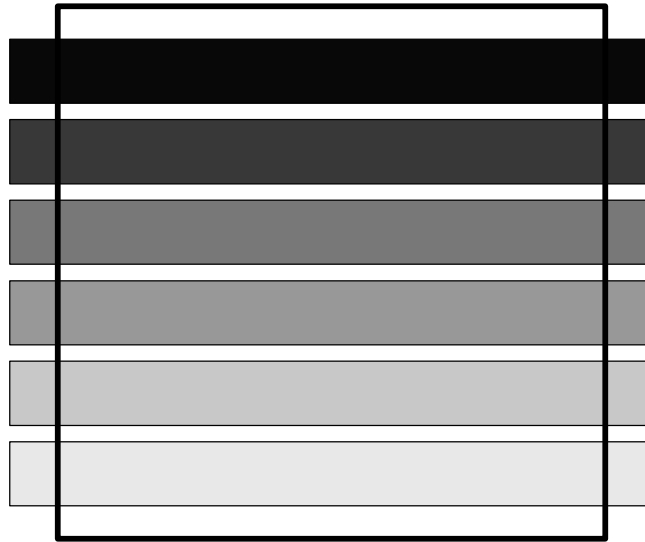


LES CAHIERS DU LANCI



TROIS TYPES DE REPRÉSENTATIONS COGNITIVES

Jean-Guy Meunier

N° 2002-02

UQÀM
Université du Québec à Montréal

Le Laboratoire d'ANalyse Cognitive de l'Information (LANCI) effectue des recherches sur le traitement cognitif de l'information. La recherche fondamentale porte sur les multiples conceptions de l'information. Elle s'intéresse plus particulièrement aux modèles cognitifs de la classification et de la catégorisation, tant dans une perspective symbolique que connexionniste.

La recherche appliquée explore les technologies informatiques qui manipulent l'information. Le territoire privilégié est celui du texte.

La recherche est de nature interdisciplinaire. Elle en appelle à la philosophie, à l'informatique, à la linguistique et à la psychologie.

Publication du Laboratoire d'ANalyse Cognitive de l'Information

Directeur : Jean-Guy Meunier

Université du Québec à Montréal

Volume 3, Numéro 2002-02 – Juillet 2002

Document disponible en ligne à l'adresse suivante : www.unites.uqam.ca/lanci/

Tirage : 15 exemplaires

Aucune partie de cette publication ne peut être conservée dans un système de recherche documentaire, traduite ou reproduite sous quelque forme que ce soit - imprimé, procédé photomécanique, microfilm, microfiche ou tout autre moyen - sans la permission écrite de l'éditeur. Tous droits réservés pour tous pays. / All rights reserved. No part of this publication covered by the copyrights hereon may be reproduced or used in any form or by any means - graphic, electronic or mechanical - without the prior written permission of the publisher.

Dépôt legal – Bibliothèque Nationale du Canada

Dépôt legal – Bibliothèque Nationale du Québec

ISBN : 2-922916-02-2

© 2002 Jean-Guy Meunier

Mise en page : L. Quevillon et D. Forest

TROIS TYPES DE REPRÉSENTATIONS COGNITIVES¹

Jean-Guy Meunier
Université du Québec à Montréal

I. L'objet d'étude des sciences cognitives : la représentation

Lorsque, dans son environnement, un organisme vivant se retrouve devant une pomme, il peut entamer deux actions distinctes. Dans un premier cas, il peut tout simplement la manger, c'est-à-dire l'ingérer physiquement. Les modèles biochimiques de la nutrition qui auront à décrire cette action proposeront divers modèles de transformations physiques de cette pomme au sein de l'appareil digestif de l'organisme. Dans un deuxième cas, l'organisme peut aussi entrer dans une relation cognitive avec cette pomme. Les modèles qui décriront cette relation diront tout naturellement que la pomme n'est plus alors manipulée dans sa structure physique originale, mais que l'agent l'a remplacée par un quelconque substitut. Dans une description *généreuse* de ce substitut, on dira que la pomme est « représentée » dans l'agent, c'est-à-dire que l'agent cognitif rend *présent à nouveau* (i.e. re-présente) en lui ce qui était présent à l'extérieur de lui, mais sous une autre « forme ».

Dans les deux cas, il y a une manipulation de formes, car il y aura création d'une forme originale qui consistera essentiellement en la création d'un *état interne* propre à l'agent. Dans le premier cas, on dira que l'agent est *transformé* physiquement. Dans le second, on dira qu'il est *informé*. Dans les deux cas, l'état interne entretiendra une relation spécifique avec son origine. Dans le premier cas, on soutient que la forme interne est *causalement* liée à cette origine. Mais dans le deuxième cas, on pose que la forme interne est *sémantiquement* liée à son objet externe. C'est-à-dire qu'elle « renvoie » à cet objet. En termes classiques, on dira que cet état interne représente son objet externe. Ainsi, ce concept représentation, malgré toute sa généralité, se retrouvera-t-il au cœur de presque tous les modèles de la cognition du plus métaphysique au plus naturaliste qui y sont confrontés ?

¹ Je remercie mon collègue Pierre Poirier ainsi que mes étudiants pour les nombreuses discussions qui ont pu faire germer les présentes réflexions.

« Representation is one of the most central concepts in cognitive science: there is no cognition without representation, and no cognitive science either. » (Billman, 1998, p.658)

Ainsi, ce concept de représentation sert à désigner de manière très générale ce simple fait : lorsqu'un agent est mis en relation cognitive avec son environnement, il construit un *état interne* qui lui rend présent cet environnement ou, à l'inverse, qui le « renvoie » à cet environnement.

Cette relation est, en effet, le cœur du concept de représentation. On la retrouve dans presque toutes les théories occidentales de la connaissance. Depuis Aristote, en passant par Augustin, Descartes, Kant et Hegel, la *cognition* a toujours été vue comme un acte de « re-présentatio », c'est-à-dire de présence nouvelle d'un objet interne dans l'agent. Cette re-présentatio est souvent traduite en termes modernes plus techniques d'*intentionnalité* (Brentano, 1874), de « *noeme* » (Husserl, 1959), de « *Umwelt* » (Von Uexküll, 1940), ou tout simplement de *signes* (Peirce, 1931; Eco, 1976), c'est-à-dire de quelque chose qui « *se tient à la place de ...* ».

On retrouve ce même type de relation dans les théories contemporaines de la cognition. Par exemple, l'IA parlera de représentation symbolique : Marr (1982), Newell (1980) Pylyshyn (1984) en feront un concept-clef de leur hypothèse symbolique :

« Representation is simply another term to refer to a structure that designates. » (Newell, 1980 :176)

« There is the notion that symbols actually have symbolic character. What this means is that symbols give access to what they designate. » (Newell et Simon, 1986, p. 34)

Il en est de même en psychologie et en philosophie cognitive (Johnson Laird, 1988; Chalmers, 1991; Brunner, 1990; Denis, 1993; Haugeland, 1986; Dennett, 1978; Putnam, 1988; Cummins, 1989; Pollock, 1989 Stich, 1983; Bermudez, 1996) et même en linguistique (Chomsky, 1980; Jackendoff, 1987; Desclées, 1990).

« Psychology is the science that investigates the representation and processing of information by complex organisms. » (Holyoak, 1991, p. 862)

En centrant l'attention sur le niveau fonctionnel du comportement cognitif par opposition au niveau matériel et computationnel, ces auteurs identifieront cependant, souvent, le concept de *représentation* à celui de *symbole* : de fait, ils rendront presque identique le traitement de l'information, avec la manipulation des représentations et, donc, des symboles.

« We are dealing with information processing machines, and the way such machines work is by using symbols to stand for things to represent things. » (Marr, 1982, p. 21)

Ce modèle dit représentationnel symbolique sera contesté par des modèles représentationnels dits non symboliques. Par exemple, les modèles connexionnistes parleront de sous-symboliques. Mais, malgré ce refus du terme symbolique, ils verront toujours le processus cognitif comme mettant en jeu une relation entre une certaine entité ou état interne et des entités externes, c'est-à-dire relevant de l'environnement.

« The subsymbolic characterization of a cognitive system intrinsically binds cognitive systems both to states of the environment and to goal conditions. » (Smolensky, 1988, p. 15)

Par ailleurs, certains, encore plus radicaux, refuseront le concept de représentation lui-même. En effet, plusieurs chercheurs (Brooks, 1991; Van Gelder, 1993; Van Gelder et Port, 1995; Scott-Kelso, 1995; Franklin, 1995; Globus, 1992; Moravec, 1988; Varela, 1988; Dreyfus, 1988) soutiendront que ce concept de représentation est trop général :

« It is the concept of representation which is insufficiently sophisticated. (van Gelder, 1993, p. 6)

« We are not building representations at all ! » (Thelen & Smith, 1994, p. 338)

Pour ces auteurs, ce concept de *représentation* sera tout simplement la mauvaise catégorie explicative pour le comportement cognitif. Ainsi, il serait à rejeter. Et on le remplacerait par d'autres, plus particulièrement par celui d'information ou, comme le dit Smolensky (1987), par des « grappes d'informations qui supportent des inférences dans des situations prototypiques ». L'information n'apparaîtrait plus alors comme une représentation, c'est-à-dire un symbole, mais comme des *dispositions* du corps (Freeman et Skarda 1987) des *réponses* à des stimuli², des *affordances* (Gibson, 1950), des *actions situées* (Clark, 1997), des *perturbations*, (Varela, 1988) des schémas *morphologiques* orientés (Thom 1988; Petitot, 1985).

Ce débat entre les approches représentationnelle, symbolique, non symbolique et non représentationnelle est très actif en sciences cognitives. Il n'en demeure pas moins, cependant, que lorsque vient le temps de décrire l'originalité de la relation cognitive, on retrouve toujours un même grand schéma de base. L'acte cognitif d'un agent implique toujours la création d'un état

² Cette position reprend les positions behaviorales classiques qui traduisent « états internes » en termes de « disposition à... ». Voir Quine 1960.

interne qui entretient une relation de dépendance avec les objets dans son environnement. Dans certaines théories de l'esprit, on verra cet état interne comme une structure symbolique. En intelligence artificielle, on le verra comme un système informationnel. Et dans les modèles dynamiques, on le verra comme un encodage distribué et/ou paramétré. Ce schéma reprend le cœur du concept de représentation, peu importe le nom qu'on lui donne.

II. Trois types de représentation

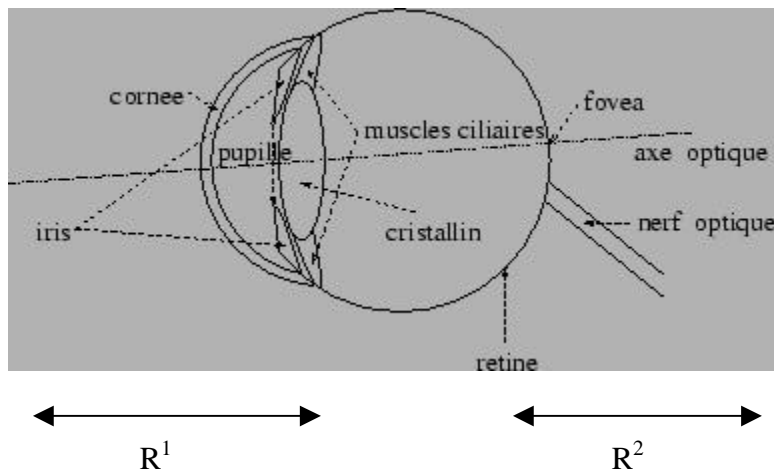
De quelle nature est cette relation entre un état interne et son objet externe? La réponse à cette question est complexe; elle n'est pas unique et elle dépend essentiellement non de l'expérience que nous en avons, mais de notre capacité à décrire ou à expliciter les opérations qui sont à l'œuvre dans les états internes créées, c'est-à-dire de la représentation. Trop souvent, cependant, le débat sur cette question se fonde sur des compréhensions, à notre avis, trop englobantes des opérations en jeu. Représenter devient synonyme de « rendre présent » à l'agent son environnement. Pourtant, comme nous le verrons ci-après, le concept de représentation désignera très souvent, dans les diverses théories cognitives, des structures de relations souvent très différentes. Nous montrerons qu'il opère dans au moins trois espaces théoriques qui construisent chacun des structures de relations radicalement différentes et mettent en jeu des opérations très spécifiques. Ainsi, distinguerons-nous trois types de concepts de représentation, à savoir la représentation comme *inscription* de l'information, la représentation comme *classification* de l'information et la représentation comme *catégorisation de l'information*. Ces trois espaces théoriques fondent des modèles radicalement différents de la représentation.

2.1. Type I : La représentation comme inscription

Dans la littérature technique des sciences cognitives, le premier espace théorique où le concept de *représentation* apparaît est celui où on explique comment un état interne est créé lorsque les organes de captation d'un agent -par exemple: l'œil, l'oreille, le nez, etc. -sont affectés par un objet externe (distal) ou interne (proximal). Ces organes constituent ce que traditionnellement on appelle les « sens ». Lorsque stimulés, ils produisent un état physique interne. Nous ne parlons pas ici de la « sensation » c'est-à-dire de la perception consciente que l'agent peut avoir de la réaction de ces sens, mais uniquement de la trace physique qui est construite dans chacun des capteurs. Bien que les opérations effectuées dans ces capteurs soient

décrites strictement en termes physiques, on parlera souvent de cet état interne en termes de *représentation*. Nous appellerons ce type (I) de représentation *inscription*.

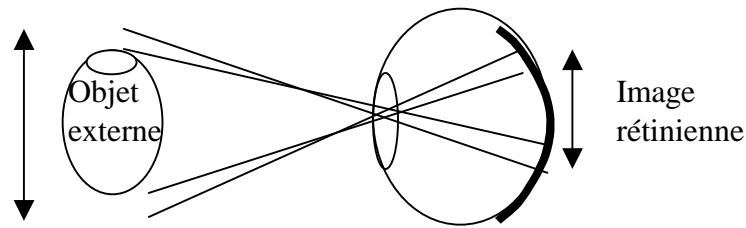
Pour illustrer ce type de représentation, prenons l'exemple de l'œil. Cet organe est une structure physiologique complexe composée entre autres de la cornée, de l'iris, de la pupille, du cristallin et de la rétine. Sur cette dernière se trouve une couche de cellules nerveuses: les cônes et les bâtonnets qui sont à proprement parler les cellules « réceptrices ». Viennent ensuite des ganglions et le nerf optique qui se rend au cerveau. En voici un bref schéma :



Le phénomène de la vision résulte de deux types de parcours radicalement différents. Un premier, noté ici R^1 , va de l'objet externe à la rétine et un second, R^2 , de la rétine au cerveau. C'est le premier qui retient ici notre attention. Dans ce premier parcours, un objet externe (stimulus distal), « produit » (i. e. reflète) des rayons lumineux, c'est-à-dire des photons, qui traversent la cornée, la pupille, le cristallin, et se projettent sur la rétine, excitant chacun des cônes et des bâtonnets qui s'y trouvent. Ce parcours est strictement physique.

Bien qu'on ne reconnaisse pas à ce processus une dimension cognitive à proprement parler, ce parcours n'en demeure pas moins un moment essentiel de la cognition. Et son résultat est souvent traduit en termes de représentation. On dira facilement (cf. Sloman, 1983; Marr, 1982) que la rétine imprime une « image » de l'objet externe ou encore qu'il y a sur la rétine une « représentation » de l'objet externe. Cette représentation est une con-« figuration » particulière de l'excitation des récepteurs sensibles que sont les cônes et les bâtonnets. Cette « représentation » est évidemment une métaphore qui est inspirée de la photographie. La rétine est comme une pellicule ou une matrice de pixels qui reçoit une « image » de l'extérieur. Elle

reproduit certaines caractéristiques de l'objet, telles que sa couleur, ses rapports de proportionnalité, son intensité lumineuse, etc.



Cette image rétinienne n'est en fait que la distribution du flux lumineux sur l'ensemble des capteurs de la rétine (dans une caméra numérique, on parlerait des pixels de l'écran sensible interne). Cette « image » est simple. Elle est un état interne provoqué par l'impact lumineux d'un objet externe sur chaque capteur. On dira souvent cependant qu'elle constitue une « description » de l'objet externe parce que la réaction des multiples capteurs forme un « pattern » reproduisant celui de l'objet externe (ici : un « pattern » lumineux). D'où sa traduction en termes représentationnels.

Ce modèle de représentation que nous avons appelé ici *inscription*, n'est pas toujours reconnue par certains cognitivistes (Haugeland, 1987) comme une « vraie représentation ». Elle ne serait qu'une réaction physique et non une représentation. Nous pensons au contraire que, bien que l'inscription ne soit qu'un processus physique cela n'empêche pas qu'elle soit une forme de représentation, quoique limitée et élémentaire. Qui plus est, elle est très souvent le prototype de la conception de ce qu'est une représentation. Et si nous mettons en évidence ce type de représentation ici, c'est qu'il se retrouve aussi implicitement au cœur de plusieurs théories de la représentation. Il présentera des caractéristiques très spécifiques que notre analyse ultérieure tentera de mettre en évidence.

2.2. Type II : La représentation comme classification

Le deuxième espace théorique en sciences cognitives en appelle au concept de représentation pour expliquer les opérations qu'un agent doit effectuer pour traiter la multitude d'informations différentes qu'il reçoit de ses organes de captation. En effet, un agent ne peut pas simplement enregistrer et archiver une à une les inscriptions reçues. Même si les 10^{14} connexions synaptiques (Ellman, 1999; Churchland, 1989) qui peuvent prendre au moins une multitude de valeurs et qui sont contrôlées par 10^6 gènes pouvaient archiver la myriade d'inscriptions qu'un organisme vivant peut accumuler pendant sa vie, il faudrait expliquer comment cet organisme réussit en un

temps relativement court à les retrouver, les comparer et les relier à des actions. Autrement dit, il faut montrer avant tout comment le cerveau regroupe ces inscriptions multiples et différentes en des formes synthétiques qui lui permettent une manipulation efficace, en temps réel et associables à une multitude d'actions différentes. Pour comprendre cette efficacité, on doit postuler que des opérations complexes de mise en forme quelconque, de regroupement c'est-à-dire de classification sont à l'œuvre. De fait, tout ce qui s'inscrit dans les capteurs ne se rend pas nécessairement au cerveau. Ce qui s'y rendra (parcours R2) subira d'importantes transformations et présentera une forme quelconque de synthèse. La représentation qui sera finalement produite ne contiendra qu'une infime partie de ce qui a été capté. Ce sont ces représentations qui sont à proprement parler celles *enregistrées* par le cerveau et qui posent les problèmes cognitifs les plus intéressants. Ce sont elles qui, par exemple, permettront à un agent de reconnaître des objets ou des événements déjà rencontrés (Edelman, 1999) d'ajuster une action à une fin, ou encore certaines d'entre elles feront l'objet de la conscience phénoménologique.

Ainsi, pour réaliser de telles fonctions, ces représentations doivent être éminemment synthétiques. Cependant cette forme synthétique n'est pas chose simple à comprendre. En effet, c'est ici que notre conception intuitive d'une représentation est souvent influencée par le modèle du type I de représentation. En effet, une représentation nous apparaît comme une entité concrète singulière - un genre d'image - qui est vue comme une réponse interne à un stimulus. Pourtant, comme nous le rappellent si bien Wittgenstein (1953) et Goodman (1988), représenter ne consiste pas simplement à rendre présent à nouveau un état particulier du monde. Ceci se réalise dans l'inscription. Il faut aussi le rendre présent *sous un ou l'autre aspect* (représentation « as »).

« L'usage ordinaire néglige souvent la distinction entre une représentation et une représentation comme (as) ... Si je vous dis que je possède un cheval noir et que je vous montre une photo où il apparaît comme un point noir dans le lointain, vous ne pourrez dire que j'ai menti, mais vous penserez certainement que je vous ai déjoué. Vous vous attendiez à ce que l'image soit celle du cheval noir c'est-à-dire qu'elle le représente comme un cheval noir. » (Goodman, 1988, p. 14)

Autrement dit, le processus cognitif construit des états internes qui sont plus que des inscriptions. Ceux-ci doivent saisir l'objet capté *sous un aspect* ou un autre.³ D'ailleurs, lorsqu'ils deviennent conscients, nous les percevons sous un aspect ou un autre.

³ Ce problème de l'aspect dans une représentation perceptive est un problème théorique profond. (voir Clementz, 2001) Ces aspects sont-ils des propriétés de l'objet ou des dimensions de notre expérience?

« On ne peut pas voir, sentir ou goûter un objet sans qu'il nous apparaisse comme ayant certaines propriétés déterminées » (Dretske, 1995, p 30).

De plus, comme le dit Lewis, le concept de système d'état interne « réalise les patterns de généralisation mis de l'avant par la psychologie populaire. » (Lewis D., 1980, p. 124). Autrement dit, ce que la psychologie populaire appelle un état psychologique comme, la douleur, le plaisir, le désir, la peur, etc. n'est en fait que le nom donné à un processus complexe de généralisation qui crée des formes synthétiques. En termes logiques, ces formes synthétiques ces « aspectualisations » mettent en œuvre des opérations très spécifiques radicalement différentes de celles que l'on rencontre dans l'inscription. La représentation II est le résultat de l'application d'une multitude de contraintes qui formellement sont des fonctions discriminantes (Nilson, 1990) ou une morphodynamique complexe (Petitot, 1990) où jouent de l'élimination, de la sélection, de la fusion, etc.

« Les résultats [de recherche] ont montré expérimentalement et démontré mathématiquement que dans de nombreux systèmes naturels organisés au moins deux niveaux, le niveau macro (global, grossier, en général finement descriptible) émergeant à travers des comportements collectifs ordonnés et coopératifs, du niveau « micro » sous-jacent (local, complexe, en général non finement descriptible) est essentiellement organisé autour des singularités des processus physiques « micro ». Les singularités structurent morphologiquement les phénomènes. Elles sont phénoménologiquement dominantes et soumises à des contraintes abstraites et universelles (platoniciennes) mathématiquement formulables et largement indépendantes de la physique « fine » des substrats. » (Petitot, 1990, p. 144)

En termes classiques, ces représentations sont le résultat de fonctions classificatoires c'est-à-dire de fonctions organisatrices de formes ou de classes. Elles résultent de fonctions d'équivalence sur les informations captées dans l'inscription.

Pour illustrer ce type de représentation II, reprenons notre exemple de l'œil. Chaque cône et bâtonnet – et il y en a quelques 250 000 000 pour chaque œil – réagit et produit un état interne propre qui ne dure qu'une fraction de seconde. « L'image » rétinienne est composée de cette multitude d'excitations juxtaposées et sans ordre. Pour arriver à un état cognitif, les modèles cognitifs ont tous postuler des processus par lesquels sont extraites de cette myriade de stimulations et excitations des informations pertinentes. Par exemple, extraire de l'information sur la position, sur l'intensité lumineuse, sur la proportionnalité, etc. Cette extraction peut être faite en série ou en parallèle. Elle peut aussi se réaliser tantôt très près de la rétine (les multiples types de cellules ganglionnaires), tantôt tout au long du parcours qui va de ces dernières au cortex visuel. Divers modules (Fodor, 1982) et parcours seraient possibles. Il y aurait même

plusieurs sous-modules pour la vision (Pacherie, 1993). Lee et Dacey (1997) ont montré, par exemple, que simplement pour identifier la couleur, le parcours se divisait en trois voies (couleurs) spécialisées selon leur origine spécifique dans les cônes et bâtonnets. Plus loin encore, certaines cellules (ganglions parasols) opéreraient des transformations sur l'information transmise. Dans cette logique on peut très bien comprendre, comme l'ont montré les théories de la mémoire (Atkinson et Shiffrin (1968)), qu'interviennent tout au long de ce parcours d'autres informations qui ne sont pas issues de la captation originelle. Et il pourra y avoir aussi retour excitateur ou inhibiteur, les cônes et bâtonnets et même les ganglions brouillant ainsi l'image rétinienne initiale. Bref :

« Quelle que soit la quantité d'informations contenue dans la lumière frappant les rétines, il faut des mécanismes mentaux pour reconnaître les objets d'une scène et les propriétés de ceux-ci que la vision rend explicites à la conscience ». (Johnson Laird, 1988, p.67)

Ainsi, ce qui se rend au cortex cérébral est le résultat d'un long processus de sélection, d'ajustement, de fusion, d'influence, de rencontre, mais surtout de réorganisation qui crée une représentation synthétique ou généralisante, c'est-à-dire de classification. L'objet représenté n'est plus l'objet capté originellement.

« L'objet dont une représentation est la représentation n'est pas déterminé par les propriétés que cet objet est représenté comme possédant ». (Drestske, 1995, p. 24)

Autrement dit, il s'agit de la représentation d'un objet, décanté, organisé, filtré, influencé, mais surtout synthétisé, c'est-à-dire présent sous une forme généralisante.

Traditionnellement en sciences cognitives et surtout en philosophie, ce sont ces formes générales qui, lorsqu'elles parviennent à la conscience, sont appelées « sensations » ou « perceptions ». En termes phénoménologiques, elles produisent les qualia, l'aesthesis, le percipi, le noème etc. dont on parlera souvent dans les descriptions phénoménologiques.

« It has been suggested that « our senses of taste and smell are yoked together phenomenologically... the resulting conscious sensations... [are] the products of an elaborate process of integration of information from a variety of sources ». (Dennett, 1991)

Ainsi, ce type de représentation II, de par les transformations qui sont opérées rend présent l'objet sous un aspect ou un autre parce qu'elle est classifiante.

2.3. Type III : La représentation comme catégorisation

Le troisième espace théorique où apparaît le concept de représentation est celui où il faut expliquer quelles opérations un agent effectue sur les représentations de type II. Plus spécifiquement, il faut expliquer comment l'agent se situe par rapport à ces représentations et, surtout, comment elles sont mises en interrelations les unes avec les autres pour construire des systèmes complexes de « croyances » et, finalement, les situer par rapport à des actions dans l'environnement. D'un point de vue cognitif, inscrire et classer les stimuli sont des opérations de base. Pour des agents cognitifs complexes, la cognition exige plus que cela. Il lui faut pouvoir manipuler les résultats de la classification de manière *catégorielle*, c'est-à-dire les situer par rapport aux autres représentations plus spécifiquement encore en leur assignant une relation spécifique à des actions.

Par exemple, par ses capacités cognitives, tout oiseau peut facilement percevoir un objet dans son environnement (une graine, une brindille, un grain de sable, etc.) Mais, pour survivre, il doit être cognitivement capable de faire plus que de percevoir. Il doit aussi reconnaître cet objet comme appartenant à une catégorie, par exemple, à la catégorie « MANGEABLE »⁴ ou « NON MANGEABLE ». Qui plus est, il peut aussi voir ce même objet à un autre moment comme étant de la catégorie « UTILISABLE POUR LE NID » ou non « UTILISABLE POUR LE NID ». De même, une personne pourrait reconnaître un certain organisme vivant comme étant de la catégorie⁵ « CANICHE ». Ce qui lui permettrait alors de mettre cette catégorie en relation avec la catégorie « CHIEN » et de mettre cette dernière en relation avec la catégorie « ANIMAL » et ensuite, « VIVANT », etc. Ou encore, si pour une oie dans une basse-cours, un évènement qui

⁴ La catégorie « MANGEABLE » ou « NON MANGEABLE » ne doit pas être lue ici comme un prédicat linguistique. Catégoriser n'implique pas immédiatement la nomination de la catégorie dans une langue. On peut catégoriser quelque chose sans les nommer. Il existe diverses formes de catégorisation que ne sont pas nécessairement nominative (Voir Montague, 1973 ; Meunier, 1998).

⁵ Nous distinguons nettement entre classe et catégorie. Toute catégorie implique une classe mais toute classe n'implique pas une catégorie. Une machine à sous peut classer des pièces de monnaie selon des gabarits, mais elle ne les catégorise pas pour autant. Un réseau de neurones peut classer mais ne catégorise pas nécessairement. En psychologie cognitive (Harnad, 1987 ; Rosch, 1973) la distinction n'est pas faite parce qu'il est impossible pour un agent cognitif de séparer les deux. Mais logiquement il faut les distinguer. Une théorie des classes ne livre pas nécessairement une théorie des ensembles. Une algèbre des ensembles est par ailleurs différente d'une algèbre des catégories. Sur un plan formel, une classe met en œuvre une fonction d'équivalence sur des éléments c'est-à-dire que les arguments de la fonction sont des entités d'un type élémentaire alors qu'une catégorisation est une fonction dont les arguments peuvent être autant des éléments que des classes. A titre d'exemple en linguistique on peut classer ensemble les multiples occurrences du mots « jean ». On ne reconnaît pas habituellement cette classe de manière catégorielle. Mais la catégorie « SN » implique une fonction d'équivalence sur les classes de mots elle-mêmes et surtout de certaines classes de relation. Vg. la classe « ART » en relation de concaténation «+» avec la classe « NOM ».

met en jeu un objet de la catégorie « HUMAIN » et qu'un tel événement entraîne toujours un événement de la catégorie « OBJET LANCÉ » et que ce type d'événement entraîne aussi celui « OBJET MANGEABLE » alors la simple apparition de l'événement de la première catégorie entraînera possiblement celui de la troisième catégorie.⁶

Ce type de représentation dont la fonction est catégorisante est essentiel à l'activité cognitive. Elle repose sur la capacité d'un agent cognitif à construire des méta-représentations, c'est-à-dire d'appliquer des opérations sur des représentations. C'est par une telle catégorisation que la puissance cognitive supérieure d'un agent cognitif se manifeste. En lieu et place des signaux captés, mais surtout des classes construites, un agent cognitif doté de ce pouvoir peut alors parcourir des structures internes, explorer de manière créatrice les réseaux, mais surtout lorsqu'elles sont traduites en des signes conventionnellement partagés, c'est-à-dire en un langage, il peut en communiquer le contenu aux autres agents. Cette capacité est évidemment partagée de manière différente selon les espèces d'agent cognitif. Un vers de terre, ne catégorise pas son environnement comme un chien guide ou un dauphin.. Et ceci est d'autant plus vrai chez les humains. Comme nous le montre la vie quotidienne, deux personnes ne catégorisent pas nécessairement leur monde de la même manière et n'en tirent pas les mêmes conclusions !

La présence de ce type de représentation a été abondamment illustrée en sciences cognitives récentes.

Pensons ici à tout le processus que Rosch (1973) a mis en évidence dans la construction des classes perceptives. Celles-ci sont manipulées comme via des représentations prototypiques ou des exemplaires typiques qui sont des catégories. C'est, semble-t-il, la manière dont un agent peut inclure dans des classes des individus atypiques, déviants, etc. Exemple: la catégorie « OISEAU » pourra inclure le moineau (le prototype) et le pingouin (atypique). Autrement dit, des objets aux propriétés variées ne deviennent cognitivement manipulables que lorsque l'agent réussit à réunir sous un prototype les individus de cette classe.

Karmiloff Smith (1992) a, dans des expériences nombreuses, montré que la cognition dépendait de l'habileté d'un agent cognitif à redécrire son information dans des langages de hauts niveaux. Elle a, par exemple, demandé à des enfants de divers âges de faire des dessins représentant un certain concept (*un homme à deux têtes, une maison avec des ailes, etc.*). Aucun enfant ne pouvait évidemment avoir eu antérieurement une telle expérience perceptive. Or il s'est avéré

⁶ On reconnaît là, évidemment, le processus déductif. Tous ne seront pas d'accord pour reconnaître que l'inférence peut se fonder sur une association. Quoi qu'il en soit, il est clair que l'association peut être transitive. Si $A \rightarrow B$ et si $B \rightarrow C$, alors $A \rightarrow C$. Et une inférence peut très bien se fonder sur la transitivité, Mais ce n'est pas sa seule possibilité.

que la possibilité de construire de tels dessins dépendait de la capacité de l'enfant à utiliser des langages (naturels ou iconiques) catégorisateurs. C'est-à-dire de pouvoir associer aux classes des catégories présentant des propriétés sémantiques diversifiées. Par exemple, pour lier la catégorie « MAISON » à celle de « AILES » ou la catégorie « HOMME » à celle de « DEUX TÊTES », il faut avoir des classes perceptives que l'on peut manipuler par leurs catégories qui alors peuvent être réunies via leur contenu sémantique. C'est par la catégorisation qu'une maison avec des ailes est pensable. On ne peut en avoir de perception.

Un troisième exemple de catégorisation encore plus intéressant se trouve dans l'acquisition d'expertise par les agents travaillant dans des domaines techniques spécialisés. Holyoak (1991) a montré que cette expertise ne pouvait reposer sur la simple habileté à manipuler des heuristiques de classification. Au contraire, l'expert devait pouvoir naviguer à travers des domaines en compilant les solutions de problèmes en des formes de plus en plus générales et catégorisées. Autrement dit, il doit manipuler ces formes de manière catégorielle. Cependant, l'intérêt de cet exemple est que cette catégorisation ne passe pas nécessairement par des catégories verbalisées dans un langage naturel, mais par une sémiotisation motrice ou corporelle. Cette catégorisation, souvent idiosyncrasique, permet à l'expert de réaliser des tâches qui exigent une sensibilité à la moindre variation dans l'espace. Par exemple, un petit mouvement de la manette dans un jeu vidéo, dans l'opération de la machinerie lourde (les grues mécaniques), etc. L'habileté cognitive ici est catégorisante, mais elle ne se trouve aucunement dans une forme linguistique. Elle est dans le geste de touche angulé d'un outil ou d'une manette etc.

Un quatrième exemple nous vient de la linguistique. La sémantique et la sémiotique cognitives (Landgacker, 1986; Talmy, 1984 ; Fauconnier, 1984; Lakoff, 1983 ; Fontanille, 1995) ont montré exemples après exemples comment le langage naturel était essentiellement une forme très élaborée de catégorisation de l'expérience perceptive (plus particulièrement spatiale) du monde.

Ainsi, dans ce type de représentation cognitive, les opérations consistent à « manipuler » des classes comme des objets en soi (inclure, transformer, relier, comparer, ajuster, adapter, etc.). En termes classiques, un agent aux habiletés cognitives supérieures doit pouvoir catégoriser, c'est-à-dire les inclure dans un système de relations, et ainsi de pouvoir en parcourir les liens internes. En termes piagétien, c'est par cette catégorisation que l'agent cognitif *s'adapte et assimile son expérience* du monde. (Piaget, 1967) Ce type de représentation catégorisante posera des problèmes théoriques des plus difficiles. Par exemple, cette catégorisation est-elle conceptuelle, non

conceptuelle? Doit-elle passer ou non par le langage? Ces opérations sont-elles aussi physiquement soutenues ?

III. Les caractéristiques sémiotique, ontologique et logico-épistémique de trois types de représentation

Comme nous le voyons, malgré l'utilisation par les sciences cognitives d'un même terme, le concept de représentation ne recouvre pas une même réalité. La représentation peut être pensée tantôt comme inscription, tantôt comme classification ou comme catégorisation. Nous allons étudier maintenant comment chacun de ces types de représentation repose sur une architecture conceptuelle différente. En fait comme nous avons tenté de le montrer ailleurs (Meunier 1999), elles réaliseront chacune à leur manière les caractéristiques de trois thèses, à savoir : une première de type sémiotique, une deuxième ontologique et une troisième de nature logico-épistémique.

3.1. La thèse sémiotique

Comme on l'a souligné à plusieurs reprises dans la littérature critique (Dennett, 1978; Dreyfus; 1998, Cummins, 1989; Dadesio, 1995, Jorna, 1990; Meunier, 1989), le concept de représentation repose sur une thèse de nature essentiellement sémiotique. En effet, ce qui définit une représentation est qu'un état interne soit en relation de dépendance avec un objet externe. Et on parle de cette dépendance en terme sémiotique de renvoi, c'est-à-dire qu'un état interne de type cognitif *signifie* toujours autre chose que lui-même. Il est *porteur d'information à propos* de quelque chose.

« By representational system (RS) I shall mean any system whose function is to indicate how things stand with respect to some other object ». Drestke (1982, p. 52) (Nous soulignons)

De fait, cette relation⁷ reprend la formulation sémiotique augustinienne du *aliquid stat aliquo*, ou la formule frégéenne ou peircienne du « something that stands for something else ». Et dans la tradition sémiotique, cette relation est déployée d'une double manière. Elle est en effet pensée

⁷ Malheureusement, en sciences cognitives, lorsqu'elle est explicitée, cette thèse est souvent réduite à ses formes les plus élémentaires, telles par exemple : une représentation cognitive est une relation à un objet, souvent physique et externe et elle est essentiellement de nature symbolique. De telles explicitations ne rendent pas justice à la complexité des thèses sémiotiques.

comme une relation *non naturelle* ou une relation *naturelle*. La première est *non naturelle* si, entre le représentant (aliquid) et le représenté (aliquo), la relation est de nature intentionnelle, c'est-à-dire fondée sur l'intention d'un agent. Ce type de relation, faut-il le rappeler, est, dans la tradition sémiotique classique, celle qui caractérise essentiellement la forme dite *symbolique* de représentation. La seconde est dite *naturelle* si entre le représentant et le représenté, il existe une relation de nature causale. Ce type de relation est appelé dans la tradition sémiotique un signe naturel ou, dans la conception peircienne, un indice :

« L'indice est lié physiquement à son objet, ils forment une paire organique mais l'esprit qui interprète n'a rien à faire avec ce lien ». (Peirce, CP, 2 :299, Écrits, p 165)

« Un indice est un signe qui réfère à l'objet qu'il dénote en raison du fait d'être affecté par celui-ci ». (Peirce, CP, 2 :231)

Pour plusieurs théories cognitives, une représentation, et plus particulièrement des états mentaux, est conçue avant tout comme une relation du premier type, c'est-à-dire une relation non naturelle. Elle est intentionnelle⁸. Le problème de cette position est qu'elle en appelle nécessairement à un concept d'interprétation qui demeure problématique pour une théorie cognitive. En effet, le terme intentionnel possède habituellement le sens que Brentano lui avait originellement donné, à savoir qu'une représentation est intentionnelle parce qu'elle est « à propos de » quelque chose. C'est-à-dire qu'il n'est qu'une autre formulation du concept de « renvoi ». S'il en est ainsi, il faut expliquer comment ce « renvoi » existe. Et les réponses à cette question postulent un interprète quelconque (désir, croyances, etc.).

Pour plusieurs chercheurs des sciences cognitives contemporaines⁹, ce modèle dit interprétatif de la représentation, bien que riche, est problématique. Premièrement, seuls certains types de représentation possèdent cette propriété. Ce sont celles de nature symbolique (langages, icônes, etc.). Deuxièmement, de quelle nature est cet interprète interne? Fonctionne-t-il lui aussi avec des représentations qui doivent à leur tour être interprétées. Et ainsi de suite? C'est l'argument de l'homoncule (Dennett, 1978). Et troisièmement, ce dont une théorie cognitive doit rendre compte, ce sont les représentations qui naissent ou émergent « naturellement » au sein

⁸ De plus comme le dit Cummins (1989) on ne doit pas ramener la question de la représentation à l'intentionnalité. Cette dimension intentionnelle est fort probablement importante dans les théories cognitives humaines, mais du point de vue de l'investigation de la question de la représentation dans un système informatique et cognitif en général, cette question demeure secondaire et peut même obnubiler la véritable analyse. Pour certains (Dennett, 1978; Cummins, 1983) l'intentionnalité (au sens de désir et croyance) n'est pas une marque essentielle d'une représentation.

⁹ De plus en plus de chercheurs pensent que le modèle peircien de l'esprit met aussi en jeu une dimension indicieuse importante (voir DeTienne, 1996)

même de l'agent en raison de sa nature cognitive. Comment se produisent « naturellement » les représentations que sont la sensation, la perception, l'aesthesis, les qualia et même pour certains les désirs, les croyances, les concepts, etc.? Une théorie cognitive pourrait expliquer leurs modalités de présence à la conscience mais elle n'explique pas leur genèse dans la cognition.

Devant ces difficultés, on voit de plus en plus la recherche cognitive reprendre la question de la représentation dans la deuxième perspective, à savoir comme une relation de type naturel, c'est-à-dire, en termes sémiotiques, comme une relation non pas de nature intentionnelle ou symbolique, mais une relation de nature indicielle¹⁰. Ce qui revient à poser la question suivante: comment une relation causale peut-elle fonder une représentation?

Aussi, quel type de relation sémiotique est à l'œuvre dans les trois types de représentations que nous avons distinguées? Sont-elles de nature intentionnelle ou indicielle?

3.1.1. L'inscription

Il nous semble évident que ce type de représentation n'est aucunement de nature intentionnelle, c'est-à-dire symbolique au sens sémiotique de ce terme. La réaction d'un capteur ne se fonde aucunement sur une intention, encore moins sur une convention. L'inscription est en fait une représentation de type indiciel. Elle en est un excellent exemple parce que la relation entre l'objet externe et l'état interne est essentiellement comprise comme une relation de nature causale. À chaque dimension à laquelle est sensible un capteur correspond un « point » dans l'état interne de ce capteur. Par exemple : dans l'image rétinienne, il y a correspondance une à une entre les cellules de la rétine et les points lumineux de l'objet. Dans le vocabulaire de Cummins (1989), elle en serait une représentation de type *covariance*. Dans un modèle formel, elle serait une représentation fonctionnelle projective.

Ce type de modèle a été souvent critiqué parce que purement physique et à ce titre il ne peut pas fonder la cognition. Dans la tradition classique, on le refusa parce qu'il ne pouvait expliquer l'illusion argument qui a été repris par la philosophie cognitive contemporaine (Fodor 1981), mais en termes de « misrepresentation » (Dretske, 1982; Dennett, 1978; Millikan 1984). Wittgenstein (1922) le refusa sur la base de l'impossibilité d'explicitier l'implicite du morphisme qu'est la ressemblance. Mais ces critiques nous semblent porter sur la perception et la sensation

¹⁰ La thèse sémiotique reçoit une formulation particulière en science cognitive. Elle est souvent traduite en termes d'information. Comme l'ont bien saisi Clarke (1987), Dretske (1982), le concept d'information n'est que la reformulation du concept classique de signe naturel.

« *Talking about information is yet a third way of talking about the fundamentally important relation of indication or natural meaning* ». (Dretske, 1982, p. 58).

qui sont des processus qui surviennent beaucoup plus loin dans le processus cognitif. Ils relèvent de la dimension épistémique liée à la sensation, la perception et la conceptualisation.

Il faut accepter, nous semble-t-il, qu'une inscription soit une *représentation*. Sinon il faut rejeter du domaine de la cognition tout ce travail que font les capteurs sensibles. Et en conséquence, il y a risque de ne plus pouvoir expliquer comment le processus cognitif est en relation avec des objets du monde externe. Mais il faut voir ce type de représentation comme élémentaire et surtout il ne faut pas confondre ce type de représentation avec l'expérience consciente qu'un agent peut en avoir.

3.1.2. La classification

Sur le plan sémiotique, ce type de représentation est le prototype de ce qu'est une représentation parce que, selon plusieurs sémioticiens, elle met justement en jeu un travail de généralisation /classification. Comme le dit Peirce, toute représentation est généralisante. (Peirce, CP 7 647, 5 (1903), 115 (1903)).

Mais la question qui est alors posée est l'origine de cette classification. Celle-ci est-elle d'origine intentionnelle ou naturelle? C'est-à-dire en termes cognitif, symbolique ou indiciel? La réponse varie. Pour la tradition classique des sciences cognitives, ce type de représentation doit être symbolique et plus précisément prédicative. Car la généralisation est sur le plan logique une classification et en conséquence elle exige un critère d'équivalence et donc elle met en jeu de la prédication. Toute classe repose sur un prédicat nous dit la logique.

Pour d'autres chercheurs, et c'est la tendance de plus en plus marquée des sciences cognitives et peut-être au contraire de ce qu'intuitivement nous pouvons penser, ce type de travail peut aussi être pensé de manière causale. De fait, il existe de nombreux systèmes physiques qui causalement classent des domaines et ceci sans passer par de la prédication et du symbolique. On peut ainsi penser qu'une représentation généralisante est causale et donc de nature indicielle.

3.1.3. La catégorisation

Comme nous l'avons dit plus haut, la catégorisation est plus qu'une classification, elle est aussi une représentation qui organise en des systèmes complexes. Pour ce faire, il faut que l'agent puisse manipuler ces classes comme des objets, ce qui implique qu'il utilise une forme sémiotique quelconque. La question, à laquelle la recherche ne peut encore répondre, est de savoir si cette

opération sur des classes exige que la catégorisation soit explicite (c'est-à-dire passe par des symboles et donc soit intentionnelle) ou implicite (c'est-à-dire indicielle et donc causale). Les conceptions de catégorisation que nous présentent les sciences cognitives vont dans ces deux directions. La catégorisation peut prendre la forme de *proposition* (Fodor, Newell), de *schéma* (Piaget), de *prototype* (Rosch), de *nom* (Landgacker), de *morphisme* (Petitot), d'*attracteur* (Varela, Van Gelder, Ellman) et même d'*analogie* (Paivio, 1971). Pour la tradition philosophique (Aristote et surtout Kant), cette catégorisation est habituellement de type conceptuel. Mais de plus en plus on pense que la catégorisation n'est pas nécessairement et immédiatement de nature conceptuelle. Toutes ces formes de catégorisation sont sémiotiques, mais certaines sont formellement symboliques, d'autres formellement indicelles. La nature spécifique de cette catégorisation est objet intense de recherche. (Peacocke, 1992 ; Evans, 1982 ; Drestke, 1995 ; Pacherie, 2000 ; Clementz, 2000 ; Dorik, 1996 ; Brandom, 1994)

3.2. La thèse ontologique

Une des thèses les plus importantes que doivent soutenir toutes les théories de la représentation est de nature ontologique. En effet, chaque théorie doit préciser le type d'existence des entités constituant une représentation. Ce qui exige d'identifier les propriétés et les relations qui affectent ces entités. Et le débat sur ces questions est ici isomorphe à celui sur le plan sémiotique. Ces systèmes reposent-ils sur une ontologie de type non naturel (intentionnelle) ou de type naturel?

Pour les uns, seules des entités cognitives sont de nature matérielle et même elles sont essentiellement de nature physique. De plus, elles sont affectées par des rapports de causalité. Pour d'autres, ces entités sont de nature intentionnelle et les rapports qu'elles entretiennent sont sémantiques. Les trois types de représentation que nous avons identifiés sont engagés vis-à-vis de l'une ou l'autre de ces ontologies.

3.2.1. L'inscription

Dans la représentation de type inscription, il est évident que l'ontologie est naturaliste. Ce qui existe concrètement sont des capteurs qui réagissent à une excitation spécifique. Une représentation n'est que le conglomérat de ces réactions individuelles. Et il en existe autant qu'il y a de capteurs. Par exemple, comme nous l'avons illustré plus haut, pour chaque œil il y en aura quelques 250 000 000. De plus, chaque réaction, qui est dépendante de la structure des capteurs

eux-mêmes, ne dure qu'un temps limité. Pour l'œil par exemple, un minimum de 24 excitations par seconde seront requises pour créer la perception de mouvement. En conséquence, il existera pour chaque micro-moment et chaque position une instance d'une représentation spécifique. Autrement dit, il existe des myriades d'occurrences (token) de ce type de représentation. Ainsi, l'inscription ne crée que des représentations individuelles. C'est pourquoi on dira qu'il n'y a pas de généralité à ce niveau.

Par ailleurs, chaque capteur est *spécifique* et *autonome*, c'est-à-dire que sa réaction est locale (non distribuée). Par exemple, dans l'œil, seuls les photons peuvent provoquer une réaction des cônes et bâtonnets, lesquels sont régis par des paramètres et des contrôles internes propres. Certes il peut exister de l'interaction, par exemple de l'inhibition, mais celle-ci, lorsqu'elle existe, ne modifie pas le « mécanisme interne » du capteur (la cellule) comme tel. Il bloque l'extrait de neurones¹¹. Sur le plan causal, un capteur n'est rien d'autre qu'un transducteur. Il prend une énergie d'un type et la transforme en une autre. En conséquence, il suit les lois de la physique qui sont habituellement pensées dans un cadre newtonien.

Comme on le voit, à toute variation dans le temps et l'espace correspond une nouvelle représentation. Évidemment, une telle multitude de représentations entraîne immédiatement une surcharge de la mémoire. En conséquence, ce type de représentation ne permet aucunement d'expliquer comment un agent cognitif réussit à construire un état interne synthétique de la multiplicité des projections qu'il construit au fur et à mesure de la continuité du temps et de la variation spatiale.

3.2.2 La classification

Ce deuxième type de représentation pose des problèmes ontologiques sérieux. En effet, elle reprend une question (millénaire) : quel est le statut ontologique d'une représentation dont la propriété essentielle est d'être générale, c'est-à-dire qui se donne sous la forme de classe? Une classe est en effet une entité d'un type différent de celui des éléments qui la composent. Une classe est plus que la somme des individus. L'interrelation entre les individus est autre chose que la classe comme entité. Mais alors, comme le soutient une position nominaliste, s'il n'existe que des entités individuelles, et que, en l'occurrence, ce sont des inscriptions individuelles, quel statut

¹¹ L'inhibition latérale est constante dans les modèles connexionnistes. Mais on observera que cette inhibition s'applique à l'entrée du stimulus. Un neurone est dit inhibé lorsqu'il ne produit pas d'extrait. Mais la raison n'est pas que les paramètres qui régissent les neurones ont été changés. Ce sont les poids qui en fait ont été tellement modifiés que le neurone ne présente plus d'extrait. En tant que tel, le neurone reste le même.

d'existence peut-on donner à une représentation classificatoire? La réponse traditionnelle à cette question fut de leur octroyer le statut d'entités intentionnelles en raison de leurs propriétés logico-sémantiques.

Il nous semble qu'une meilleure explication du processus de classification lui-même peut nous aider à mieux comprendre ce statut. De fait, la recherche cognitive contemporaine offre un certain nombre de modèles montrant la production et l'émergence causale de classes. (Nilson, 1990) Pour les uns (l'hypothèse du symbolisme matériel, par exemple), le processus est linéaire et repose sur un concept de causalité efficiente classique. Pour les autres (les hypothèses connexionnistes et dynamiques), le processus est parallèle et en appelle à des modèles causaux non -linéaires et souvent même chaotiques et dissipatifs. Ici s'insère paramétrage du contrôle et des attracteurs. Les entités concrètes sont toujours des individus, mais ils sont de divers types en ce sens que leurs rôles fonctionnels sont variés. Tous ces modèles permettent de comprendre de mieux en mieux comment causalement des représentations de formes synthétiques peuvent exister.

3.2.3 La catégorisation

La représentation catégorielle pose aussi des problèmes ontologiques originaux. Quel est, en effet, le statut ontologique d'une catégorie? Pour nous, la réponse est dans la suite de ce qui a été dit à propos de la représentation classificatoire. En effet, l'assignation d'une catégorie à une classe relève de la nature de la relation que peut entretenir cette classe avec d'autres classes. La catégorisation est une classification de classification. Elle sert essentiellement à définir pour un agent cognitif les rôles et rapports que les représentations classifiantes peuvent avoir dans ses actions. En ce sens, elles sont plus que des classes : elles sont des modèles. Traditionnellement, on assigne à ces catégories une ontologie qui relève du monde intentionnel, c'est-à-dire sémantique. Mais de plus en plus (Gardenfors, 1994, 2000 ; Pacherie, 1993) on se demande si cette catégorisation ne pourrait pas aussi être expliquée comme un mode de fonctionnement naturel.

3.3 La thèse logico-épistémique

La dernière thèse concerne les propriétés logiques qu'un agent cognitif entretient avec ces représentations. Il est en effet, presque impossible de parler de *représentation* sans préciser la nature logique du rapport que l'agent entretient avec les représentations qui sont construites. Et

ce rapport entraîne des propriétés logico-épistémiques très spécifiques. Évidemment, les positions sur la nature de ce rapport varieront selon l'agent cognitif qu'on modélise c'est-à-dire selon s'il s'agit d'un robot, d'un animal, d'un humain.

Pour les uns, toute représentation est nécessairement objet de désir ou de croyances. C'est-à-dire qu'elles sont construites soit pour répondre à des actions soit pour être évaluées. Des croyances peuvent ainsi être bonnes ou mauvaises, adéquates ou non, vraies ou fausses, conscientes ou non, etc. Pour d'autres, la représentation interne n'est qu'un autre état physique parmi d'autres et toute relation à d'autres états internes n'est que de nature causale physique.

Ainsi, il faut identifier comment nos trois types de représentation se comportent par rapport à trois questions importantes qui classiquement caractérisent des états épistémiques : à savoir le principe de substituabilité, les transformations admises et finalement le statut des classes.

3.3.1 L'inscription

Comme nous l'avons dit plus haut, l'inscription est une représentation fondée sur les automatismes des capteurs. À moins qu'il ne soit défectueux¹², un capteur ne réagit qu'à ce à quoi il est sensible, c'est-à-dire il est affecté *nécessairement* par ce qu'il lui est propre (e. g. un bâtonnet ne réagit qu'à un photon). C'est ce qui assure la stabilité de leur fonctionnement. Et ce qui, comme le montre si bien Drestske (1986), le rend porteur d'information au sens technique de ce terme. La relation causale entre l'intrant et l'extrant du capteur est une relation nécessaire et sa probabilité est 1 !

La conséquence immédiate de ceci est que l'inscription fonctionne selon un principe logique d'extensionnalité. Elle n'admet pas le principe de substituabilité des identiques. C'est-à-dire que seul un stimulus de même type peut provoquer une même réaction. Par exemple, dans les bâtonnets, seul un photon provoque une réaction. Et un photon en vaut un autre.

Vu l'automatisme des opérations dans l'inscription, les transformations que ce type de représentations peuvent subir sont relativement limitées. Leur réaction peut varier, mais uniquement à l'intérieur de paramètres. Elles ne présentent habituellement pas la complexité des transformations de types syntaxiques. Enfin, relativement à la classification, l'inscription sert d'intrant. L'inscription n'est pas une classe, mais sera l'élément d'une éventuelle classe. Par exemple, l'image rétinienne n'est qu'un conglomérat de millions de réactions qui devront être classées selon l'un ou l'autre critère.

¹² Même si un bâtonnet particulier est affecté, le système pourra ultérieurement s'ajuster en raison soit de la redondance du système soit d'autres mécanismes de contrôle des capteurs eux-mêmes.

Sur le plan épistémique, ces caractéristiques de l'inscription permettent de comprendre a) pourquoi un agent peut se fier à ses capteurs. Ils sont automatiques et redondants. b) pourquoi les capteurs ne produisent jamais d'erreur : d'une réaction chimique, on ne dit pas qu'elle est vraie ou fausse. Enfin c) pourquoi leurs opérations sont habituellement vues comme inaccessibles à la conscience, c'est-à-dire phénoménologiquement obscures. Aucune théorie ne soutient qu'un agent est conscient de la réaction de ses cônes et bâtonnets ou des ganglions périphériques. Pour Peirce, ce sera une raison suffisante pour ne pas la voir comme une représentation.¹³

3.3.2 La classification

La généralisation, comme nous l'avons dit plus haut, met en jeu de la classification. Or, une classification est essentiellement une partition ou l'application de fonctions d'équivalence. Classifier, c'est partitionner. Or, comme les éléments d'un domaine peuvent être classés d'une infinité de manières, il faut donc un critère c'est-à-dire qu'il faut des mécanismes de contrôle (paramètres, règles, etc.). Or ces contrôles sont autant d'interventions qui médiatisent la production du résultat final. Comme le processus peut être complexe, il arrive très souvent que du bruit s'y installe. Ceci peut évidemment entraîner des erreurs. Autrement dit, sur le plan cognitif, il peut y avoir une différence entre ce qui est capté et ce qui est perçu. C'est pourquoi, on ne rencontre pas la crédibilité épistémique que l'on trouvait dans l'inscription. Une représentation générale peut donc être erronée, illusoire, déviante, incorrecte, etc. Une des conséquences immédiates de cette complexité est que la représentation construite ne fonctionne pas à l'extensionnalité. Un même « token » et surtout un même type d'intrant ne provoqueront pas nécessairement dans l'agent une même représentation. Ce type de représentation n'admet donc pas la substituabilité des identiques.

Cette complexité entraînera aussi une compréhension des transformations particulières. Pour les uns, ces transformations seront régies par des règles. Pour d'autres, les transformations formeront un système complexe et elles sont essentiellement régies par les paramètres dynamiques.

¹³ Pour Peirce, une représentation est plus qu'un signal. Elle implique un rapport cognitif.

3.3.3 La catégorisation

La catégorisation pose aussi des problèmes logico-épistémiques parce qu'elle est une redescription, c'est-à-dire une méta-représentation fondée sur les diverses relations que les classes peuvent entretenir entre elles. En conséquence, ce type de représentation présentera aussi des propriétés logiques caractéristiques. Comme la représentation de type II, elle ne sera pas extensionnelle et ne permettra pas la substituabilité des identiques. En effet, une même classe peut être catégorisée de plusieurs façons. Et l'une ne vaut pas l'autre. Mais, une méta-représentation de la cognition ne peut se contenter de redécrire des perceptions, c'est-à-dire des représentations généralisantes (perçues consciemment ou non). Dans les organismes supérieurs, tôt ou tard l'agent devra porter un jugement de validité sur ses constructions c'est-à-dire sur ces représentations. Et ce sont ces jugements qui permettront d'explorer les inférences qui sont liées aux relations construites dans la catégorisation. Ces inférences sont les transformations que peuvent subir les méta-représentations et donc traditionnellement elles sont régies par des règles qui précisément se livrent sous la forme de manipulation de catégories.

Enfin, le rapport à la catégorisation est spécifique. Une catégorisation organise les classes en systèmes catégoriels. C'est souvent via cette systématisation que la question de la catégorisation sera reliée à la conceptualisation. Mais la question demeure de savoir si toute catégorisation est nécessairement conceptualisation.

Ainsi, si nous avons exprimé la différence des foncteurs épistémiques en jeu dans ces trois types de représentations, nous pourrions les caractériser dans les termes classiques de l'herméneutique :

Nous les représenterions par les foncteurs épistémiques suivants :

Pour l'inscription : L'agent « SAISIT » p

Pour la généralisation : l'agent « PERCOIT » p

Pour la catégorisation : l'agent « COMPREND » p .

Chacune de ces attitudes dites « propositionnelles » appartiendrait à des systèmes logiques différents. Le premier serait extensionnel et les deux autres intensionnels, mais à des degrés différents.

Conclusion générale

Dans les analyses précédentes, nous avons étudié certaines des caractéristiques du concept de représentation tel qu'on le trouve utilisé dans les sciences cognitives. Dans ce cadre, le

concept de représentation demeure très général. Il n'est qu'un terme commode pour désigner la forme d'état interne qu'un agent crée lorsqu'il est en relation cognitive avec son environnement. Cette généralité le rend souvent vulnérable à la critique. Pourtant, on ne réussit pas à le remplacer. Un des ses avantages est d'être neutre sur ces engagements sémiotiques, ontologiques et logiques. Ce n'est que son usage spécifique dans les espaces théoriques de l'inscription, de la classification, et de la catégorisation qu'il se compromet. Nous avons tenté de montrer que chacun de ces types de représentation présente des caractéristiques sémiotiques, ontologiques et logico-épistémiques différentes et qu'on ne peut les confondre l'une avec l'autre. Ce sont trois types de formes d'états internes commodément désignés sous le terme de représentation. Ces trois types de représentation sont fondés sur la différence des opérations qui sont mis en œuvre dans ces trois types d'état interne. La représentation I inscrit de l'information; la représentation II classe l'information; la représentation III catégorise l'information. Malheureusement, les chercheurs ne retiennent souvent du concept de représentation que l'une ou l'autre des caractéristiques propres à l'un ou l'autre de ces trois types. Ce qui fait qu'en dernière instance, le débat porte sur la différence entre le symbolique et le non symbolique, entre le causal ou l'intentionnel.

Pour les uns, la représentation n'étant que symbolique, doit donc être intentionnelle. Ce qui pose la question de son origine causale. Pour d'autres, la représentation au contraire étant une réponse causale ne semble pas être classificatoire ce qui exclut qu'elle puisse être catégorielle.

Les distinctions que nous avons élaborées ici sur la base de leurs opérations spécifiques rejoignent, nous semble-t-il de plus en plus les recherches contemporaines en philosophie de l'esprit qui chacune à sa manière reconnaît de plus en plus les opérations spécifiques des organes des sens, de la perception et de la conceptualisation. Ainsi, on redonne une place de plus en plus importante aux opérations des sens en la distinguant nettement de la *perception* dont on voit de plus en plus la fonction classificatoire (causale) et que l'on diffère de l'expérience phénoménologique qu'on peut en avoir (les qualia). Et on reconnaît aussi l'originalité de la fonction catégorisante où la conceptualisation ne serait qu'une forme parmi d'autres de cette opération catégorisante¹⁴.

Pour aller plus loin dans notre analyse, il nous aurait aussi fallu aborder un quatrième espace théorique où la question de la représentation se pose en sciences cognitives, à savoir celui où des

¹⁴ Evans, Dretske (1995), Peacocke (1992), Clementz (2000), Thagard (1990), Bermudez (1998), Tye (1995) sont parmi ceux qui, de plus en plus en philosophie, acceptent que certaines opérations cognitives soient catégorielles sans par ailleurs être de nature purement conceptuelle au sens classique du terme c'est-à-dire impliquant la prédication, la nomination dans un langage et l'inférentialité. On parlera alors de proto concept.

agents cognitifs qui s'échangent entre eux des représentations. C'est la question des représentations sociales (Jodelet, 1989) ou des représentations collectives (Vygotsky, 1966; Dawkins, 1982; Sperber, 1996; Luhman, 1982) et finalement du langage. Nous n'avons ici touché que la question de la représentation dans les modélisations des agents cognitifs pris individuellement. Nous sommes convaincus cependant que, dans ces modèles de représentations, on transpose souvent les concepts des théories cognitives individuelles et on y retrouverait des distinctions similaires. La véritable différence apparaîtrait cependant dans les engagements sémiotiques, ontologiques et logiques. ♣

Références bibliographiques

- Atkinson, R. C. & Shiffrin. (1968). *The psychology of Learning and Motivation*. Academic Press.
- Bechtel, W. & Grapham, G. (1998). *A Companion to cognitive science*. Blackwell.
- Bermudez, J. L. (1996). *The paradox of self-consciousness*. Cambridge : MIT Press.
- Billman, D. (1998). Representations. In : *A Companion to Cognitive Science*. W. Bechtel and G. Georges (eds), pp. 649-659.
- Brandson, R.B. (1994). *Making it explicit*. Cambridge : Harvard University Press.
- Brentano, F. (1944). *Psychologie du point de vue empirique*. Paris : Aubier (1874).
- Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence* 47: 139-159.
- Bruner, J. (1990). *Acts of Meaning*. Cambridge, Mass: Harvard University Press / MIT Press.
- Chalmers, David, J. (1991). *Consciousness and Cognition*. Center for Research on Concepts and Cognition. Indiana University Press. Preprint.
- Chomsky, N. (1980). *Rules and representations*. New York : Columbia University Press.
- Churchland, P. M. (1989). *The Neurocomputational Perspective*. Cambridge. MIT Press.
- Clark, A. (1997). *Being There : Putting the Brain Body and World together Again*. Cambridge UP.
- Clarke, D. S. jr. (1987). *Principles of Semiotic*. Routledge and Keagan Paul, New York.
- Clementz, F. (2000). La notion d'aspect perceptif. In: *De la perception à l'action*. Paris : P. Livet (ed), p. 17-56.
- Cummins, R. (1989). *Meaning and Mental Representation*, Cambridge: MIT Press.
- (1983). *The nature of psychological explanation*. Cambridge Mass : MIT Press, Bradford Books.
- Dadesio, T. C. (1995). *On Minds and Symbols, The relevance of cognitive Science for Semiotics*. Mouton de Gruyter.
- Dawkins, R. (1982). *The Selfish Gene*. Oxford Unive Press.
- De Tienne, A. (1996). *L'analytique de la représentation chez Peirce*. Bruxelles : Publications des Facultés universitaires ST Louis.
- Denis, M. (1993). « Pour les représentations » In M. Denis & G. Sabbah (Ed.). *Modèles et concepts pour la science cognitive*. (pp. 95 -107). Grenoble: Presses Universitaires de Grenoble.
- Dennett, D. (1978). *Brains Storms*. Cambridge : MIT Press.
- (1991). *Consciousness Explained*. Boston : Little, Brown.
- Desclées, J. -P. (1990). «Archétypes cognitifs, transivité, et intentionnalité», *Protée*, vol. 18, n^o 2. pp. 7-19.
- Dorik, J. (1998). The Ontology of Perception : Bipolarity and Content. *Erkenntnis*, 48, 153-169.
- Drestke, F. I. (1982). *Knowledge and the flow of information*. Cambridge : MIT Press.
- Dretske (1995). *Naturalizing the Mind*. Cambridge : MIT Press.
- Dreyfus, L. H. (1979). « From Micro-World to Knowledge Representatio ». in Haugeland, J. (ed) *Mind Design*. Cambridge : MIT Press.
- Dreyfus, H. A. and Dreyfus, L. (1988). « Towards a Reconciliation of Phenomenology and AI ». In *Foundational Issues in Artificial Intelligence*. D. Partridge and Y. A. Wilks (eds.), Cambridge : Cambridge University Press.
- Eco, U. (1976). *A Theory of Semiotics*. Bloomington : Indiana University Press.
- Edelman, S. (1999). *Representation and Recognition in Vision*. Cambridge : MIT Press.
- Ellman, J. (1999). « Connectionism, artificial life and dynamical systems ». In : *A companion to Cognitive Science*. W. Bechtel and G. Graham (eds), Oxford : Blackwell, pp. 488-506.
- Evans, G. (1982). *The Varieties of Reference*. Clarendon Press, Oxford
- Fauconnier, G. (1984). *Espaces mentaux, Approche de la construction du sens dans les langues naturelles*. Paris : Minuit.
- Fodor, J. (1982). *The Modularity of Mind*. Cambridge, MA : MIT Press.

-
- Fodor, J. A. (1975). *The Language of Thought*. New York : Crowell.
- (1981). *Representations*. Cambridge : MIT Press.
- Fontanille, J. (1995). *Sémiotique du visible. Des mondes de lumière*. Paris : P.U.F Oxford.
- Franklin, S. (1995). *Artificial Minds*. MIT Press.
- Freeman, W. J. & Skarda, C. (1987). «How Brains Make Chaos in Order to Make Sense of the World». *Behavioral and the Brain Sciences*. 10: pp. 161-173.
- Gardenfors, P. (2000). *Conceptual Space: The geometry of thought*. MIT Press.
- (1994). « How logic emerges form the dynamics of information », in Eijck, J. van and Visser, A. *Logic and Information Flow*. Cambridge : MIT Press.
- Gibson, J. J. (1950). *The Perception of the Visual World*. Boston: Houghton Mifflin.
- Globus, G. G. (1992). « Toward a noncomputational cognitive neuroscience », *Journal of Cognitive Neuroscience*. (4): 299-310.
- Goodman, N. (1988). *Reconception in Philosophy and Other Arts in Sciences*. Routledge.
- Harnad, S. (1987). « Category Induction and representation ». In: *Categorial perception*. S. Harnad (ed), Cambridge University Press, pp. 535-565.
- Haugeland, J. (1987). *Mind Design*. (ed) MIT Press 1987.
- (1986). *Artificial Intelligence : The very Idea*. Cambridge : MIT Press; Bradford Book.
- Holyoak, K. (1991). « Symbolic connectionism : Toward third generation theories of expertise ». In: *Towards a General Theory of Expertise : Prospect and Limits*. K Ericson & J Smith (ed), Cambridge : Cambridge University Press.
- Husserl, Edmund. (1959). *Recherches logiques*. Paris : Presses universitaires de France.
- Jackendoff, R. (1983). *Semantic and Cognition*. Cambridge : MIT Press.
- Jodelet, D. (1989). *Les représentations sociales*. Paris : PUF.
- Johnson-Laird, P. N. (1988). *The Computer and the Mind*. Harvard : Harvard UP.
- Jorna René, J. J. (1990). *Knowledge Representation and Symbols in the Mind*. Tubingen : Stauffenburg Verlag.
- Karmiloff-Smith, A. (1992). *Beyond Modularity : A Developmental Perspective on Cognitive Science*. Cambridge : MIT Press.
- Lakoff, G. (1985). *Les métaphores de la vie quotidienne*. Paris. Ed Minuit.
- Landgacker, R. W. (1986). « An Introduction to cognitive Grammar ». *Cognitive Science*, vol. 10, n° 1, pp. 1-40.
- Lee B. B. & D. M. Dacey. (1997). « Structure and function in primate retina ». In: *Color Vision Deficiencies*. C. R. Cavonius (ed), Holland : Kluwer, Dordrecht. pp. 107-118.
- Lewis, D. (1980). « Mad pain and martian pain ». In N. Block, Ed. *Readings in the Philosophy of psychology*. vol. 1. Cambridge, MA : Harvard University Press.
- Luhman, N. (1982). *The Differentiation of Society*. New York : Columbia University Press.
- Margolis, H. (1987). *Patterns, Thinking and Cognition*. University of Chigaco Press.
- Marr, D. (1982). *Vision : A computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco : W. H. Freeman.
- Meunier, J G. (1989). « Artificial intelligence and the theory of Signs », *Semiotica*. September, p. 43-63.
- (1998). « Categorial structure of Iconic Languages ». *Theory and Psychology*. vol 8, no 6, p 805-827.
- (1999). « Représentation en Sciences Cognitives ». *RSSI*. vol 19, no 2-3, p 83-105.
- Milikan, R. C. (1987). *Language, Thought and other Biological Categories*. MIT Press.
- Montagüe, R. (1972). *Formal Philosophy*. Selected Papers of Richard Montagüe, Thomason editor. Yale University Press Newell (1983).
- Moravec, H. (1988). *Mind Children*. Cambridge, Mass : Harvard University Press.
- Newell, A. (1980). « Physical symbol systems », in *Cognitive Science*, 4, 135-183.

-
- Newell, A. and Simon, H. (1976). « Symbol Manipulation » in *Encyclopedia of Computer Science*. New York: Petrocelli/Charter.
- Newell, A. (1986). « The symbol level and the knowledge level ». In Demopoulos W. (Ed.). *Meaning and cognitive structure. Issues in the Computational Theory of Mind*. Norwood : Ablex Pub. pp. 169-175.
- Nilsson, N. J. (1990). *Learning Machines*. San Mateo : Morgan Kaufman.
- Pacherie, E. (1993). *Naturaliser l'intentionnalité*. PUF.
- (2000). « Les différents niveaux de contenu perceptif ». In: *De la perception à l'action*. Paris : P. Livet (ed), pp. 59-81.
- Paivio, A. (1971). *Imaginary and Verbal Processes*. New York: : Holt Rinehart & Winston.
- Peirce, C. S. (1931-58). *Collected Papers*. Cambridge : Harvard University Press.
- Petitot (1990)
- (1985). *Morphogénèse du sens*. Presses Paris : Universitaires de France.
- Piaget, J. (1967). *La naissance de l'intelligence chez l'enfant*. 4^e édition. Neuchâtel : Delachaux et Niestlé.
- Pollock, J. L. (1989). *How to Build a Person : A Prolegomenon*. Cambridge, Massachusetts : MIT Press.
- Putnam, H. (1988). *Representation and Reality*. MIT Press.
- Pylyshyn, Z. W. (1984). *Computation and Cognition. Towards a Foundation for Cognitive Science*. MIT Press.
- Quine, W. O. (1960). *Word and Objects*. Cambridge: MIT Press.
- Rosch, E. (. 1973). « Natural Categories ». *Cognitive Psychology*. 21(3- 4): 328-350.
- Ryle, G. (1949). *The concept of mind*. Londres : Hutchinson.
- Scott Kelso, J. A. (1995). *Dynamic Patterns*. MIT Press.
- Searle, J. (1992). *The rediscovery of the Mind*. MIT Press.
- Slovan, A. (1983). « Image interpretation : The Way Ahead », in O. J. Braddick & A. C. Sleight (eds) *Physical and Biological Processing of Images. PP*. New York: : Springer Verlag.
- Smolensky, P. (1988). « On the Proper Treatment of Connectionism », *The Behavioral and Brain Sciences*. II 1-74.
- Sperber, D. (1996). *La contagion des idées : théorie naturaliste de la culture*. Paris : O Jacob. 243 p.
- Stich, S. (1983). *From Folk Psychology to Cognitive Science*. Cambridge : MIT Press.
- Talmy, L. (1983). « Spatial Orientation : Theory, Research and Application » Pick, L. Acredolo, (eds) *How Language Structures Space*. Plenum Press.
- Thagard, P. (1990). « Concepts and conceptual change ». *Synthese*. 82: 255-74.
- Thelen, E. & L. B. Smith. (1994). *A dynamic systems approach to the development of cognition and action* Cambridge : MIT Press.
- Thom, R. (1988). *Esquisse d'une Sémiophysique*. Paris : InterÉditions.
- Tye, M. (1995). *Ten Problems of Consciousness : A Representational Theory of the Phenomenal Mind*. Cambridge, MA : The MIT Press.
- Uexküll, Jakob von. (1940). *Bedeutungslehre*. Berlin: Fisher.
- Van Gelder, T. (1993). *What might cognition be if not computation?* Cognitive Sciences Indiana University Research Report 75.
- Van Gelder, T. et R. Port. (1995). « It's about time : An overview of the dynamical approach to cognition », in *Mind as motion : Explorations in the dynamics of cognition*. Cambridge, MA : MIT Press.
- Van Gelder, T. (1997). « Dynamics of Cognition ». In Haugeland, (ed). *Mind design II*. MIT Press.
- Varela, F. J. (1988). *Invitation aux sciences cognitives*. Paris : Seuil.
- Vygotsky, L. S. (1966). *Mind in Society*. Cambridge : Harvard Univ. Press.
- Wittgenstein, L. (1953). *Investigations philosophiques*. Gallimard (1961).
- (1922). *Tractatus logico-philosophicus*. Routledge and Kegan.