

# 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 effet<sup>1</sup> :

- 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 continu<sup>2</sup> ;

- 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" ;

<sup>1</sup>Nous 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].

<sup>2</sup>Il 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 part<sup>3</sup>.

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

<sup>3</sup>C'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 autre<sup>4</sup> ;

- 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 *systématicité* 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) ?

<sup>4</sup>On 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 systématicité 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 systématicité 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*<sup>5</sup>.

<sup>5</sup>On 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 lesinstancient 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*<sup>6</sup>, 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.

<sup>6</sup>Les 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.

<sup>7</sup>Elle 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 suivante<sup>7</sup> :

— 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

<sup>8</sup>Smolensky 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.

<sup>9</sup>En 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'avons 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 différents<sup>8</sup>, 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 recouplement, 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)<sup>9</sup>.

- 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 *abstraites*, 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.

<sup>10</sup>Ces 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 schématisation 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 ?

<sup>11</sup>Il 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 schématisation 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 généralisée<sup>10</sup>.

• 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 générale<sup>11</sup>. 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 computationnelle<sup>12</sup>. 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.*

<sup>12</sup>On 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. Bourguin, 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 :

<sup>13</sup>Ce 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 propriétés<sup>13</sup>.

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*

<sup>14</sup>Bien 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'autonomie<sup>14</sup>.

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 empirique<sup>15</sup>, 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].*

<sup>15</sup>*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<sup>e</sup> 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<sup>e</sup> 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-Soulie 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<sup>e</sup> 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.