by multivariational analysis from the statistical regularities of the stimuli samples *does not* imply that they are structured. Actually,

“you have constituent-structure when (and only when) the parts of semantically evaluable entities are themselves semantically evaluable”.

“Complex spatially-distributed implementation in no way implies constituent-structure”.

The main error of CNC, its “major misfortune”, is to have mixed up a componential analysis of micro-features with a combinatorial structure.

“The question whether a representational system has real-constituency is independent of the question of micro-feature analysis”.

“It really is very important not to confuse the semantic distinction between primitive expressions and defined expressions with the syntactic distinction between atomic symbols and complex symbols”.

In short, from the moment when the semantically evaluable entities (nodes, activation-patterns, etc.) are conceived of as *atomic* and *holistic* Gestalts related only by causal relations, it becomes impossible to account for the fundamental features of cognition, namely, productivity, generativity, systematicity, compositionality, and inferential coherence (cf. A3).

“The connectionist architecture (...) has no mechanism to enforce the requirement that logically homogeneous inferences should be executed by correspondingly homogeneous computational processes”.

CNC presupposes a systematic organization of cognition. But it should also be able to explain it.

"It's not enough for a connectionist to agree that all minds are systematic ; he must also explain *how nature contrives to produce only systematic minds*".

Hence the final verdict :

"The only mechanism that is known to be able to produce pervasive systematicity is classical architecture. And (...) classical architecture is not compatible with connectionism since it requires internally structured representations".

A5. Further, according to the authors, CNC's main criticism against CLC is not acceptable. It claims that, for CLC, the behavioral regularities must come from explicitly encoded rules. But this is false. In fact, for CLC, several functions can be encoded implicitly (for example, as part of the hardware). What should be explicit are only the data structures that the cognitive machines transform and not the rules (the grammar) of transformations.

A6. As a consequence, the CN perspective should be rejected as a *cognitive* theory. It relies upon a "bad" associationist and empiricist psychology against which one can repeat the well known rationalist criticisms formulated since Kant.

A7. The authors conclude that the only real interest of CNC is to provide an alternative theory of *implementation* for the classical functional architecture. They stress the fact that most of the arguments put forward by CNC bear only on the limitations imposed on competence by the concrete constraints of performance. According to them, the material limits of performance result from an interaction between an unlimited formal competence (unlimited but finitely describable by generative rules (see A3 (i)) on the one hand and the limited resources on the other. Adopting a functionalist perspective radically opposed to the emergentist CN one, they separate drastically the functional architecture (the software algorithms) from its implementation (the hardware). This is, for them, a "principled distinction". The (micro-level) models of implementation are neutral with respect to the nature of cognitive (macro-level) processes and to deny this fact is to confuse structure and function.

Such a confusion leads to catastrophic consequences. For example,

(i) from the evident existence of neuronal networks, one will conclude to an associationist psychology (networks of representations), or

(ii) from the not less evident anatomic distributivity of neurons, one will conclude to a functional distributivity of the mental representations themselves (componential analysis in micro-features), or

(iii) from the reinforcement of the connection of two neurons by their co-activation, one will conclude to associationist statistical models of learning, or still

(iv) in the other direction, from a functional locality (position of a symbol in an expression, for example), one will conclude to a physical localisation in instantiation.

The “brain style” of CNC is quite definitely a dramatic epistemological error : it makes

“the implicit —and unwarranted— assumption that there ought to be similarity of structure among the different levels of organisation of a computational system”.

It projects the neuronal level onto the cognitive one, and so doing, reactivates “the worst of Hume and Berkeley”.

A8. Thus, the CN stance may only provide

“an account of the neural (or “abstract neurological”) structures in which classical cognitive architecture is implemented”.

The symbolic structures of CLC are of course physical ones. They are neurally encoded and instantiated and it is their physical implementation which causes the operational behavior of the cognitive system. The CN arguments become therefore valid if we interpret them as arguments in favour of a physical implementation in massively parallel networks. For example, the fact that cognitive processes are fast, whereas neuronal phenomena are slow, or the fact that a considerable amount of forms (words, faces, etc.) stored in memories can be quickly recognized, or still, the continuity, the fuzziness, the approximation and the structural stability properties of cognitive processes, all these facts support the thesis of a CN implementation of the algorithms making up the CL functional architecture. But if the CN models are rather to be seen as concerning only implementation, then they should surrender their cognitive claims. They should in particular refuse to assign

“a representational content to the units (and/or aggregates) that they postulate”.

A9. A key argument which is not made explicit by the authors is that “structural” necessarily means “formal-symbolic”. If mental representations share a combinatorial syntax and semantics then they are *ipso facto* “symbol systems”. As we will see, it is this *formalist dogma* — the dogma of logical form — which is the Achille’s heel of all their arguments.

2. Comments : the problem of a dynamical structuralism

F-P arguments are well constructed and apparently forceful. However, we can question their effective validity at different levels.

C1. Arguments A1 and A2 (as regards the characterization of CLC), A3., A5. and A8 (except for its conclusion) are, we think, excellent and convincing. But they *do not* imply a rejection of CN as a cognitive theory. They simply impose to it certain constraints and additional requirements (as explained in A1) : to be able to develop what we will call the *structural hypothesis*.

C2. The presentation and characterization of CNC given in A2 and A4 are caricatures. They “demonstrate” only the following “syllogism” :

- (i) a “good” CNC should be able to develop the structural hypothesis ;
- (ii) for intrinsic reasons, the caricature of CNC presented in A2 and A4 does not fulfill this requirement ;
- (iii) therefore CNC, whatever can be its further developments, is *de jure* a “bad” cognitive theory.

But of course, nothing proves that the caricature can be identified with the full theoretical power of CNC. It is why the argument is dogmatic.

Let us continue to call “dynamical” the CN cognitivism. The central question is the following :

Question : Just as it is possible, using appropriate formal theories, to develop a symbolic structuralism, is it also possible, using the mathematical theories of dynamical systems, to develop a dynamical structuralism ?

If we reduce *a priori* all possible CN models to graphs of causal relations between holistic units lacking internal structure, then the response is of course trivially negative. *But these elementary models are only a very tiny part of the mathematical theory of dynamical systems.* We shall return to this point later. It is essential.

C3. Even if we could accept that, contrary to the assertion of the authors, it is possible to elaborate an authentic dynamical structuralism, this would not lead to transform the CL/CN opposition into a Manichean alternative. There is certainly higher processing levels of the cognitive system which are of symbolic nature. But this does not entail that there do not exist lower levels which are of a dynamical nature. Associationist processes are certainly not sufficient to explain the structure of cognition, but they can be nevertheless necessary. Logico-symbolic superstructures can possess associationist infrastructures. The question is not whether

CNC should overcome CLC (or if the later should excommunicate the former), but to find out whether *the structural hypothesis can or cannot be already elaborated at the dynamical level of cognitive processes*. Such a dynamical structuralism must be clearly distinguished from the formal symbolic one. It must be :

(i) an authentic structuralism,

(ii) a *proto-symbolic* one, that is, one which is compatible with the symbolic level.

If one still wants to criticize it, one must develop more refined arguments than the F-P ones (cf. C6).

C4. Fodor and Pylyshyn seem deeply unaware of the true nature of the *emergence* of a macro-level from a micro-level. By separating the functional level of algorithms from the level of implementation, they disregard what is really the central issue in the point of view they are attacking. However, the physical parallel they suggest should have incited them to more circumspection.

"The point is that the structure of 'higher levels' of a system is rarely isomorphic, or, even similar, to the structure of 'lower levels' of a system. No one expects the theory of protons to look very much like the theory of rocks and rivers, even though, to be sure, it is protons and the like that rocks and rivers are 'implemented in'".

The argument is fallacious. Physics is radically anti-functionalist. In physics, the relation between micro-levels and macro-levels is a matter of emergence. No physicist would separate the levels and postulate, as the authors do, that micro-levels are "neutral" in relation to macro-levels and that the later are thus independent of their "implementation". The "form/matter" opposition, which is of Aristotelian origin, has been eliminated in modern science. The very physical problem here is to understand how an emergent — and therefore *non-independent* — macro-level can present nonetheless a certain *autonomy* of structure. That two levels are of a different nature *does not* imply they are independent and "neutral" in relation to each other. To assert this is to misunderstand totally the epistemology and ontology of emergence. CN cognitivists are thus right when they distinguish the problem of implementation from the *intra-cognitive* problem of the emergence of a symbolic level out of a dynamical subsymbolic level. But we repeat that CL cognitivists are right when they assert that this dynamical level, in order to be considered as a cognitive one, should be *structural*.

C5. This last point is the crux of the problem. The authors attack CNC because it assumes the systematic organization of cognition without

explaining it (A4). But the same argument can be turned against them. For they themselves *do not* explain this systematicity. They only *describe* it formally. By reducing the performance constraints to the material concreteness of implementation, by separating, according to their functionalist perspective, the levels and by autonomizing competence they can surreptitiously identify a formal logico-combinatorial description of competence with the development of the structural hypothesis. *But this identification is possible only if we assume the thesis A9 according to which structural $\equiv$ symbolic (logico-combinatorial).* But, if we admit this equivalence, then the argument becomes trivial : CNC is not symbolic (by definition), “hence” it is not structural, “hence” it cannot account for the structural character of cognitive processes.

A formal symbolic description of mental representations and of mental processes is clearly possible. But as such, it should not be mixed up with an explanation. To get an explanation, we must :

(i) model the semantically evaluable entities by *mathematical* structures — perhaps very sophisticated — of a certain type, that is, belonging to a certain mathematical universe ;

(ii) show that a theory of structures can be developed within this universe.

The question of CNC then becomes (cf. C2 and C3) :

Question : If semantically evaluable entities are modeled by attractors of dynamical systems, is it or is it not possible, within the framework of the theory of dynamical systems, to develop a theory of structure ?

C6. If CL cognitivists content themselves with a mere formal symbolic description, it is because the explanation of cognitive structures must be for them of an innatist nature. Behind the CL/CN controversy and the conflict of arguments, behind the rationalist criticism of empiricist associationism, lies, in fact, an epistemological alternative. It was brilliantly sketched by Massimo Piatelli-Palmarini in his paper “Sélection sémantique et sélection naturelle : le rôle causal du lexique” [Piatelli-Palmarini, 1988]. The argument goes as follows.

In all biological domains, one has progressed from *instructivist* (“Lamarckian”) theories to *selective* (“Darwinian”) ones. In every case, both experimentally and theoretically, one has reached the conclusion that there cannot be a transfer from the structure of environment to the organism, and that only mechanisms of *internal* selection can be mechanisms of learning. This internal selection involves filtering and fixation of parameters which selectively stabilize certain possibilities

among a *very rich* universe of genetically *determined* possibilities. From the *instructivist* point of view, the genetic constraints are poor and structuration comes from general capacities, such as adaptations, resolutions of problems by trial and error, etc. For the selective point of view, on the contrary, the genetic constraints are essential, the structuration is strongly innate and modular, and adaptation is replaced by "exaptation", that is, by the fact that the characters can be selected independently of all adaptive value, even if, later on, they acquire such a value. For the selective thesis, the impossibility for an organism to assimilate external structures is a nomological one : it is nomologically improbable that

"structures external to the organism might possibly be 'internalized' through a 'learning' process" ;

it is but nomologically very probable that

"a process of selection, of triggering and parameter-fixation, acting on a vast, profligate and highly articulated repertoire of innate structures may prove to be the most productive explanatory hypothesis" [Piatelli-Palmarini, 1988, p. 23].

It is such an innatist and selective point of view which is now further developing in the domain of the cognitive sciences, in syntax as well as in semantics. Hence the radical critique undertaken by Chomsky, Fodor and their colleagues against empiricist theories of learning by imitation, association, assimilation, induction, problem-solving, etc. Many results seem to indicate that there exists a rich syntactico-semantic architecture of language whose universality is of a genetic origin :

"our species innately possesses a rich, specific, modular and highly articulate capacity for language, organized around certain universal 'principles' ".

This cognitive capacity would be independent of perception and action. It would manifest

"a very intricate and closely inter-dependent process, full of 'deductive' consequences that are known to each of us in a totally unconscious way".

That is why, genetic constraints being contingent, a formal description can amount to an explanation.

It is this formalist dogma we are criticizing. For its arguments are relevant only in the symbolic framework. They do not imply at all that the innate symbolic form of the cognitive system exhausts its structure. It is indeed perfectly legitimate to assume :

(i) that there exists an objective content on which this form operates ;

(ii) that a *dynamical* functional architecture can also be innately constrained.

3. The main point of the F-P argument

Let us come now to the central argument of Jerry Fodor and Zenon Pylyshyn. They consider the way in which some CN cognitivists (Hinton, McClelland, Rumelhart) treated a sentence like "John loves Mary". The problem is, evidently, that of

"the role relations that traditionally get coded by constituent-structure".

They accept therefore with great fair-play the conception of syntax which is the least symbolic and the most akin to the CN sensibility, namely that of *case grammars*. But they stress the fact that, to be acceptable and amenable, CNC must provide a good CN account of the semantic roles which select cases. This is the main problem : to model in a CN framework what European linguistics and semiotics call *actantial relations*⁹.

⁹We use here the gallicisms "actant", "actantial" and "actantiality" for what concerns the case semantic roles.

The CN cognitivists mentioned above represent actantial relations by a set of activated units such as {John-subject ; +loves ; +Mary-object}, where the descriptors *J-S*, *L*, *M-O* are labels of holistic units *without* internal syntactic structure and *without* structured inter-relations. These descriptors combine an identity (an actant *J*, *M*) with an actantial role (*S*, *O*) and represent the syntactic structure of the sentence in a set-theoretic manner. Fodor and Pylyshyn can easily show that such a representation immediately leads to a lot of inescapable difficulties which can be solved only by a "grotesque" proliferation of the number of descriptors :

"the idea that we should capture role relations by allowing features like John-subject thus turns out to be bankrupt".

"It is of course no accident that the connectionist proposal for dealing with role relations runs into these sorts of problems. Subject, object and the rest are classically defined *with respect to the geometry of constituent-structure trees*. And the connectionist representations don't have constituents".

If we just superpose additively the activated holistic entities in order to account for the sentences, then it becomes, for example, impossible to account for the relation between *J-S ; +L ; +M-O* and *M-S ; +L ; +J-O* (argument of *systematicity*, cf. A3 (ii)).

"This consequence (...) offers a particularly clear example of how failure to postulate internal structure in representations leads to failure to capture the systematicity of representational systems".

Further, in the case of a conjunction of sentences, it becomes impossible to retrieve the initial structures. The superposition leads to an irreversible destructuring (what is called the "superposition catastrophe").

This is really the decisive point :

"when representations express concepts that belong to the same proposition, they are not merely simultaneously active, but also *in construction with each other*".

And, in order to be in a relation of "construction" — that is, to be related by dependence relations —, representations should be constituents of more complex representations (cf. the arguments A2 and A3).

"Representations that are 'in construction' form parts of a geometrical whole, *where the geometrical relations are themselves semantically significant*".

The main problem is therefore to build-up what we will call a *configurational definition of case roles*.

Of course, for the CL paradigm, the problem of a configurational definition of actantial relations is *a priori* solved by the use of formal and combinatorial symbolic structures. But this does not entail at all that every such configurational definition *must* be, for *de jure* reasons, of a symbolic nature.

4. Towards a geometry of syntax

We see that Fodor and Pylyshyn's criticisms against the possibility of working out a connectionist theory of high level cognitive abilities — and in particular a theory of syntactic constituency and compositionality — are based on two thesis.

a. *A thesis concerning the internal constituent structures — that is the internal form — of mental representations.*

The structures they consider are principally case-structures where semantic roles (the "actantial relations") are "in construction with each other". These semantic roles "classically defined *with respect to the geometry of constituent-structure trees*" are geometrical and syntactical relations which constitute the roots of constituency. Here, constituency is more basic than combinatorial compositionality.

b. *A thesis concerning the necessary symbolicity of constituent structures and the reducibility of constituency to combinatorial compositionality.*

For the authors, representations can manifest internal constituent structures if and only if they are symbolic representations. Their internal

constituency is therefore a combinatorial one analogous to that found in formal languages. The “geometrical” form of constituency is therefore reduced to a pure combinatorial one.

Now, the point is that thesis (a) does not entail thesis (b). Actually, there exist many natural structures which are not symbolic but present nevertheless constituent structures. The most evident case is that of atoms and molecule in quantum mechanics : electronic orbitals provide a typical example of constituents which are not symbolic in nature and can be dynamically modeled as solutions of partial differential equations¹⁰. Therefore, if constituent structures are *natural* structures they don't need to be necessarily symbolic.

¹⁰Moreover, physical 'causality' is sensible to such physical constituent structures.

In the perceptive realm, another spectacular example is provided by the use of anisotropic diffusion equations in computational vision. They are able to carry out in a unified way two contradictory operations on images : smoothing and segmenting ! They yield very powerful bottom-up and data-driven algorithms which can achieve a morphological constituent analysis without any symbolic means.

If we admit thesis (a) without admitting thesis (b), the nature of the problem raised by the F-P arguments changes radically. It becomes to know whether the main concepts of relation, structure, constituency and compositionality can be mathematically interpreted in a purely dynamical framework. Is it possible to work out a “syntactic geometry” of structures in the CN context ?

III. Refutation of the main point of the F-P argument and the main problem

R1. The central F-P argument is valid only as far as it is applied to a *weak* CN. Remind that we call weak CN a CN which models uniformly semantically evaluable entities of different syntactic types by mathematical structures of the same type, without taking into account their differences of grammatical categories. On the contrary, we call strong CN a CN which has the capacity of modeling the differences and the relations between different grammatical categories.

R2. Let us clarify this a little bit further. The F-P argument points out a *category mistake*. Its “syllogism” is as follows :

• *Syllogism S1 :*

(i) Let A_i ($i = 1, \dots, n$) be the actants of a sentence and let V be the verb organizing the actantial interactions. Let us *model* the actants A_i by means

of mathematical structures A_i of a certain type (for example activity patterns) on which is defined an abelian (associative, commutative, with neutral element and inverse elements) operation of composition $\oplus$, that is an abelian group law (for example, the superposition of activity patterns).

(ii) Let us model the verb V by a structure V of the *same* type as that of the A_i .

(iii) Let us model the actantial interaction V of the A_i by the sum :

$$V \bigoplus_{i=1}^{i=n} A_i.$$

(iv) Experimental observation : such a modeling strategy runs into unavoidable difficulties.

(v) Conclusion : the modeling of the actants A_i by the structures A_i should be rejected since it is experimentally refutable. $\square$

Further, the authors oppose this syllogism to another one aimed at showing the superiority of CLC.

• *Syllogism S2 :*

(i) Let us *symbolize* the actants A_i by symbols A_i^* .

(ii) Let us symbolize the verb V by a symbol V^* .

(iii) Let us symbolize the actantial interaction V of the A_i by syntagmatic relations between the A_i^* and V^* (for example, by a syntagmatic tree produced by a generative grammar or a constituent-structure grammar).

(iv) Empirical observation : such a symbolization is "good".

(v) Conclusion : it has to be accepted since it is experimentally valid. $\square$

The problem is that the first syllogism (S1) is fallacious and the second (S2) tautological.

• *S1 is fallacious.*

In fact, it is equivalent to say that the (logico-combinatorial) structures of syntagmatic-tree type being non-associative and non-commutative (and therefore non-abelian), they cannot be modeled adequately by abelian algebraic structures of group type.

Let us mimic the F-P argument for another theory, for instance a physical theory (ultra-simple, so fictional) of elementary particles (e. p.).

(i) Let us model the free e. p.'s P_i ($i = 1, \dots, n$) by irreducible representations G_i of the Poincaré group in an Hilbert space.

(ii) Let us model the concept of an interaction V between the P_i by another irreducible representation F .

(iii) Let us model the interaction of the P_i by the sum

$$F \bigoplus_{i=1}^{i=n} G_i .$$

(iv) Empirical observation : such a modeling strategy runs into unavoidable difficulties and is experimentally refutable.

(v) Conclusion : the modeling of the P_i by the G_i must be rejected. $\square$

In such a case, the fallacy is striking. One has made a category mistake in confusing the concept of *interaction* in (ii) and (iii) with an additional free e. p. (iv) is trivial since an interaction of P_i is not the same thing as the system of the free P_i to which V has been added. The inference (iv) $\Rightarrow$ (v) is completely illegitimate.

It is the same case with the F-P argument. It also points out a category mistake : an interaction of actants is modeled by a mathematical structure V of the *same type* as those A_i which are used to model the actants themselves. Even if Fodor and Pylyshyn are then perfectly right in denouncing such an error in weak CN, nevertheless their drastic and dramatic conclusion is fallacious. The only correct conclusion is that, if actants are dynamically modeled by attractors, then verbs expressing interactions of actants cannot be modeled by attractors of the same type, but only by interactions of attractors.

• *S2 is tautological.*

It is clear that if one symbolizes constituents by means of formal symbols, then one can *a priori* symbolize their structural relations by means of formal relations. But, as we have already seen, such a formalization does not explain at all the relations.

R3. It is thus necessary to clarify and work out further F-P central argument. This can be done in the following way.

(i) First of all, we must be aware of *the distance that separates a true mathematical modeling from a mere formal symbolization*. Modeling a certain class of natural phenomena is to interpret them by sophisticated mathematical theories which allow to *reconstruct* their properties *mathematically* ; whereas to symbolize them, in contrast, is to *represent* such properties *formally*. The requirement of modeling has quite nothing to do with symbolization : for instance a mathematical physics of elementary particle interactions has nothing to do with symbolic representations of the type $V(P_i)$ where V is an n -ary relation. The main limitation of symbolic-formalist points of view in cognition is their confusing of a formal description with a mathematical explanation (see

[Petitot, 1982a, 1982b, 1985a, 1986b, 1986c]). We see here the dramatic consequences of the argument A9 criticized in C5 above.

(ii) Up to now, CNC constitutes the most decisive attempt to move from a formal symbolization to a mathematical modeling in cognitive sciences. And, as far as it aims to provide an explanation only for *proto-symbolic* structural phenomena, the lack of formalism cannot be imputed to it.

(iii) We can refute F-P arguments if we can answer positively the question whether, in the case of syntactic structures, there are *two* structural levels which do correspond respectively to the dynamical and symbolic ones. In a number of works, I have tried to show that such is effectively the case. Underlying the strictly grammatical level of grammatical relations (which are quite adequately described in terms of symbolic structures : syntagmatic trees, etc.) there does exist, in fact, the level of actantial relations where the actants are defined by their semantic (casual) roles and where the verbs express the actantial interactions. The differences between formal grammars and case grammars are well known. Now — it is the main result of these works — it is possible to work out a dynamical theory of actantial grammars.

(iv) But, we noticed that F-P arguments are *neutral* with respect to this difference of levels. They concern as well the actantial roles as the grammatical relations *and only refer to the "geometry" of structures* where "the geometric relations are themselves semantically significant". Thus we may apply the results of R2 regarding actantial syntax. Whence the question :

Main question : *If the actants A_i of a process are modeled by attractors A_i of a dynamical system, is it possible, within the framework of the mathematical theory of dynamical systems, to elaborate a theory of actantial interactions — that is a theory of the verb ?*

(v) Let us develop further this question. We have seen that, in many dynamical models, the situation can be greatly simplified if one makes the hypothesis that the dynamics X defining the attractors A_i admits a global Lyapunov function or, even more simply, that X is gradient : $X = -\text{grad } f$. The A_i are then the minima m_i of the potential function f .

Simplified main question : *If the actants A_i of a process are modeled by the minima m_i of a potential function, is it possible, within the framework of the dynamical theory of potential functions, to elaborate a theory of actantial interactions — that is a theory of the verb ?*

The mathematical challenge is therefore to develop a theory of *interactions of attractors*. What we call an *attractor syntax*. René Thom, Wolfgang Wildgen, Per Aage Brandt and ourselves have shown that bifurcation theory provides adequate tools for solving it.

IV. From connectionist binding to a configurational definition of semantic case-roles

Since 1988, many CN cognitivists have proposed strategies for taking up Fodor and Pylyshyn's challenge. One of the most interesting, but still insufficient, is Smolensky's idea of using the *tensorial product* operation (see [Smolensky, 1990] and [Visetti, 1990]).

1. Smolensky's tensorial product

Smolensky's main idea is to take for granted the CL finitist and combinatorial view of symbolic structures and to *represent* them in a CN way — in much the same way as one represents abstract groups in linear groups in the well known group representation theory. To do this, he first adopts a case conception of syntax and thinks of syntactic structures as composed of three sorts of entities :

- (i) semantic case-roles r_i ;
- (ii) fillers f_j ;
- (iii) binding relations between roles and fillers.

He supposes then that the roles and the fillers are *already* represented in a CN (local or distributed) way and solves the problem of representing *the binding relations* using the linear device of tensorial product.

Suppose that the roles r_i (resp. the fillers f_j) are vectors belonging to the vector space V_R (resp. V_F) of the global states of a network R (resp. F). Let u_p (resp. v_φ) be the units of R (resp. F). One connects R and F using connections $u_p \leftrightarrow v_\varphi$ with Hebbian weights $w_{p,\varphi} = \sum_i r_{i,p} \cdot f_{i,\varphi}$ where $r_{i,p}$ (resp. $f_{i,\varphi}$) is the activity of the unit u_p (resp. v_φ) in the global activity pattern of R (resp. F) representing r_i (resp. f_i). The tensorial product device consists in introducing *new* units $b_{p,\varphi}$ between R and F , $b_{p,\varphi}$ being connected by two weights = 1 to u_p and v_φ and having $w_{p,\varphi}$ as activity (see figure 1, on the opposite page).

It is easy to see that we get that way a CN implementation $R \otimes L$ of the tensorial product $V_R \otimes V_L$ with basis $b_{p,\varphi} = u_p \otimes v_\varphi$. With $w_{p,\varphi} = \sum_i r_{i,p} \cdot f_{i,\varphi}$, the state of $R \otimes L$ becomes :

$$\sum_{p,\varphi} w_{p,\varphi} b_{p,\varphi} = \sum_{p,\varphi} \left(\sum_i r_{i,p} \cdot f_{i,\varphi} \right) u_p \otimes v_\varphi = \sum_i \left[\left(\sum_p r_{i,p} u_p \right) \otimes \left(\sum_\varphi f_{i,\varphi} v_\varphi \right) \right] = \sum_i r_i \otimes f_i$$

We get therefore a *representation* $\Psi : S \rightarrow V$ of a set S of structures in the state space of a network, where a representation is defined in the following manner. We suppose that there exists a *role decomposition* F/R of S that is a truth function which assigns to each pair (f_j, r_i) the truth value

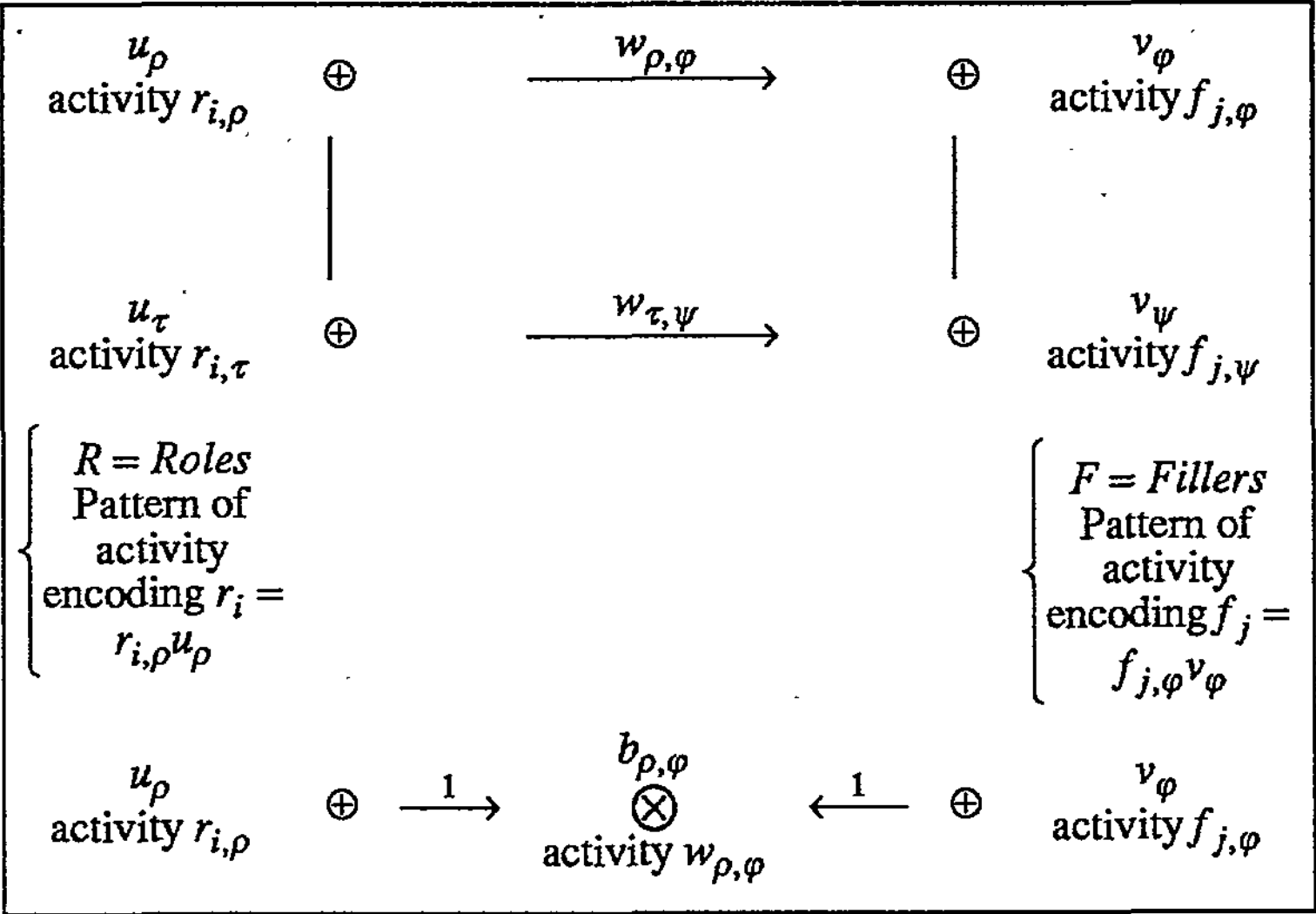


figure 1

of the predicate $f_j \mid r_i$ on S : “ f_j fills the role r_i in $s \in S$ ”. We suppose as given a CN representation of the fillers/roles bindings $\Psi_{bind} : \{ f_j \mid r_i \} \rightarrow V$ and we define the representation of the role decomposition F/R by the map :

$$\Psi : S \longrightarrow V$$
$$s \longrightarrow \sum_{\{(f_j, r_i) \mid (f_j \mid r_i)(s)\}} \Psi(f_j \mid r_i) .$$

In a tensorial product, r_i (resp. f_j) is identified with an activity pattern $r_{i,\rho} u_\rho$ (resp. $f_{j,\phi} v_\phi$) and $f_j \mid r_i$ is identified with $r_i \otimes f_j$. As far as a structure s is a conjunction of $f_i \mid r_i$ and a conjunction is represented by addition, we get finally :

$$\Psi(s) = \sum_i r_i \otimes f_i .$$

Paul Smolensky has shown in a detailed and convincing manner that this type of procedure allows to represent in various ways operations and transformations on symbolic structures. According to him, this shows that it is possible to integrate

“in an intimate collaboration, the discrete mathematics of symbolic computation and the continuous mathematics of connectionist computation”.

This view is clearly anti-eliminativist. Smolensky does not want to *reduce* all symbolic structures and processes to CN ones. He wants to *represent* in

a CN way these symbolic descriptions in order to explain "higher thought processes".

For instance, in a recent joint work with G. Legendre and Y. Miyata [Smolensky *et al*, 1992], he applies this strategy of understanding

"how symbolic computation can arise naturally as a higher-level virtual machine realized in appropriately designed lower level connectionist networks"

to binary trees. Let r_{x_i} be the *positional* roles in a binary tree with nodes x_i . A tree s with atom f_i at node x_i is represented by the tensorial product $s = \sum r_{x_i} \otimes f_i$. This representation is in fact a *recursive* one. The x can be coded by binary strings using the code 0 = "left child" and 1 = "right child". Let $r_{x_0} = r_x \otimes r_0$ and $r_{x_1} = r_x \otimes r_1$. Using such a binary coding, a tree s becomes represented by a vector of the vector space $\bigoplus_{k=1}^{k=\infty} V_R^{\otimes k} \otimes V_F$ where V_R is generated by r_0 and r_1 . Smolensky shows how to implement a programming language such as LISP in such a framework. He gives also a CN representation of Context Free Grammar theory.

2. Dynamical binding

With regards to the implementation of the binding relations between roles and fillers, we want also to mention the *dynamical binding* through synchronized oscillatory neural groups developed by Gerald Edelman, Christof von der Malsburg, Charles Gray, Wolf Singer, Elie Bienenstock, and Lokendra Shastri¹¹. The key idea is that the terms of sentences (objects) are internally represented and encoded by distributed arrays of rhythmic neural assemblies — oscillators — and the bindings by processes of *synchronization* — phase locking — between such oscillators. In this perspective :

"reasoning is the transient but systematic propagation of a *rhythmic* pattern of activation, where each *phase* in the rhythmic pattern corresponds to an object in the dynamic or short-term memory, where bindings are represented as the *in-phase* or *synchronous* firings of appropriate nodes, where long-term facts are subnetworks that act as temporal pattern matchers, and where rules are interconnection patterns that cause the propagation and transformation of rhythmic patterns of activation" [Shastri-Ajjanagadde, 1990, p. 14].

Using such oscillatory patterns and phase-locking processes, one becomes able to embody dynamical bindings in *the fine temporal structure* of firing patterns in the brain.

¹¹See e. g. [Edelman, 1987] and [Sporns *et al.*, 1991], [von der Malsburg & Bienenstock, 1986], [Bienenstock, 1992], [Shastri & Ajjanagadde, 1990]. Synchronization of oscillatory neural groups is also used for modeling constituency relations (what is called the "labeling hypothesis"). This rises very technical mathematical problems. See e.g. the works of Yoshiki Kuramoto, Hiroaki Daido, George Bard Ermentrout and Nancy Kopell. There is now strong evidence concerning the possibility of explaining that way the cognitive Gestalt laws (see e.g. [Engel *et al.*, 1992]).

This idea is akin to the works concerning the role of *fast* synapses in neural networks, that is the possibility for a neural network to change non-adiabatically its synaptic weights during its transient functioning.

We see that the way by which one can bind a role label with a filler term rises effectively some fundamental issues. But these are not central for taking up F-P challenge. Indeed, the main problem is that of the *configurational* definition of roles which can substitute for the classical role *labels*. In such a configurational definition, roles are identified with *positions* — places — in configurations of positions. Of course, they have to be filled by fillers, but the key difficulty is to elaborate an effective CN theory of such positional relations *without taking for granted any prior CL representation of them*.

As was strongly stressed in [Smolensky, 1990], CNC must

“find ways of naturally instantiating the *sources of power* of symbolic computation within fully connectionist systems”.

Such a CN instantiation is much more than a mere CN implementation of a symbolic stuff. It is an “extended version of connectionist computation”.

This is the core of the main problem because it is the “syntactic geometry” of the internal form of constituent structures which defines the semantic relations characterizing the roles. For solving it, it is not sufficient to bind noun fillers and case-frame slots of verbs, the roles being themselves tensorial product of semantic verb features and case-roles. We must be radical, and *instantiate* — and not only implement or represent — the concept of role in a purely dynamical way.

3. The core of the debate : the need for a configurational definition of the roles

In their response to Smolensky's response, Jerry Fodor and Brian McLaughlin [1990] reconsider the systematicity problem and the fact that :

“cognitive processes are causally sensitive to the constituent structure of mental representations” (p. 185).

They summarize their main point claiming that :

“all we really need is that propositions have internal structure, and that characteristically, the internal structure of complex mental representations corresponds, in the appropriate way, to the internal structure of the propositions that they express” (p. 187).

More precisely, they introduce a condition (C) which “expresses a *psychological law* that subsumes all systematic minds” :

(C) "If a proposition P can be expressed in a system of mental representations M , then M contains some complex mental representation (a "mental sentence") S , such that S expresses P and the (classical) constituents of S express (or refer to) the element of P " (p. 187).

Condition (C) plus the fact

"that mental processes have access to constituent-structure of mental representations"

allows to explain the cognitive systematicity of mind.

Against this theoretical background, Fodor and McLaughlin can provide an evaluation of Smolensky's tensorial product device. Their main criticism is that it is impossible to retrieve from tensorial product representations and from abelian superposition operations a constituent-structure whose constituents are endowed with a *causal* status. Indeed, in a vector space the choice of a basis and hence of a vector decomposition is *not canonical*. Every vector decomposition is therefore *counterfactual* and the constituents (components) it generates *cannot* have causal efficiency.

We think that this negative argument is essentially right even if it is over-drastring. For instance, it is true that there exists no canonical basis in a vector space V (that is, V has a non-trivial symmetry group, the linear group $GL(V)$). But, nevertheless, the vector space V_R of the states of a network R does possess a distinguished basis, namely that defined by its units. In that case, vector decompositions are not counterfactual operations. But notwithstanding, F-M criticism points out a major difficulty which can be expressed in the following manner.

We have seen that for Smolensky, the basic problem of a CN theory of symbolic structures is that of the *binding relations* between roles and fillers. He succeeded in solving this problem, but in a way which replies to only *half* of F-P challenge. Indeed, it says nothing about the possibility of reaching in the CN framework a configurational definition of actantial roles. Moreover, it takes for granted a symbolic pre-definition of the roles. As was stressed by Yves Marie Visetti, in the tensorial product approach "the associative conception of memory as a relaxation to a preferential state" together with "the concept of attractor as an intrinsic meaningful state" disappear [Visetti, 1990, p. 186].

Now, we have just seen that the problem is not only to represent semantic roles as local or distributed activity patterns of some appropriate network. It is also to give a correct CN account of the relations of actantial interaction which are involved in syntactic structures. *These relations are not binding relations*. They concern the roles independently of their fillers. The PTC agenda which, according to Smolensky,

"consists in taking [the] cognitive principles and finding new ways to instantiate them in formal principles based on the mathematics of dynamical systems"¹².

¹²See Smolensky's contribution to [COMPCOG I, 1991].

must also be applied to the configurational definition of the actantial roles.

In some sense, it is such a requirement which is stressed by Fodor and McLaughlin [1990] when they claim that in order to build up a CN theory of constituency and systematicity one must :

- (i) find "some property D, such that if a dynamical system has D its behavior is systematic" ;
- (ii) "make clear what property D is" ; and
- (iii) "show that D is a property that CN systems can have by law" (p. 201).

But suddenly, the authors become dogmatic. They state that such a requirement is impossible to satisfy. Even when Smolensky stresses the fact that constituent structures do exist in physics (for instance a molecule with its atomic nucleus and its electrons), they reply that

"since *being a representation* isn't a property in the domain of physical theory, the question whether mental representations have constituent-structure has no analog in physics" (p. 200).

Of course, it is then very easy for them to conclude

"that Fodor and Pylyshyn's challenge to connectionists has yet to be met. We still don't have *even a suggestion* of how to account for systematicity within the assumptions of connectionist cognitive architecture" (p. 204).

This claim resumes some previous ones in the F-P 1988 paper :

"so far as we know, there are no worked out attempts in the Connectionist literature to deal with the syntactic and semantical issues raised by relations of real-constituency" (p. 22) ;

"There doesn't seem to be any other way to get the force of structured symbols in a Connectionist architecture. Or, if there is, nobody has given any indication of how to do it" (p. 24) ;

"There are no serious proposals for incorporating syntactic structure in Connectionist architectures" (p. 67).

However, these peremptory judgements are not at all true. Indeed, let us recall again that the concept of *bifurcation* of attractors provides an *effective* property D, and that it permits therefore to work out a CN

attractor syntax and a dynamic theory of constituency. It is a rather technical stuff. But this is another story. Our purpose here was only to discuss the epistemological issues of dynamical constituency.

EHESS-CREA (Paris)
e-mail: petitot@poly.polytechnique.fr

Bibliography

ALVAREZ (L.) & LIONS (P. L.) & MOREL (J. M.)
1992, "Image Selective Smoothing and Edge Detection by Non-Linear Diffusion", *SIAM J. Numer. Anal.*, 29, p. 845-866.

AMIT (D.)
1989, *Modeling Brain Function*, Cambridge University Press.

ANDERSON (J. M.)
1971, *The Grammar of Case, towards a Localistic Theory*, Cambridge University Press.

ANDLER (D.)
1990, "Connexionnisme et Cognition : à la recherche des bonnes questions", *RS* 1990, p. 95-127.

ANDLER (D.), ed.
1992, *Introduction aux sciences cognitives*, Paris, Gallimard.

ANDLER (D.) & PETITOT (J.) & VISETTI (Y. M.)
1991, "Dynamical Systems, Connectionism and Linguistics", *COMPCOG I* [1991].

AZENCOTT (R.)
1988, "Simulated Annealing", *Séminaire Bourbaki*, 697, Astérisque 161-162, Paris.

BIENENSTOCK (E.)
1992, "Suggestions for a Neurobiological Approach to Syntax", *COMPCOG II* [1992].

BRANDT (P. A.)
1986, *La Charpente modale du Sens*, Doctoral Thesis, University of Paris III.

COMPCOG I
1991, *Interdisciplinary Workshop on Compositionality in Cognition and Neural Networks*, D. Andler, E. Bienenstock, B. Laks eds., Abbaye de Royaumont.

COMPCOG II
1992, *Interdisciplinary Workshop on Compositionality in Cognition and Neural Networks*, D. Andler, E. Bienenstock, B. Laks eds., Abbaye de Royaumont.

DESCLÉS (J.- P.)
1990, *Langages applicatifs, langues naturelles et cognition*, Paris, Hermès.

[Doyon et al.]
DOYON (B.) & CESSAC (B.) & QUOY (M.) & SAMUELIDES (M.)
1993, "Chaos in Neural Networks with Random Connectivity", *International Journal of Bifurcation and Chaos* (forthcoming).

EDELMAN (G. M.)
1987, *Neural Darwinism : The Theory of Neuronal Group Selection*, New York, Basic Books.

[EMG]
1992, *Emergence dans les modèles de la cognition*, Paris, Telecom.

[Engel *et al.*]

ENGEL (A.) & KÖNIG (P.) & GRAY (C.) & SINGER (W.)

1992, "Temporal Coding by Coherent Oscillations as a Potential Solution to the Binding Problem : Physiological Evidence", *Non-Linear Dynamics and Neural Networks*, H. Schuster ed., Berlin, Springer.

FILLMORE (Ch. J.)

1977, "The Case for Case Reopened", p. 59-81, in *Grammatical Relations*, P. Cole and J. M. Sadock eds., *Syntax and Semantics*, vol. 8, New York, Academic Press.

FODOR (J. A.) & McLAUGHLIN (B. P.)

1990, "Connectionism and the Problem of Systematicity : Why Smolensky's Solution Doesn't Work", *Cognition*, 35, p. 183-204.

FODOR (J. A.) & PYLYSHYN (Z. W.)

1988, "Connectionism and Cognitive Architecture : A Critical Analysis", *Cognition*, 28, 1-2, p. 3-71.

HARRIS (C.)

1990, "Connectionism and Cognitive Linguistics", *Connection Science*, 2, 1-2, p. 7-33.

HJELMSLEV (L.)

1972 (1935), *La Catégorie des cas*, München, Wilhelm Fink Verlag.

HOLENSTEIN (E.)

1992, "Phenomenological Structuralism and Cognitive Semiotics", R. Benatti ed., *Scripta Semiotica*, 1, p. 133-15.

JACKENDOFF (R.)

1983, *Semantics and Cognition*, Cambridge (MA), MIT Press.

1987, *Consciousness and the Computational Mind*, Cambridge (MA), MIT Press.

KITCHER (P.)

1988, "Marr's Computational Theory of Vision," *Philosophy of Science*, 55, p. 1-24.

KOPELL (N.) & ERMENTROUT (G. B.)

1990, "Phase Transitions and Other Phenomena in Chains of Coupled Oscillators", *SIAM J. Appl. Math.*, 50, 4, p. 1014-1052.

KURAMOTO (Y.) & NISHIKAWA (I.)

1987, "Statistical Macrodynamics of Large Dynamical Systems. Case of a Phase Transition in Oscillator Communities", *Journal of Statistical Physics*, 49, 3-4, p. 569-605.

LAKOFF (G.)

1988, "A Suggestion for a Linguistics with Connectionist Foundations", *Proceedings of the 1988 Connectionist Models Summer School*, M. Kaufman.

LANGACKER (R. W.)

1987, *Foundations of Cognitive Grammar*, vol. 1, Stanford University Press.

1991, *Foundations of Cognitive Grammar*, vol. 2, Stanford University Press.

[LTC]

1989, *Logos et Théorie des Catastrophes* (the René Thom's Cerisy Conference, J. Petitot ed.), Genève, Éditions Patino.

MARR (D.)

1982, *Vision*, San Francisco, Freeman.

[Mézard *et al.*]

MEZARD (M.) & PARISI (G.) & VIRASORO (M. A.)

1987, *Spin Glass Theory and Beyond*, World Scientific Lecture Notes in Physics, vol. 9, Singapore.

PARISI (D.)

1991, "Emerging Compositionality in Neural Nets", *COMPCOG I*, p. 90-98.

PETTITOT (J.)

1979, "Hypothèse localiste et théorie des catastrophes", *Théories du langage, Théories de l'apprentissage*, M. Piatelli éd., Paris, Le Seuil.

1985, *Morphogenèse du Sens*, Paris, PUF.

1986, "Structure", *Encyclopedic Dictionary of Semiotics*, Th. Sebeok ed., vol. 2, p. 991-1022, New York, de Gruyter.

1989a, "On the Linguistic Import of Catastrophe Theory", *Semiotica*, 74, 3-4, p. 179-209.

1989b, "Hypothèse localiste. Modèles morphodynamiques et Théories cognitives : remarques sur une note de 1975", *Semiotica*, 77, 1-3, p. 65-119.

1989c, "Catastrophe Theory and Semio-Narrative Structures", *Paris School of Semiotics*, P. Perron, F. Collins eds., p. 177-212, Amsterdam, J. Benjamins.

1989d, "Modèles morphodynamiques pour la grammaire cognitive et la sémiotique modale", *RSSI* (Canadian Semiotic Association), 9, 1-2-3, p. 17-51.

1990, "Le Physique, le morphologique, le symbolique. Remarques sur la vision", *Revue de Synthèse*, 1-2, p. 139-183.

1991a, "Why Connectionism is such a Good Thing. A Criticism of Fodor's and Pylyshyn's Criticism of Smolensky", *Philosophica*, 47, 1, p. 49-79.

1991b, "Syntaxe topologique et grammaire cognitive", *Langages*, 103, p. 97-128.

1991c, "Le Schématisme morphodynamique de l'aspectualité", *Le discours aspectualisé*, J. Fontanille éd., p. 177-193, Amsterdam, J. Benjamins.

1992, *Physique du Sens*, Paris, Éditions du CNRS.

1993a, "Modeling : Formalization or Mathematization ? The Example of a Morphodynamical Approach to Language", *Structures of Signification*, vol. III, H.S. Gill ed., New Delhi, Wiley Eastern, p. 700-709.

1993b, "Natural Dynamical Models for Cognitive Grammars", *Language and Cognition*, F. Manjali ed., New Delhi, Bahri Publications.

1994, "Attractor Syntax : Morphodynamics and Cognitive Grammars", *Continuity in Linguistic Semantics*, C. Fuchs and B. Victorri eds., p. 167-187, Amsterdam, J. Benjamins.

forthcoming, 1995, "Morphodynamics and Attractor Syntax. Dynamical and Morphological Models for Constituency in Visual Perception and Cognitive Grammar", *Mind as Motion*, T. van Gelder, R. Port eds., Cambridge (MA), MIT Press.

PETTITOT (J.) & SMITH (B.)

1991, "New Foundations for Qualitative Physics", *Evolving Knowledge in Natural Science and Artificial Intelligence*, J. E. Tiles, G. J. McKee, G. C. Dean eds., p. 231-249, London, Pitman.

PIATELLI-PALMARINI (M.)

1988, "Sélection sémantique et sélection naturelle : le rôle causal du lexique", *RS* [1990], p. 57-94.

PYLYSHYN (Z. W.)

1984, *Computation and Cognition : Toward a Foundation for Cognitive Science*, Cambridge (MA), MIT Press.

REGIER (T.)

1988, "Recognizing Image-Schemas Using Programmable Networks", *Proceedings of the 1988 Connectionist Models Summer School*, M. Kaufman.

RENALS (S.) & ROHWER (R.)

1990, "A Study of Network Dynamics", *Journal of Statistical Physics*, 58, 5-6, p. 825-848.

RS, 1990, *Sciences cognitives : quelques aspects problématiques*, J. Petitot éd., *Revue de Synthèse*, IV, 1-2.

SHASTRI (L.) & AJJANAGADDE (V.)

1990, *From Simple Associations to Systematic Reasoning : A Connectionist Representation of Rules, Variables and Dynamic Bindings using Temporal Synchrony*, Technical Report, Computer and Information Science Department, University of Pennsylvania.

SIMON (H.)

1994, "Literary Criticism : A Cognitive Approach", *Bridging the Gap*, S. Franchi, G. Güzeldere eds., *Stanford Humanities Review*, 4, 1 (Supplement).

SMITH (B.)

1993, "The Structures of the Common-Sense World", *Gestalt Theory. Its Origins, Foundations and Influence*, A. Pagnini, S. Poggi eds., Florence, Olschky.

SMOLENSKY (P.)

1988, "On the Proper Treatment of Connectionism", *The Behavioral and Brain Sciences*, 1, n°11, p. 1-23.

1990, "Tensor Product Variable Binding and the Representation of Symbolic Structures in Connectionist Networks", *Artificial Intelligence*, n°46, p. 159-216.

[Smolensky *et al.*]

SMOLENSKY (P.) & LEGENDRE (G.) & MIYATA (Y.)

1992, *Principles for an Integrated Connectionist/Symbolic Theory of Higher Cognition*, Technical Report, University of Colorado at Boulder, Department of Computer Science.

[Sompolinsky *et al.*]

SOMPOLINSKY (H.) & CRISANTI (A.) & SOMMERS (H.-J.)

1988, "Chaos in Random Neural Networks", *Phys. Rev. Lett.*, 61, p. 259-262.

[Sporns *et al.*]

SPORNS (O.) & TONONI (G.) & EDELMAN (G. M.)

1991, "Modeling Perceptual Grouping and Figure-ground Segregation by means of Active Reentrant Connections", *Proc. Natl. Acad. Sci. USA*, 88, p. 129-133.

TALMY (L.)

1978, "Relation of Grammar to Cognition", *Proceedings of TINLAP-2*, D. Waltz éd., Urbana, University of Illinois.

1983, "How Language Structures Space", *Spatial Orientation : Theory, Research, and Application*, H. Pick and L. Acredolo eds., Plenum Press.

1985, "Force Dynamics in Language and Thought", *Parasession on Causatives and Agentivity*, Chicago Linguistic Society (21 st. Regional Meeting).

1990, "Fictive Motion in Language and Perception", Workshop *Motivation in Language*. University of San Marino : International Center for Semiotic and Cognitive Studies.

THOM (R.)

1972, *Stabilité structurelle et Morphogenèse*, New York-Paris, Benjamin-Ediscience.

1980, *Modèles mathématiques de la Morphogenèse*, Paris, Christian Bourgois.

1988, *Esquisse d'une Sémiophysique : physique aristotélicienne et théorie des catastrophes*, Paris, InterÉditions.

TIROZZI (B.) & TSODKYS (M.)

1991, "Chaos in Highly Diluted Neural Networks", *Europhy. Lett.*, 14, p. 727-732.

Van GELDER (T.)

1992, "Connectionism and Dynamical Explanation", *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, Hillsdale, L. Erlbaum.

VISETTI (Y. M.)

1990, "Modèles connexionnistes et représentations structurées", *Modèles Connexionnistes*, D. Memmi, Y. M. Visetti eds., *Intellectica*, n°s 9-10, p. 167-212.

Von der MALSBURG (Ch.) & BIENENSTOCK (E.)

1986, "Statistical Coding and Short-Term Synaptic Plasticity : A Scheme for Knowledge Representation in the Brain", *Disordered Systems and Biological Organization*, (E. Bienenstock, F. Fogelman, G. Weisbuch eds.), p. 247-272. Berlin, Springer.

WILDGEN (W.)

1982, *Catastrophe Theoretic Semantics*, Amsterdam, Benjamins.

ZEEMAN (C.)

1965, "Topology of the Brain", *Mathematics and Computer Science in Biology and Medicine* (Medical Research Council), p. 240-256.

1976, "Brain Modelling", *Structural Stability, the Theory of Catastrophes and Applications in the Sciences*, Lecture Notes in Mathematics, 525, p. 367-372. Berlin, Springer.

1977, *Catastrophe Theory : Selected Papers 1972-1977*, Redwood City, Addison-Wesley.

