



Éducation et didactique

vol 1 - n°3 | Décembre 2007
Varia

Calculer les cultures

Le « structuralisme » dans l'histoire de l'intelligence artificielle

Patrice Maniglier



Édition électronique

URL : <http://journals.openedition.org/educationdidactique/217>

DOI : 10.4000/educationdidactique.217

ISBN : 978-2-7535-1616-8

ISSN : 2111-4838

Éditeur

Presses universitaires de Rennes

Édition imprimée

Date de publication : 1 décembre 2007

Pagination : 79-104

ISBN : 978-2-7535-0597-1

ISSN : 1956-3485

Référence électronique

Patrice Maniglier, « Calculer les cultures », *Éducation et didactique* [En ligne], vol 1 - n°3 | Décembre 2007, mis en ligne le 01 décembre 2009, consulté le 19 avril 2019. URL : <http://journals.openedition.org/educationdidactique/217> ; DOI : 10.4000/educationdidactique.217

Tous droits réservés

CALCULER LES CULTURES :
LE « STRUCTURALISME » DANS L'HISTOIRE
DE L'INTELLIGENCE ARTIFICIELLE*

Patrice Maniglier,
Centre International d'Étude
de la Philosophie Française Contemporaine

*Je dédie ce texte à Yves-Marie Visetti et à Bernard
Laks, mes deux guides dans les choses de l'esprit*

Résumé : Parmi les arguments invoqués contre le paradigme cognitiviste classique depuis plus d'une vingtaine d'années, il en est un qui lui reproche d'être au fond incapable de rendre compte de la variabilité culturelle, c'est-à-dire du changement aussi bien diachronique que dialectal des langues ou des formes culturelles en général. C'est notamment au nom d'un argumentaire de ce genre que les modèles connexionnistes, morphodynamiques ou de vie artificielle se sont développés, ainsi que les critiques de la linguistique générative depuis la sociolinguistique jusqu'aux approches herméneutiques et morphologiques contemporaines. On voudrait dans cet article montrer que le «structuralisme» tel qu'il s'est constitué dans les sciences de la culture entre Saussure, pour la linguistique, et Lévi-Strauss, pour l'anthropologie, n'a été rien d'autre qu'une tentative pour penser un certain type de fonctionnement, qui aurait la propriété singulière de faire émerger des formes (comme sont les signes linguistiques), mais des formes qui se transforment alors que les mêmes processus sont à l'œuvre. On montrera qu'il s'agissait là d'une manière très singulière et très actuelle de poser le problème de l'esprit: l'esprit est un mécanisme, mais qui à force d'appliquer les mêmes règles, finit par les transformer. Processus mentaux et formes culturelles variables sont inséparables. C'est l'idée même de simulation qui en ressort transformée. »

Mots clés : Intelligence artificielle ; Structuralisme ; Saussure ; Connexionisme ; Linguistique variationniste

Patrice Maniglier

Très ancien est le projet de faire de l'esprit, c'est-à-dire des phénomènes mentaux ou spirituels comme on voudra les appeler, l'objet d'une science positive. Longtemps, ce projet s'est confondu avec celui non pas tant de la psychologie, comme une histoire restrictive a tendance à nous le faire croire (cf. Gardner 1985), mais avec ce qu'on appellera selon les pays et les époques : les sciences morales et politiques, les sciences humaines, les sciences de la culture ou les sciences sociales. Comme l'ont montré très justement Stefano Franchi et Güven Güzeldere dans l'introduction de leur recueil, *Computational Minds and Mechanical Bodies* (cf. Franchi & Güzeldere, 2005¹), une perspective historique sur l'intelligence artificielle se doit d'intégrer des dimensions trop souvent négligées par les historiens de cette discipline, et la problématique de l'intelligence artificielle dépasse les limites du projet d'ingénierie connu sous ce nom et prend ses racines très loin dans l'histoire². Mais il ne sert à rien de remonter à Platon si l'on ne voit pas que

Montesquieu écrivit un livre dont le titre est *L'esprit des lois*, si l'on oublie que le concept d'esprit (*Geist*) est au cœur de l'idéalisme allemand avec Hegel, Schelling, puis Humboldt et jusqu'à Cassirer, que Comte propose une politique de l'esprit³, ou encore que Durkheim prétend relever le défi posé par Kant à la science, faire une théorie positive de la capacité humaine à dire des vérités et à agir moralement, en sociologisant ce que Kant attribuait à la philosophie transcendantale, et en montrant par là que ce que le philosophe considérait comme des conditions a priori antérieures à tout savoir pouvait tout à fait être l'objet d'une science positive⁴. Dans le temps comme dans l'espace, il ne suffit pas d'avoir la vue longue, il faut aussi avoir un champ de vision assez large, si l'on veut voir ce qui est à voir.

Il est vrai cependant que malgré cette longue histoire aux multiples dimensions, il ne vint à à peu près aucun de ces auteurs classiques, l'idée qu'on

* Ce texte est la version française d'un essai commandé par Stefano Franchi, à paraître en italien (Franchi & Bianchini, 2007) et en anglais (Franchi & Bianchini, 2008), sous le titre : « Processing Cultures, 'Structuralism' in the history of artificial intelligence ».

puisse appliquer à l'esprit lui-même des modèles *mécaniques*. Or on peut penser que ce qui définit l'espace contemporain des sciences de l'esprit, ce n'est pas telle ou telle théorie particulière, mais ce lien nouveau, et qui aurait paru contre-nature à la plupart des auteurs anciens, entre *machine* et *esprit*. On comprend dès lors la place que l'Intelligence Artificielle y joue. Descartes lui-même, le grand Descartes, dont se revendiqueront si volontiers certains parmi les plus brillants représentants de la « révolution cognitive », rejetait explicitement l'idée que l'esprit humain puisse avoir un caractère mécanique et donc être simulé par une machine. C'est qu'il avait, au fond, une conception très datée historiquement de la machine elle-même, sur le modèle des *horloges*, réagissant à des facteurs purement mécaniques, de sorte que si l'on pouvait imaginer simuler le comportement d'un animal qui réagissait à des forces extérieures (selon un modèle qui deviendra celui de l'arc réflexe), on ne pouvait le faire pour les hommes qui, avant d'agir, décomposent les stimuli et recomposent leurs comportements à partir de petites unités recombinaisons, de sorte qu'ils connaissent ce que la machine ne connaît pas : la liberté⁵. Or l'on sait que c'est au nom de ce modèle combinatoire même que le paradigme symbolique dans les sciences cognitives s'est imposé⁶, s'appuyant sur l'existence d'un nouveau type de machine, la machine universelle de Turing. Beaucoup ont donc pensé que l'objectivation théorique des phénomènes mentaux à travers des modèles mécaniques était inséparable d'un certain type de machine, l'ordinateur classique, et d'un certain type de théorie de l'esprit, celle qui le représente comme un calculateur opérant sur des représentations internes – modèle qu'on a appelé « *computo-représentationnel* » (Fodor, 2000).

Cependant, ce modèle théorique n'aurait sans doute pas connu la fortune dont il a bénéficié, s'il n'avait pas entretenu un rapport de circularité vertueuse avec les sciences humaines elles-mêmes. En particulier, la linguistique générative a contribué de manière décisive à crédibiliser le modèle de l'esprit comme un calculateur sur des représentations internes, en même temps qu'elle tirait argument de son interprétabilité (au moins théorique) dans ce modèle pour se justifier elle-même : une linguistique qui pouvait donner lieu à l'écriture de programmes sur ordinateur, qui, eux-mêmes, permettent de simuler certains aspects du comportement verbal humain de manière plus ou

moins satisfaisante, n'avait-elle pas sur ses concurrentes un avantage immédiat, si du moins l'on acceptait de se placer dans l'épistémologie constitutive des sciences cognitives depuis le fameux « test de Turing », qui fait de l'intelligence artificielle non pas seulement un projet technologique, mais aussi un terrain expérimental pour des *théories* à prétention scientifique sur l'esprit ? On posera donc, dans la suite de ce texte, que l'espace de ce qu'on appelle les « sciences cognitives » se définit par la polarité induite par ces deux exigences ou ce double terrain expérimental : d'une part l'esprit doit être justiciable d'une *simulation* par les moyens d'une forme d'intelligence artificielle, d'autre part, on n'aura pas saisi les phénomènes de l'esprit tant qu'on n'aura pas rendu compte aussi des phénomènes *culturels*. Bref, que le problème de l'esprit posé dans des termes contemporains est celui d'une simulation des phénomènes culturels.

La question est maintenant de savoir si ce lien favorise nécessairement le modèle cognitif dominant. Car, depuis quelques années, on constate une double contestation du modèle *computo-représentationnel*, tant du côté des modèles technologiques d'intelligence artificielle que du côté des théories linguistiques qui s'y rattachent plus ou moins lointainement. Il faut bien que la contestation soit double d'ailleurs, s'il est vrai que c'est dans ce cercle complexe entre simulations technologiques et modèles théoriques que se définit l'originalité du « problème de l'esprit » dans le contexte contemporain. Du côté technologique, la critique du paradigme symbolique s'identifie aux différentes approches *dynamicistes*, d'abord en termes de réseaux connexionnistes (Rumelhart, D. E., Mc Clelland, J. L. & the PDP Research Group, 1988), puis de modèles *morphodynamiques* et *néo-cybernétiques*, comme ceux de la théorie des catastrophes (Petitot, 1985a, 1985b) ou de l'*énaction* (Varela, 1989), et enfin les modèles robotiques de la vie artificielle (Brooks R. & Steels L., 1995)⁷. Du côté théorique, on pensera d'abord aux différents dépassements internes dont la linguistique générative a fait l'objet (en particulier l'approche « Principes et Paramètres »), puis aux contestations radicales qui se sont cristallisées dans les linguistiques cognitives qui ont essayé de changer véritablement le cadre de la théorie linguistique (Lakoff, 1987; mais surtout Langacker, 1987), jusqu'au retour contemporain des linguistiques de corpus et des approches *herméneutiques* du langage⁸.

De nombreux arguments tant technologiques que théoriques motivent cette double contestation. Il en est un pourtant sur lequel je voudrais ici m'arrêter, précisément parce qu'il est assez rarement mis en avant : c'est celui qui dénonce justement l'incapacité du paradigme symbolique à rendre compte d'une propriété centrale des phénomènes mentaux, à savoir leur caractère *culturel et donc variable*. Nous allons en effet montrer, dans un premier temps, en quel sens on peut dire que le modèle symbolique est impuissant à rendre compte de la variation culturelle, et nous exposerons les critiques récentes qui en ont été développées d'abord de l'intérieur de la linguistique et dans un cadre qui ne prétendait pas rompre avec les concepts fondamentaux de la linguistique générative (celui de la sociolinguistique de Labov), puis au sein des modélisations d'I.A. proposées soit dans un cadre encore computationnel (par le connexionnisme), soit en posant les problèmes d'emblée du côté de la robotique et en pointant vers une autre épistémologie de l'intelligence artificielle (Luc Steels).

Dans un deuxième temps, on montrera que cette critique n'est au fond qu'une redécouverte (parfois parfaitement consciente et assumée, souvent inconsciente et involontaire) d'un projet qui fut précisément celui que refoula, dans les sciences humaines, la grammaire générative – à savoir le projet structuraliste. Celui-ci en effet, contrairement à l'idée reçue qu'on en a trop souvent, n'a pas consisté en une recherche des invariants communs à chaque communauté linguistique (*la langue, c'est-à-dire la structure du français par-delà la variation dialectale ou stylistique des manières de parler*), mais au contraire en un effort pour comprendre pour quelles raisons les mêmes mécanismes qui nous permettent d'apprendre une langue (ou une autre dimension de la vie culturelle), nous conduisent inévitablement à la transformer ou plutôt à la laisser se transformer à travers nos usages sans même que nous nous en rendions compte. Si on peut faire du structuralisme, dont la véritable date de naissance comme mouvement transversal aux sciences humaines est contemporaine des années d'apparition de l'Intelligence Artificielle à proprement parler, puisqu'on peut la fixer à la rencontre de Jakobson et de Lévi-Strauss à New York en 1943⁹, un précurseur de ces recherches contemporaines, c'est précisément parce qu'il privilégie une approche de l'esprit en termes de *mécanismes* (bien qu'il ne définisse pas de modèle

technologique précis), mais tout en tenant qu'il n'y a d'esprit qu'au sein de cultures aussi bien socialement qu'historiquement en variation. En somme, à travers ce petit montage archéologique, on se propose de montrer que le paradoxe qui résume au plus près le défi de l'intelligence artificielle, depuis son origine, n'est autre que celui-ci : penser une machine à histoire, dépasser l'opposition entre le mécanique et le culturel, penser l'esprit comme une machine vouée à générer des cultures avec ce que cela implique d'incontrôlable variabilité – bref des machines qui, loin d'être contrôlées et contrôlables, instruisent de nouveaux espaces d'imprévisibilité, de contingence, d'histoire...

Le paradigme symbolique et ses critiques variationnistes

Le générativisme aux prises avec la variation historique et dialectale

On pourrait s'étonner du reproche de fixisme adressé à la linguistique chomskyenne qui a notamment fait de la *créativité* un trait essentiel, définitoire même, du langage, s'opposant ainsi d'ailleurs frontalement au structuralisme alors en vogue. Mais Chomsky oppose deux sortes de créativités : la *créativité gouvernée par les règles*, qui ne fait guère sortir de la langue qu'on parle ; et la *créativité changeant les règles*. Or le point du débat porte précisément sur la légitimité de cette distinction. Certes, rien ne rend la variation culturelle au sens strict *impensable* dans le modèle génératif. Mais la manière même dont ce changement est théorisé, la représentation des mécanismes du changement, le soustrait pour ainsi dire à l'observation empirique. Car changer de langue, c'est, dans ce cadre théorique, nécessairement, changer les règles. Ce ne saurait être l'application des mêmes règles qui conduit à les transformer. Le seul fait de se représenter le fonctionnement du langage comme l'application de règles à caractère impératif et univoque opérant sur des représentations mentales (conformément au modèle symbolique développé conjointement par Simon et Newell à la même époque) empêche de penser une transformation continue et involontaire des langues. Le changement ne peut être que discontinu : on passe d'un coup d'une langue, immobile, à une autre, tout

aussi immobile, mais entièrement différente puisqu'elle n'a pas les mêmes règles et que ce sont les règles qui définissent l'identité de chaque langue. Tout se passe donc comme si la langue n'était jamais vraiment *dans le temps*. Dans un article fondateur de la sociolinguistique, Weinreich, Labov et Herzog (1968) en faisaient le constat paradoxal : dans le cadre générativiste, les langues que nous parlons ne peuvent pas, par nature, avoir d'histoire : « en effet, s'il est nécessaire qu'une langue soit structurée pour fonctionner efficacement, comment les gens peuvent-ils continuer à parler pendant qu'elle se transforme, c'est-à-dire pendant qu'elle traverse des périodes de moindre systématisme ? ¹⁰ ». Si l'on voulait pousser le paradoxe, on pourrait dire : les langues donc *ne changent* pas vraiment ; elles se succèdent.

Le problème ne porte d'ailleurs pas seulement sur la variation diachronique, mais aussi sur la variation dialectale. On sait en effet qu'un des corollaires de la définition des langues comme ensembles de représentations de base et de règles de transformation opérant sur ces représentations mentales et permettant d'engendrer l'ensemble infini de toutes les performances possibles, est la thèse dite de l'*homogénéité* de la compétence, qui suppose que tous les sujets parlent la même langue et qu'ils n'en parlent qu'une seule à chaque fois. La variation dialectale est interprétée comme coexistence d'une pluralité de langues qui se ressemblent, certes, mais qui sont numériquement distinctes. Ainsi, on admettra bien que l'étude de l'anglais qu'on trouve dans les livres de grammaires vaut surtout pour l'anglais standard et normalisé des speakerines de la télévision ou des professeurs de linguistique, et qu'il existe d'autres anglais, comme celui des noirs new-yorkais étudié par Labov. Mais on dira qu'il s'agit au fond de deux *langues* différentes, aussi dignes d'études l'une que l'autre, mais qui ne partagent pas les mêmes règles. On dira ainsi qu'en Black English Vernacular (BEV), il existe une règle qui stipule qu'il n'y a pas de marque de la troisième personne de l'indicatif présent (*she write*) : ce n'est pas une faute, c'est une autre règle, corrélative d'autres et cohérente avec un système.

Or Labov a fait remarquer que cet usage d'un dialecte ou de l'autre n'a rien d'exclusif, et qu'un même individu peut utiliser préférentiellement, mais non pas systématiquement, le BEV dans tel contexte, dans une mesure d'ailleurs statistiquement évalu-

lable. De sorte qu'on peut noter qu'il n'y a pas tant changement de code, que mise en variation des règles même des langues, qui tantôt sont appliquées, tantôt sont suspendues ou remplacées par d'autres, selon des taux relativement stables d'occurrences dans des contextes identiques. C'est pour rendre compte de ces phénomènes qu'il élabore le concept quasi-paradoxal de règles variables : non pas que les règles changent à mesure qu'elles s'appliquent, mais elles peuvent être suspendues selon une fréquence calculable, ce qui montre qu'elle ne s'applique pas de manière impérative et mécanique, mais sous une forme statistique qui relève plus de la *régularité* sociologique que de la *règle* symbolique. De plus, Labov notait qu'un jeune garçon de douze ans peut permuter environ 18 fois d'un « code » dans un autre dans une séquence aussi courte que celle-ci, par laquelle il explique un jeu, le skelly : “An’ den like if you miss onesies, de othuh person shoot to skelly ; if he miss, den you go again. An’ if you get in, you shoot to twosies. An’ If you get in twosies, you go to threesies. An’ if you miss threesies, then the person tha’ miss seklley shoot the skellies an’ shoot in the onesies: an’ if he miss, you go f’om threesies to foursies ¹¹.” Quel sens y a-t-il à parler ici de deux langues différentes ? Le concept de dialecte n'est-il pas mal construit ? Ne doit-on pas plutôt dire, avec Weinreich : « une maîtrise quasi innée de structures hétérogènes n'a rien à voir avec la connaissance de plusieurs dialectes ni avec la simple performance, mais fait partie de la compétence linguistique de l'individu unilingue. » (art. cit., p. 100-101) ? Weinreich était clair : il faut « rompre l'identification entre structure et homogénéité », et penser les structures linguistiques non pas comme la somme des pratiques individuelles, ni comme une entité séparée de ces pratiques, mais comme quelque chose d'intrinsèquement hétérogène, d'intrinsèquement variable, d'intrinsèquement multiple.

Mais il fallait sans doute pour cela renoncer tout à fait au vocabulaire de la règle et donc au soutien idéologique que pouvait apporter le paradigme symbolique en intelligence artificielle. Ce fut, comme l'on sait, l'enjeu de ce qu'on appela les « linguistiques cognitives », en particulier avec les synthèses des années quatre-vingt, celle de Lakoff (1987) d'abord, et celle de Langacker (1987) surtout (qui est plus systématique, plus scientifique, plus radicale aussi par bien des aspects), que de montrer qu'une autre linguistique était possible, qui partagerait le projet

d'inscrire le langage dans l'ensemble des phénomènes cognitifs, mais qui aurait un autre concept de la cognition que celui du traitement formel des représentations mentales. Pour les linguistiques cognitives, connaître, ce n'est pas calculer, c'est reconnaître, c'est-à-dire d'une certaine manière percevoir, catégoriser, réidentifier. Un des arguments en faveur de l'abandon du concept de règle rejoignait les remarques de la sociolinguistique : l'appartenance d'une forme à une langue n'est pas aussi tranchée que le modèle de la règle (qu'on applique ou qu'on transgresse) ne l'exigerait. Par exemple, pour reprendre un exemple de Langacker, on hésitera à juger « non grammaticale » des expressions comme : « un vélo est désiré par moi », « ce vélo est désiré par mon frère », (par opposition à la claire acceptabilité de « ce vélo doit être désiré par quelqu'un »). Ou encore qu'on sentira qu'il y a des tensions de plus en plus fortes dans l'acceptabilité des phrases suivantes :

Sheila was kicked by her mother.

? *Sheila was wanted by her mother.*

?? *Sheila was resembled by her mother*¹².

Pourtant, trancher sur de tels exemples est essentiel, puisqu'il s'agit de décider du fonctionnement du *passif*. L'appartenance d'une forme à une langue n'est donc pas de l'ordre du tout ou rien, mais du plus ou moins. C'est en se fondant sur des exemples semblables que les linguistiques cognitives proposèrent de ne plus définir la langue comme un ensemble de règles appliquées sur des symboles ininterprétés, mais comme ensemble de formes symboliques qui émergent de routines cognitives très générales (et non pas spécifiques au langage), et qui ne sont ni « formelles » ni « aveugles au sens », mais bien au contraire directement engagées dans un certain type de prise symbolique sur la réalité perçue. Ainsi, ici, « *to resemble* » souffre mal le passif pour des raisons sémantiques, parce qu'il a beau être un verbe transitif en anglais, il n'est pas un verbe d'action. La grammaire, dira Langacker (1987) est une forme de *schématisation* du réel. Le langage apparaît ainsi comme une forme particulière de la *perception*, une manière de voir et de faire voir, et non pas comme un module de la capacité de *calcul*. La linguistique ne communique plus avec l'ensemble des sciences cognitives par l'analogie formelle des processus à l'œuvre dans

le traitement syntaxique ou formel des représentations du langage, mais par celui, plus général, de la *reconnaissance des formes*. Une des grandes forces des linguistiques cognitives, et en particulier de celle de Langacker, est d'avoir posé que les catégories grammaticales même les plus hautes, loin d'être innées, sont des effets émergents de la mise en ordre d'une grande diversité de patterns linguistiques. On retrouvera, à propos de la linguistique de Saussure, une intuition extrêmement proche.

Mais certains pourraient se demander si ce problème n'est pas aussi celui auquel répond le profond remaniement que la linguistique chomskienne a connu avec l'approche dite « Principes et Paramètres »¹³, qui, tout en abandonnant le concept de règle, maintient une représentation du langage comme processus de traitement de symboles ininterprétés. Cette nouvelle vague de la grammaire générative, après tout, sacrifie tout à fait la notion de langue. Au lieu d'avoir des ensembles de règles, il existe des contraintes universelles innées (et donc présentes dans chaque organisme humain individuel), qui définissent la Grammaire Universelle, sont multiples, pas forcément cohérentes, les « langues » se caractérisant par leur manière de paramétrer ces contraintes et de résoudre les problèmes de compatibilité qu'elles posent¹⁴ (cf. Chomsky 1991). Mieux : en réalité, il n'y a pas à proprement parler de « langue » : on communique directement, si j'ose dire, de grammaire universelle à grammaire universelle. Les « langues » ne sont que des épiphénomènes, comme un nuage vague émergent de la confusion de la bataille, mais rien de substantiel en soi. La variation est intégralement du côté de la performance, et les variations significatives sont celles qui reposent sur une nouvelle hiérarchie des contraintes.

On voit pourtant que ce modèle comporte un défaut symétrique au précédent : au lieu d'hypostasier les langues dans des règles trop rigides, il en nie tout à fait l'existence : entre la faculté du langage et les actes de parole, il n'y a rien. On ne discutera pas cependant ici sur le fond ce modèle ; on relèvera simplement un paradoxe philosophique, qui vaut pour les différents états de la linguistique générative. D'une manière générale, on sait que pour Chomsky, une langue n'est rien d'autre qu'une théorie intérieure qu'un sujet construit à propos de la langue en usage dans la collectivité à laquelle il appartient. Les langues (au sens des objets

théoriques que reconstruit le linguiste) ne sont donc jamais en tant que telles des faits sociaux; ça reste des phénomènes purement psychologiques et intérieurs à chaque individu. D'où l'extrême psychologisation et individualisation du fait linguistique, et l'usage de l'expression Langue Interne (LI) pour désigner la langue comme objet de la linguistique (cf. Pollock, 1997, ch.1). Cette position semble cependant conceptuellement paradoxale. Car si connaître une langue c'est connaître la langue en usage dans la collectivité, mais que cette langue en usage n'est rien d'autre que la langue interne d'un tel et d'un tel, cela veut dire que je connais quelque chose qui n'est soi-même rien d'autre qu'une connaissance d'autre chose, voire une connaissance de la connaissance même que je suis censé avoir... Bref, définir la langue comme un système de connaissances *intérieures*, c'est la définir comme connaissance *de rien*. On voit bien le paradoxe d'une connaissance spéculaire absolument sans critère : c'est avec de tels arguments d'ailleurs que Wittgenstein, dans *Les Investigations Philosophiques*, remettait en question la pertinence de la notion de « calcul mental » : il est impossible de suivre une règle *tout seul*, car en ce cas nous n'aurions plus aucun critère pour évaluer l'application de cette règle ou son manquement; il n'y a de règle que là où il y a une collectivité et des usages collectifs¹⁵... Il faut donc que ce soit la langue *externe* que les individus connaissent; mais pour cela, il faut que cette langue externe ait une autonomie, et que ce soit *elle* que les linguistes étudient, et non pas seulement les vues que tentent d'en prendre, chacun de leur côté, les sujets parlants.

Mais quoi qu'il en soit de la pertinence de ces objections philosophiques, il se trouve que les modèles en termes de « principes et paramètres » ne sont de toutes manières plus vraiment un argument en faveur d'une approche de la cognition sur le modèle du calculateur symbolique. On sait en effet que les architectures à la von Neuman exigent un calcul séquentiel : des représentations internes sont transformées par l'application d'un algorithme qui efface ou remplace certains caractères par d'autres en ne se fondant que sur des contraintes syntaxiques, c'est-à-dire mécaniques. On sait aussi que Jerry Fodor aura généralisé ce mode séquentiel de traitement de l'information avec le concept de modularité, en représentant l'ensemble de la cognition comme un processus séquentiel (Fodor, 1983). Or il est clair qu'un modèle linguistique en termes de contraintes agissant en parallèle pour déter-

miner comme une solution optimale la bonne représentation linguistique (sans qu'il y ait véritablement de « transformation », au sens classique, puisque c'est bien ce concept qui a été à l'origine de toutes les contestations), qu'un tel modèle, donc, ne trouve plus dans le paradigme symbolique le cercle vertueux qui caractérisait les sciences cognitives des années soixante. Fodor lui-même en a d'ailleurs fait le constat très clair récemment, contre ceux qui croient que les sciences de l'esprit sont enfin arrivées très près du Graal (cf. Fodor, 2000). Il n'y a pas d'abord le module phonologique (qui applique ses règles pour transformer une représentation phonétique en représentation phonologique), puis le module morphologique, puis le module syntaxique, puis le module sémantique ou pragmatique; il y a une collaboration de l'ensemble de ces dimensions dans la détermination d'une *bonne forme*, dont on ne dira pas qu'elle est acceptable en vertu de règles déterminées, mais simplement qu'elle peut être formée malgré l'exercice de contraintes hétérogènes. Or un tel modèle réclame un mode de calcul non plus séquentiel comme le privilégiait le paradigme symbolique où chaque règle s'applique après l'autre, mais parallèle. De ce point de vue donc, il y a rupture entre la linguistique générative et le paradigme cognitiviste pour lequel elle avait tellement fait. Ce n'est d'ailleurs pas un hasard si une des approches les plus convaincantes pour rendre compte du changement linguistique en termes de contraintes et d'optimisation des « principes » est due à un auteur qui se trouve être une des grandes figures du connexionnisme, c'est-à-dire d'un des rivaux les plus sérieux qu'ait connu le paradigme symbolique en intelligence artificielle – je veux parler de Paul Smolensky et de sa théorie de la phonologie harmonique (avec Géraldine Legendre), bientôt généralisée dans la Théorie de l'Optimalité (avec Prince)¹⁶. Mais il est vrai qu'il s'agit là déjà d'un argument sur les meilleures modélisations des phénomènes cognitifs, puisqu'il s'agit de savoir quel type de calcul, séquentiel ou parallèle, est le plus à même de modéliser les données linguistiques. Il vaut donc mieux passer directement à une discussion de l'histoire récente de l'intelligence artificielle face à ce défi : modéliser la variation culturelle.

Machine à culture :

Il se trouve que les objections rapides que nous venons de faire à un modèle qu'on peut dire, gros-

sièrement, « computo-représentationnel » en linguistique se retrouvent dans la critique faite au paradigme symbolique classique en intelligence artificielle. Là aussi, l'incapacité à rendre compte de la variabilité inhérente à certains comportements est apparue comme un argument en faveur d'un changement radical de paradigme, qui passait aussi par l'abandon des notions de règle et de représentation interne.

Parmi les motifs du succès que connurent les modèles connexionnistes dans les années quatre-vingt, on a trop rarement mentionné cette supériorité qu'ils pouvaient afficher dans la modélisation du changement. Certes, cet argument n'apparaît pas de manière explicite dans la somme du connexionnisme des années quatre-vingt, *Parallel Distributed Processing* (cf. Rumelhart D.E., McClelland J. L., and the PDP Research Group, 1988). Il intervient pourtant au cœur de la théorie, mais d'une manière discrète : sur la question de l'apprentissage. Parmi les principaux acteurs du renouveau des réseaux des années quatre-vingt, Jeffrey Elman est sans doute un de ceux qui ont le mieux compris l'importance de cet aspect, tout en se revendiquant d'ailleurs d'une forme de néo-structuralisme incarné pour lui par l'œuvre de Piaget¹⁷. Il l'exposa notamment dans un article synthétique écrit avec Elizabeth Bates, « Connectionism and the study of change ». L'idée est au fond fort simple : il s'agit de montrer que l'apprentissage est non pas l'acquisition de nouvelles règles ou de nouveaux objets sur lesquels appliquer ces règles (c'est-à-dire des représentations, par exemple d'items lexicaux si on est dans un cadre linguistique), mais une véritable *transformation de soi* c'est-à-dire des structures mêmes de fonctionnement de l'esprit. Il s'agit de dépasser l'opposition simpliste entre l'empirisme et l'innéisme, avec, d'un côté, des contenus qu'on se contenterait d'engranger, et, de l'autre, des formes dont on se demanderait si elles sont données par la nature, ou creusées par les contenus même de l'expérience vécue. C'est la distinction entre forme et contenu qu'il s'agit de remettre en question. En réalité, la forme elle-même change, et surtout *permet son propre changement*. Le système doit être ainsi fabriqué qu'il peut se transformer, et non pas seulement en subissant de l'extérieur le changement, comme une table de boucher qui se creuse à force d'avoir été frappée, mais en l'organisant positivement, en le préparant. On ne sera jamais satisfait d'une théorie de l'esprit qui ne serait pas capable de rendre compte du fait

que l'architecture même de l'esprit l'ouvre à cette capacité de transformation. Ce sont les *mêmes* mécanismes qui permettent à la machine de fonctionner au présent, et qui la conduisent à transformer l'architecture même de ces mécanismes. C'est l'application des mêmes procédures qui entraîne leur transformation. Tel devrait être le réquisit minimum de toute théorie de l'apprentissage. Apprendre, ce n'est pas acquérir *des contenus*, c'est se transformer *soi-même*. Ce n'est pas seulement ajouter dans des cases toutes prêtes, c'est modifier l'architecture de l'ensemble. « Learning is structural change », écrivaient Bates et Elman (1993). Il faut donc avoir une conception de la « structure », qui, par nature, soit ouverte à sa propre transformation.

Or c'est la réponse à cette question que Elman croit trouver dans les modèles connexionnistes. Pour ceux-ci, en effet, l'apprentissage n'est pas une propriété secondaire, mais l'essence même de leur fonctionnement. Leur réalité *est* le changement structural. Il n'y a pas en eux de différence *de nature* entre une performance ordinaire et un changement de compétence. Notons aussi que Elman insiste souvent sur le fait que ce sont les mêmes architectures, au fond, les mêmes mécanismes, qui permettent de donner lieu à des comportements très variés : « Connectionists models illustrate how the same learning mechanism can give rise to behaviors which differ dramatically at various points in time. » (Elman, 1996, p. 43). Il ne s'agit pas en effet d'opposer d'un côté une théorie des structures innées et donc invariantes, et de l'autre une théorie de l'acquisition par l'expérience et donc aussi variables que le sont les contextes : il s'agit au contraire de construire une théorie des structures ouvertes, des structures capables de changer. Qu'est-ce qui, dans le connexionnisme, permet de l'envisager ? Tout d'abord l'abandon des notions de représentations et de règles. On ne stipule pas dans l'écriture du modèle les règles auxquelles il doit obéir ; on le laisse les trouver. De même, il n'y a pas de mémoire spéciale pour les items, qu'il s'agirait ensuite d'aller chercher lorsque des problèmes de reconnaissance se posent. La mémoire est toute entière dans les connexions, dans la manière dont elles ont été corrigées. Là encore, on doit dire : la mémoire n'est pas affaire de contenus objectifs (comme les colombes dans l'ironique volière du *Théétète* de Platon), mais d'identité subjective. Ce qui ne veut pas dire qu'un système

capable d'un tel apprentissage n'aurait pas une architecture bien déterminée à l'origine et qu'il serait comme la tabula rasa des empiristes. Au contraire, il a une architecture tout à fait spécifique : elle consiste en un réseau de neurones multi-couche doté d'une règle dite de rétropropagation qui, à partir de l'évaluation d'une « erreur » (c'est-à-dire d'un écart entre une réponse produite et ce qu'aurait été la bonne réponse), modifie les connexions entre les neurones. Les réseaux apprennent non pas en acquérant des règles ou des symboles nouveaux, mais en changeant les connexions, c'est-à-dire en se transformant eux-mêmes. La reconnaissance de formes par exemple, ne consiste pas à décomposer les stimuli en un certain nombre de traits et à chercher à quelle catégorie déjà connue correspond ce stimulus dans la mémoire, mais à laisser le réseau réagir à une grande variété de patterns d'activation d'entrée par un petit nombre de patterns d'activation de sortie, qui dépendent de la structure des connexions. De plus, un comportement n'est pas correct ou incorrect, mais plus ou moins correct, plus ou moins semblable. De ce fait même, un tel système peut tenir compte de situations qui ne sont pas prévues et qui constituent au fond des transgressions à son système catégoriel. Un exemple très simple de ce phénomène est celui que donnent Rumelhart, Smolensky, McClelland et Hinton, dans *PDP* : un logiciel de reconnaissance des pièces d'un appartement (chambre, salon, cuisine, etc.) sur la base de certains traits qui sont essentiellement des meubles (présence ou absence d'un lit, présence ou absence d'un téléviseur, etc.), trouvera une configuration stable pour quelque chose qui est en fait un « studio », voire quelque chose de beaucoup plus complexe (avec le lit dans la cuisine, etc.). Un module cognitif classique aurait simplement été dans l'incapacité de classer l'objet, et donc de se stabiliser, alors que le modèle connexionniste n'a eu aucune peine à générer ce qui devient au fond une nouvelle catégorie, précisément parce que les catégories ne sont jamais inscrites dans la mémoire, mais émergent localement, de chaque opération concrète¹⁸.

Mais nul n'a été plus clair et plus profond dans cette défense du connexionnisme contre le cognitivisme symbolique classique au nom de sa capacité à modéliser, cette fois clairement, le caractère intrinsèquement *culturel* et donc variable de la cognition humaine, que le philosophe Paul M. Churchland. On peut se reporter au texte à tous égards remar-

quable intitulé « Reductionism, Connectionism and the plasticity of human consciousness » (in Churchland, 1989). Il s'agissait de montrer qu'il n'y avait pas d'opposition entre une approche naturaliste et même matérialiste de l'esprit (que Churchland défend), réduisant l'esprit au cerveau et contestant à la psychologie naturelle (folk psychology) toute validité, et l'idée selon laquelle il n'y a pas de phénomènes mentaux sans culture. Mais l'intérêt de ce texte de Churchland est d'avoir remarqué que la véritable objection à laquelle un naturalisme exigeant est mis au défi de répondre ne porte pas tant sur le problème de l'explication des comportements que, de manière beaucoup plus immédiate et radicale, sur le problème de l'observation même des phénomènes qui constituent nos actions de tous les jours : les phénomènes culturels auxquels nous sommes sensibles, par exemple un regard, une parole, un geste, etc., ne sont pas constitués de traits physiques observables, isolés les uns des autres et mesurables par un appareil analogique simple, comme peuvent l'être les paramètres du monde auquel serait sensible, disons, la bactérie magnétotactile dont Dretske a fait le modèle non pas du plus simple organisme cognitif comme on aurait tendance à le croire, mais plutôt du plus complexe des organismes non-cognitifs¹⁹. Les *phonèmes* sont un excellent exemple de ces traits « complex, subtle, context sensitive or stimulus transcendent » (Churchland, 1989, p. 133). Rien dans le signal phonétique comme tel ne peut être reconnu comme un trait de voisement (vibration de cordes vocales qui distingue le b du p, le d du t, etc.). Car, au niveau phonétique, le trait de voisement est si variable en fonction des contextes, si distribué dans le signal, si altéré, qu'il est à peu près impossible d'imaginer qu'un détecteur de traits spécialisé (comparable à un thermostat enregistrant les variations de températures) puisse identifier la présence ou l'absence de ce trait. Dès lors, vous avez beau mettre tous les capteurs que vous voulez, la machine aura bien du mal non pas à *interpréter* les données, mais bien à les *percevoir*. Or, remarquait Churchland, un réseau de neurones multi-couche peut très bien, lui, être entraîné à reconnaître de telles variations, précisément parce qu'il analyse l'ensemble du signal en parallèle et de manière distribuée, non pas en traitant de manière isolée chaque trait. Dès lors, un argument majeur en faveur des modèles connexionnistes était non pas leur crédibilité neurologique comme on l'a souvent dit, en particulier en croyant restituer

l'opinion « matérialiste » de Churchland, mais bien leur capacité à modéliser la *capacité culturelle* des hommes. Les hommes peuvent apprendre *n'importe quelle culture*, précisément au même sens que les réseaux sont capables d'apprendre n'importe quelle différence aussi arbitraire soit-elle.

Qu'on ne pense pas néanmoins qu'il s'agisse pour nous de fournir ici un énième plaidoyer en faveur des modèles connexionnistes. Il s'agit au contraire de montrer à quel *problème* trop souvent méconnu ces modèles ont répondu, afin d'ouvrir une autre perspective historique au sein des sciences cognitives. Parmi les efforts les plus explicites pour lier une théorie renouvelée de l'intelligence artificielle et une linguistique variationniste appuyée sur de solides fondements cognitifs, on doit d'ailleurs mentionner les modèles de vie artificielle, et plus précisément les *cultures de robots* proposés par Luc Steels et appliqués par lui-même et Pierre-Yves Oudeyer à la phonologie diachronique (Steels & Oudeyer, 2000). Alors que les modèles connexionnistes laissent en dehors de leur considération le problème de la *teacher* (qui apprend au réseau la différence entre une « bonne » et une « mauvaise » réponse?), de sorte que le *sens* des différences qu'il est capable de « mémoriser » est totalement étranger au réseau, les propositions de Luc Steels ont l'intérêt de partir d'emblée du projet de mettre en réseau les ordinateurs eux-mêmes, et donc de ne pas décider d'avance ce qu'il y a à apprendre. Ainsi, les traits phonologiques ne sont pas considérés comme des sortes de données extrinsèques et intangibles, mais bien comme des résultats émergents de l'interaction de plusieurs détecteurs de traits (dont on peut, si on le veut, imaginer qu'ils soient eux-mêmes extraits par des réseaux) et des validations réciproques que les agents se donnent. L'esprit du modèle est simple et l'exemple phonologique peut encore une fois servir de point de départ : il s'agit d'un jeu de l'imitation, dans lequel un agent A produit au hasard un son codé en termes articulatoires, un agent B reçoit ce son perçu en le codant en termes acoustiques ; puis B essaie de reconvertir ce son perçu en un son produit dans l'espace articulatoire, en opérant au hasard, A donnant un feed-back positif si il estime que c'est le même (du moins dans une certaine marge de ressemblance), et négatif si il lui semble trop s'écarter. On voit qu'il s'agit d'un problème de « inverse mapping » entre l'espace acoustique et l'espace articulatoire. Une règle de correction permet

de faire varier les performances et d'obtenir progressivement des convergences entre ces productions au départ purement aléatoires. Le modèle ouvertement revendiqué est celui de la théorie darwinienne de l'évolution, avec ses variations individuelles et son principe de sélection. L'intérêt du modèle est qu'aucune catégorie phonologique n'est donnée d'avance : les catégories émergeront des interactions entre les agents. Des discontinuités apparaîtront dans les espaces acoustiques et articulatoires de sorte que l'un fera la cartographie de l'autre et réciproquement. On s'approche ici très clairement de la modélisation de l'émergence d'une *culture*, du moins si l'on admet qu'une culture puisse être considérée comme un *milieu* sensible constitué de paramètres fins qui ne correspondent à aucune fonction linéaire d'un des cinq sens (ou quelque soit le nombre qu'on leur donne), et qui ne reposent que sur les validations réciproques que se donnent des agents cognitifs capables de décomposer les signaux continus et variables en unités discrètes et invariantes²⁰.

On comprend alors que ces catégories ne soient pas stables, puisqu'elles ne reposent sur aucune nature extrinsèque, sur aucune nécessité naturelle, mais uniquement sur l'accord – ou le *sentiment* d'accord (car il n'y a, encore une fois, aucune mesure extérieure, physique, du consensus²¹) – qu'ont les agents vis-à-vis des catégorisations les uns des autres, à partir d'un corpus généré de manière aléatoire. Certes, dans un modèle à deux agents, réduit à un problème de inverse mapping un peu trop simple comme celui que nous venons d'évoquer, le système atteindra rapidement un état stable. Mais dans des systèmes multi-agents de large dimension, combinant un grand nombre d'interactions, de nombreuses solutions concurrentes aux mêmes problèmes coexisteront, les agents innoveront souvent précisément parce qu'ils n'ont pas une vue globale du réseau, et ces innovations se répandront dans certaines conditions de sorte que l'évolution de tels systèmes est inscrite dans leur mode même de constitution²². Il n'y a rien, si l'on peut dire, à *rajouter* pour penser la variation des répertoires de formes : les mécanismes qui président à la constitution de ces systèmes sont les mêmes que ceux qui conduisent à leur transformation. Pas d'ajout de règle, pas d'action transformatrice particulière : juste le fonctionnement normal d'une population cognitive qui ne cesse de transformer ses propres milieux de vie, et cela involontairement. Il est de la

nature de tels systèmes catégoriaux d'être ouverts à la variation, parce qu'ils sont avant tout des résultats localement émergents d'un grand nombre d'interactions. Telle serait donc la leçon : pour modéliser les phénomènes cognitifs, il faut peut-être *partir* de la notion même de culture.

L'intelligence est collective. L'esprit n'existe pas ailleurs que dans la masse. Il n'est pas dans ma tête, ni dans la vôtre, mais *entre nous*. Il dépend de notre capacité précisément à nous comprendre, c'est-à-dire en réalité à créer des milieux sensibles dont les paramètres pertinents ont été constitués par et à travers notre échange. En cela d'ailleurs, l'idée de simuler l'intelligence en simulant l'émergence de sociétés artificielles qui créent ce qu'il faut bien appeler des cultures (c'est-à-dire des régularités efficaces qui ne valent pourtant que pour une collectivité déterminée dont les limites, au reste, ne sont pas tranchées²³), passe brillamment le test même de Turing : celui-ci ne nous disait-il pas que l'esprit est quelque chose qu'on attribue à des agents qui *disent ce qui se dit* dans des contextes et des circonstances déterminés, et, pourrait-on dire en généralisant, *qui font ce qui se fait*, c'est-à-dire que l'esprit est intrinsèquement lié à la capacité de partager des conventions non nécessairement explicites et des usages (en l'occurrence, dans le test de Turing, des usages concernant la différence des sexes telles qu'elle peut filtrer dans le langage, qu'on doit savoir « imiter »²⁴) ? Tel serait donc l'horizon actuel de l'intelligence artificielle au moment présent de son histoire : simuler l'émergence des phénomènes culturels à travers des collectivités²⁵.

Ce diagnostic, à la vérité, est aujourd'hui assez partagé : la « mémétique » n'est après tout rien d'autre qu'un tel projet formulé d'une manière parfois un peu rustique, et de nombreuses théories sociales de l'esprit ont été développées récemment en philosophie (Dennett, 1993 ; Brandom, 1994 ; Searle, 1998 ; Descombes, 1996), corrélativement à une réévaluation de l'intérêt des philosophies de « l'esprit objectif » comme celles de Hegel (par Brandom par exemple) ou de Humboldt (par Descombes par exemple). Certes, il y a de grandes différences entre ces philosophies, leurs argumentaires sont incompatibles et elles ne partagent même pas une position analogue quant au statut et à l'intérêt de la modélisation pour la pensée philosophique. Mais elles ont en commun de soutenir, pour des raisons et avec des

moyens divers, que l'esprit est inséparable des phénomènes sociaux ou culturels. Il n'est pas cependant dans notre intention de proposer ici une thèse philosophique sur ce problème, mais uniquement une autre mise en perspective historique d'une certaine rencontre entre la problématique d'une hypothétique science de l'esprit (que la notion d'« intelligence artificielle » porte avec elle), et les sciences humaines, en mettant en évidence un espace méconnu pour cette rencontre, celui de la modélisation des phénomènes culturels. Car il se trouve que cet espace a été ouvert, il y a bien longtemps déjà, à travers ce mouvement qui fut connu – mais aussi méconnu – sous le nom de « structuralisme ». Y revenir, pourtant, permettra de préciser ce qui se joue à travers ce geste qui fait de la simulation des phénomènes de type culturel la *crux experimenti* de l'intelligence artificielle.

Fonction symbolique et variabilité culturelle : une nouvelle lecture du structuralisme

Quand bien même on nous accorderait que le défi d'un calcul capable d'engendrer des systèmes de formes essentiellement ouverts à la variation est un enjeu central de la recomposition des sciences cognitives en cours depuis la perte d'hégémonie du paradigme symbolique, on pourra se demander en quoi ce diagnostic permettrait de réinscrire ce mouvement théorique propre aux sciences humaines que fut le structuralisme dans une archéologie de l'intelligence artificielle, et quel bénéfice ce regard historique pourrait apporter à la recherche contemporaine.

Le signe, être d'esprit, être d'histoire

Que le structuralisme ait prétendu à fournir une théorie de l'esprit humain, on l'admettra sans doute aisément : ce fut un objectif explicite, aussi bien pour Jakobson que pour Lévi-Strauss. Mais on se souviendra aussi que c'était précisément par son refus de l'histoire et par son espoir d'expliquer un certain nombre de faits culturels par des contraintes d'ordre purement formel, c'est-à-dire, au bout du compte, systémiques, qu'il semblait s'associer à l'épopée des sciences de l'esprit. Quant à ses rapports avec l'intelligence artificielle, ils furent presque anecdotiques²⁶. Quant à la linguistique, ne sait-on pas que le geste constitutif du structuralisme y fut la rupture avec la

problématique strictement historique de la linguistique du dix-neuvième siècle? Qui a fait plus que Saussure pour enraciner dans la culture linguistique du vingtième siècle l'interdit de recherche sur la question si controversée de l'origine du langage? À qui, sinon à lui, doit-on précisément la distinction entre synchronie et diachronie, qui conduisit à se représenter l'histoire des langues comme une suite d'états intrinsèquement immobiles, comme si la langue ne changeait jamais et que le changement ne pouvait être un objet théorique? Certes on sait que Saussure dit de la langue qu'elle est sociale, alors que la parole est individuelle. Mais ce serait précisément parce qu'il la veut *identique* chez tous les individus, afin que, pour étudier une langue, il suffise d'étudier un individu! William Labov a appelé cet usage d'un argument « sociologique » pour écarter précisément le travail sociologique, le « paradoxe de Saussure » – que Chomsky n'aurait eu qu'à reprendre dans un nouveau cadre pour former sa conception de la « langue interne » que nous avons exposée plus haut (1972 : ch.8). De sorte qu'il paraîtrait bien surprenant d'apprendre que le structuralisme ait quelque chose à apporter à propos de la modélisation des phénomènes culturels et historiques...

Pourtant, la lecture même distraite du *Cours de Linguistique Générale* eût pu et dû donner au moins l'impression de quelque contradiction, tiraillement, incertitude, dans la pensée du maître de Genève. Saussure y écrivait en effet, dans un chapitre trop rarement lu, et intitulé (précisément) « Mutabilité et Immutabilité du signe » : « La langue est pour nous le langage moins la parole. Elle est l'ensemble des habitudes linguistiques qui permettent à un sujet de comprendre et de se faire comprendre. Mais cette définition laisse encore la langue en dehors de sa réalité sociale, puisqu'elle ne comprend qu'un des aspects de la réalité, l'aspect individuel; il faut une masse parlante pour qu'il y ait une langue. À aucun moment et contrairement à l'apparence, celle-ci n'existe en dehors du fait social, parce qu'elle est un phénomène sémiologique. Sa nature sociale est un de ses caractères internes. » (1972, p. 112). Difficile d'être plus clair : le « produit social » existe non pas uniquement dans le cerveau des individus, mais dans la collectivité, et même, disent les notes manuscrites de Saussure en personne, dans « l'âme collective ». Séparer la langue et la parole, ce n'est pas privilégier l'invariant, sur la variation : bien au contraire,

c'est parier que la langue elle-même est en variation. De fait, dans le même passage, Saussure continue : non seulement, la langue est sociale, mais en plus elle est historique (cf. Saussure, 1972, p. 112-113). Autrement dit, la langue, loin d'être définie comme ce qui est en exception par rapport à la variation historique et dialectale, semblait bien au contraire ne se constituer comme langue qu'à travers ces deux facteurs que Saussure appelait : le Temps et la Masse. Toute théorie de la langue qui ne permettrait pas de comprendre qu'elle est à la fois sociale et historique, pour des raisons d'essence, qui tiennent au type même de sa systématité – un « système complexe », dit Saussure, si complexe que nul peut en maîtriser le devenir (cf. 1972, p. 110-111), une forme de totalité sans cesse détotaillée et retotaillée – ne saurait être considérée comme adéquate pour Saussure.

Mais alors comment concilier cependant ces textes avec ceux qui semblent vouloir arracher la langue à la variation historique et dialectale? On sait que le *Cours de linguistique générale* n'a pas été écrit par Saussure, mais par deux de ses disciplines Bally et Séchehayé. La redécouverte progressive, depuis les années cinquante (avec en particulier Godel, 1957), des notes de cours et des manuscrits de Ferdinand de Saussure, permet d'avoir aujourd'hui une vision bien plus juste du maître de la linguistique moderne. L'exégèse saussurienne a établi que le problème de Saussure est non pas de tourner le dos à la linguistique historique du dix-neuvième siècle, mais de proposer au contraire une « épistémologie de la grammaire comparée »²⁷. Le véritable problème de Saussure n'est pas de retrouver des identités de langue sous la variété des manières de parler, mais bien de comprendre pourquoi, *à force de parler une langue, on finit par en parler une autre*. Car tel est bien le fait mis en évidence par la grammaire comparée, qui est la discipline qu'il pratiquait, lorsqu'elle a montré que des langues apparemment aussi différentes que le latin, le gothique et le sanskrit, ont toutes été originellement la « même » langue. Saussure le disait avec force dans ses conférences inaugurales à l'Université de Genève : « Le français ne vient pas du latin, il est le latin » (Saussure, 2002, p. 152)... Ce ne sont pas deux langues différentes, mais bien la *même langue*. C'est cette découverte que la répétition des signes linguistiques entraîne leur transformation qui le conduit, pour ainsi dire rétroactivement, à montrer que l'identité synchronique elle-même n'est pas

garantie et que c'était pour cette raison, d'essence, de structure, que les langues étaient par nature « livrées à la variation ». Tel était le projet intellectuel de Ferdinand de Saussure : expliquer ce qui, dans la structure de la langue, la livrait ainsi à l'histoire et à la variation réglée dont le comparatiste collecte les traces comme les bris d'un objet qui ne fut jamais homogène. Or les éditeurs du *Cours* ont inversé dans leur livre l'ordre d'exposition qu'on trouvait dans les leçons de Saussure : celui-ci commençait par des cours de linguistique historique et géographique, et terminait par la linguistique synchronique. À travers cette dernière, il s'agissait pour lui de répondre au problème de ce qu'il appelait « l'identité diachronique » : que doit donc être l'identité entre *calidus* et *calidus*, demandait-il, pour que *calidus* devienne *chaud* à mesure qu'on le répète²⁸ ? Car si la langue est dans le temps, encore une fois, ce n'est pas de manière extrinsèque, comme si elle était plongée dans un élément qui venait la ronger de l'extérieur, mais pour des raisons d'essence, qui tiennent à la nature même de ses identités.

Loin de reprocher à la linguistique historique de son temps d'avoir privilégié une approche historique du langage, Saussure lui impute d'avoir cru que l'historicité de la langue ne pouvait être due qu'à ses aspects « matériels », « physiologiques », aux seuls écarts de prononciation, et qu'il fallait, pour en rendre compte, renoncer à l'idée selon laquelle le langage était une activité symbolique. La variation serait la part inconsciente du langage, celle qui fait que les langues ne se réduisent pas à la fonction de communication qui est la leur, celle qui leur confère une certaine résistance, opacité – résistance et opacité qui montrent que loin d'être de simples fonctions subjectives, ce sont aussi des réalités objectives, susceptibles d'une approche théorique comme les phénomènes de la nature²⁹. Le changement ne pourrait donc être dû qu'à des altérations du geste articulatoire, à des écarts dans l'imitation, par les individus, de la parole des autres. Tel était notamment le programme du mouvement théorique connu sous le nom de « l'école des néogrammairiens » qui, à la fin du dix-neuvième siècle, dominait la linguistique internationale, avec un programme de recherches qui prétendait appuyer la reconstruction de l'histoire des langues sur une théorie physiologique contrôlée expérimentalement, permettant d'expliquer les changements linguistiques par des « lois phonétiques », reposant elles-mêmes

sur des contraintes physiologiques ou psychologiques s'exerçant également sur tous les individus.

Or Saussure est le premier sans doute à avoir compris que cette voie était une impasse, parce que jamais les phénomènes du langage ne pourraient être réduits à des événements spatio-temporels particuliers comme des gestes articulatoires individuels. Pour une raison simple et vertigineuse, totalement ignorée de l'ensemble de la tradition philosophique et théorique : ils sont en réalité *incorporels*. Le signifiant lui-même « dans son essence n'est aucunement phonique, il est *incorporel*, constitué, non par sa substance matérielle, mais uniquement par les différences qui séparent son image acoustique de toutes les autres » (Saussure, 19752, 164, je souligne). Le problème de l'identité diachronique ne fait ainsi que prolonger celui de l'identité synchronique : « Chaque fois que j'emploie le mot Messieurs, j'en renouvelle la matière ; c'est un nouvel acte phonique et un nouvel acte psychologique. Le lien entre les deux emplois du même mot ne repose ni sur l'identité matérielle, ni sur l'exacte similitude des sens, mais sur des éléments qu'il faudra rechercher et qui feront toucher de très près à la nature véritable des unités linguistiques » [Saussure, 1972 : 152]. Il faut bien comprendre que *tout* change, qu'il n'y a pas simple « variation », mais qu'il devient même impossible de conserver un « noyau » minimal, donc impossible de séparer l'accidentel de l'essentiel. L'entité de langue, si elle est une réalité, est une réalité *non observable*. Non observable, au sens précis de non mesurable, et donc de non expérimentable. Il est impossible de définir avec des critères rigoureusement physiques ni l'identité d'un signe linguistique, ni son unité, c'est-à-dire son individuation dans la chaîne phonique. L'histoire postérieure à Saussure se chargera de montrer que tous les appareils de mesure mis en place pour dégager des délimitations empiriques du signe étaient vains : Jakobson (1976) le rappelle dans la première de ces *Six leçons sur le sens et sur le son* qu'il avait données à la *New School for Social Research* à New York – devant cet auditeur très attentif que fut Claude Lévi-Strauss lui-même. La question du rapport entre substrat phonétique et valeurs phonologiques reste aujourd'hui encore ouverte, et ce fut au demeurant un des terrains sur lesquels les approches morphodynamiques, connexionnistes ou catastrophistes, ont prétendu apporter des solutions³⁰.

Or, pour répondre au problème du caractère inhérent, intrinsèque, de la variabilité aux phénomènes linguistiques, Saussure réintroduit ce qui semblait être un vieux concept : celui de signe. Si la linguistique est inscrite par Saussure dans cette discipline dont il forge le concept, la sémiologie, ce n'est pas, comme on l'a longtemps cru, parce que le langage partagerait avec les coutumes vestimentaires, les pratiques culinaires, les récits mythologiques ou les prescriptions rituelles, la même *fonction*, celle de signifier quelque chose, mais bien parce que ses objets sont de même *nature* : c'est cette forme bien singulière de l'identité comme *identité variable* qui caractérise le domaine de la sémiologie³¹. C'est d'ailleurs – curieusement, on l'a rarement noté – la définition même de la sémiologie : « science qui étudie la vie des signes au sein de la vie sociale » (1972, p. 33) – autrement dit science des manières dont les signes se cessent de se transformer au fur et à mesure qu'ils sont utilisés ou qu'ils « circulent ». À ceux qui cherchaient à reconstruire une linguistique générale à partir des résultats de la grammaire comparée, Saussure répondait donc, non sans ironie : « les philosophes du 17^e et du 18^e siècle qui ont parlé du langage nous ont déjà donné la réponse : le langage est fait de signes ». Mais cela supposait la reconstruction du concept de signe lui-même, et en particulier la rupture avec l'idée que le signe est un *moyen matériel* pour exprimer ou communiquer une idée spirituelle. Le signifiant est en fait aussi spirituel que le signifié. En affirmant que *c'est pour la même raison que le langage a un sens et une histoire*, Saussure prétend réconcilier les deux grandes traditions qui, aujourd'hui encore, se partagent le champ des études linguistiques : celle qui, issue de la tradition des grammaires rationnelles, étudie les langues comme des expressions plus ou moins parfaites d'un esprit subjectif, c'est-à-dire de la raison ; et l'autre, issue de la grammaire historique du xv^e siècle, qui conteste que les langues puissent être considérées à l'aune de la raison ou de la faculté de penser en général, étant séparées de la pensée par la pesanteur de leur dimension matérielle qui les livre à l'histoire. Saussure renvoie dos à dos ces deux traditions (cf. 1972, p. 117-118). En intégrant la linguistique dans la sémiologie, il réalise donc une double opération : d'un côté réaffirmer que le langage a un lien intime avec la pensée (tout dans la langue est psychique) ; de l'autre affirmer que ce lien intime avec la pensée n'a pas été clairement aperçu tant que l'on n'a pas tenu compte de « ce facteur imprévu, absolument ignoré

de la combinaison philosophique, LE TEMPS » (2002, p. 231) comme d'un facteur intérieur à la fonction psychologique elle-même. Le langage est bien fait de signe, mais le signe est tout entier une *pensée* (et non pas un moyen matériel pour exprimer la pensée), et, ce qui est plus étonnant, une pensée qui n'existe qu'à travers le temps et pour autant qu'elle ne cesse de changer, comme si, ce qui nous rendait capable de penser, nous amenait à penser toujours autre chose sans même que nous nous en apercevions...

Mécanique de la pensée, mécanique de l'histoire

Dès lors, la question est : qu'est-ce qui, dans la nouvelle théorie des signes proposée par Saussure sous le nom de sémiologie, permet de comprendre que le signe soit non seulement une pensée, mais encore une pensée telle qu'elle ne puisse jamais que devenir autre ? A priori, on le voit mal. Définir un signe non par des propriétés positives particulières, mais par le fait qu'il se distingue de tout autre, comme c'est notoirement l'un des axiomes majeurs du structuralisme, semble précisément devoir laisser hors de la langue tout le champ des variations dialectales et stylistiques. Peu importe en effet que certains Français roulent les « r » alors que les autres les grasseyent : il suffit qu'ils préservent autant de phonèmes que ceux avec qui ils souhaitent communiquer. Et on ne voit pas comment des variations au niveau de la parole pourraient entraîner la moindre conséquence sur la langue, puisque de telles variations seraient tout simplement ignorées, voire imperceptibles par les sujets parlants.

C'est qu'en réalité la théorie saussurienne du signe est plus complexe qu'il n'y paraît. Tout tient, au fond, à une thèse qu'on peut appeler de la *double détermination du signe*. Tout signe, en effet, est, selon Saussure, déterminé deux fois, à deux niveaux et selon deux logiques : une première fois comme « *terme* » et une deuxième fois comme « *valeur* »³². De plus, ces deux modalités objectives qui caractérisent chaque signe doivent être considérées comme des *formes émergentes* liées à des mécanismes de différenciation qui sont nettement distingués par Saussure : d'un côté ce qu'il appelle un jeu de *différences*, de l'autre un jeu d'*oppositions*³³. Ce sont les conflits de recouvrement entre ces deux modes de systématité, qui font le caractère essentiellement et intrinsèquement dynamique

de toute langue et plus généralement de tout système sémiologique. Un signe est d'abord un *ensemble de variations qualitatives* concrètes, par exemple phoniques, articulatoires, lumineuses, qui ne doivent pas encore être confondues avec ce qu'on appelle des *traits distinctifs*, car précisément elles se contentent de caractériser une expérience sonore comme une *nuance* complexe. Ces termes constitués de pures nuances sensibles sont ensuite comparés « par l'esprit » les uns aux autres, selon un processus que Saussure qualifie invariablement d'automatique, de mécanique ou d'instinctif, et dès lors *opposés* les uns aux autres. De cette comparaison oppositive naissent les « valeurs », qui peuvent être considérées comme de pures positions dans un système relationnel fermé. Alors, et alors seulement, le système peut être étudié abstraction faite de la nature même des différences, c'est-à-dire comme « forme » et non plus comme « substances ».

Précisons ces deux étapes. Soit le signe « clair » produit devant moi. C'est d'abord une certaine nuance qualitative singulière. C'est même, plus précisément, l'association entre plusieurs ensembles de nuances qualitatives, acoustiques (une tendance du son à se diffuser dans l'espace acoustique), mais aussi articulatoires (la sensation intérieure que j'ai des gestes de mon appareil articulatoire quand je parle), et aussi optiques (une impression singulière d'éclaircissement qui se mêle synesthésiquement aux précédentes), etc. Ce qui justifie qu'on appelle signes ces synesthésies, c'est que la nuance acoustique singulière qui caractérise, à un moment donné, le son « clair » prononcé devant moi, n'existerait pas si elle n'était relevée, accompagnée, d'une autre variation qualitative *hétérogène*, c'est-à-dire d'une sensation d'éclaircissement, qui l'accompagne régulièrement (par exemple, quand mes parents ouvraient les volets et disaient « il fait clair aujourd'hui »). L'inverse est vrai : mon expérience des variations de lumière est déterminée par les variations qualitatives hétérogènes qui se mêlent à elle de manière synesthésique, comme les sons du langage. De cette détermination conjointe d'une variation qualitative par une autre *émergent* des formes qualitatives, qui sont des synesthésies. Ce sont ces paquets de variations qualitatives mêlées, qui sont « comparés » les uns aux autres et acquièrent ainsi une valeur oppositive : les « termes » obtenus sont en effet classés en fonction des différences qualitatives qui les constituent, qui alors, et alors seulement,

fonctionnent comme des « traits distinctifs ». Ainsi car/autobus, train/méto, route/rue sont mis en série sous la variation URBAIN/EXTRA-URBAIN, qui n'a pas de sens en dehors de cette langue (le français), puisqu'elle est déterminée précisément comme la raison de cette double série qui permet d'ordonner en les comparant, deux séries de termes. Saussure construit ainsi une grammaire sérielle et qualitative, dans laquelle les catégories grammaticales ne sont pas des fonctions logiques a priori (comme elles sont dans les grammaires rationalistes de Port-Royal à Chomsky), mais des raisons de séries de variations concrètes (comme elles le redeviennent dans les modèles connexionnistes³⁴). Chaque terme est alors redéfini : il n'est plus seulement la somme des traits distinctifs qu'il actualise, mais aussi une position dans un système. C'est-à-dire qu'il se définit par son *opposition* aux autres : ainsi « enseignement » n'est plus seulement l'ensemble des traits distinctifs phoniques et sémiques qu'il actualise, mais seulement ce qui n'est pas « éducation », n'est pas « enseigner », etc., ou plus exactement il se définit par sa distance relative à ces autres termes³⁵. C'est cela la « valeur ». Elle résulte d'un mécanisme que Saussure appelle de « post-élaboration » (2002, p. 87-88), par lequel les *mêmes* entités sont pour ainsi dire déterminées deux fois, définies selon deux logiques, de surcroît en parallèle sur plusieurs plans qualitatifs.

Comment comprendre maintenant que cette théorie permette de rendre compte du fait que la simple répétition de ces entités complexes, plus précisément leur *usage*, puisse entraîner leur transformation, *alors même que chaque usage linguistique n'est déterminable que comme réalisation d'une possibilité linguistique qui, en un certain sens, la précède* ? C'est que l'organisation en série des « termes » (la post-élaboration) se fait sur la base des différences qualitatives, concrètes, qui les constituent. Or on peut fort bien altérer ces différences, pour une raison ou pour un autre (stylistique, ou physiologique, ou sociologique, etc.), sans modifier la structure des oppositions. Et de ce fait, tout l'équilibre du système est *susceptible d'être troublé. Mais il ne le sera que dans la mesure où intervient une interprétation*, c'est-à-dire dans la mesure où l'on créera, sur la base des nouveaux traits différentiels, une série d'oppositions qui n'existait pas auparavant, et qui modifie l'analyse que l'on peut faire des performances langagières. Saussure donne de nombreux exemples de ce phéno-

mène. Ainsi pour la création des pluriels par alternance dans les langues germaniques comme pour *tooth/teeth, foot/feet*, etc. On a d'abord un mécanisme pour exprimer le pluriel proche du latin : *fot/foti, gast/gasti*, etc. Puis l'*umlaut* qui altère la voyelle centrale au pluriel pour des raisons phonétiques : *fot/feti, gast/gesti*, etc. Puis chute du *i* final. Conséquence : création d'une forme totalement inouïe pour dire le pluriel dans la langue d'origine, non par désinence, mais par flexion du radical. Il s'agit là typiquement d'un *fait de valeur* résultant d'une *interprétation*, qui consiste à rapprocher deux formes qualitatives. Cette interprétation est seconde. Elle repose si l'on veut sur une « erreur », mais cela montre quelle sorte de logique, aveugle, qualitative, sérielle, anime l'usage. La théorie de la valeur se veut, si l'on peut dire, une *grammaire des usages*. Cette logique n'est pas indépendante de la *qualité* des termes, parce qu'elle est une logique a posteriori, dans lequel l'ordre surgit, non à partir de la projection de schémas structuraux préalablement donnés sur une réalité substantielle relativement indifférente, mais par le tracé de séries de ressemblances et de dissemblances entre les termes. Tous les phénomènes grammaticaux relèvent de cette logique, les fonctions grammaticales n'étant elles-mêmes que des super-séries entre les raisons de séries d'oppositions. Il importe de comprendre que la notion de système construite par Saussure est donc celle d'un système a posteriori, dans lequel la dimension formelle du langage émerge de la mise en ordre du donné qualitatif par lui-même. Même les fonctions grammaticales les plus élevées (comme celle de « sujet » par exemple), dépendent de la réalité qualitative des langues parlées. La grammaire, loin d'être un présupposé de toute activité de langue, est un produit émergent de la mise en ordre du sensible par lui-même.

Or ces processus de mise en ordre sont pour Saussure rigoureusement *mécaniques* : ils relèvent d'une faculté innée. La grande originalité du fondateur de la « sémiologie » – et ce par quoi il doit être inscrit dans l'histoire des sciences cognitives – c'est d'avoir fait l'hypothèse d'une faculté *non spécifique*, la faculté sémiologique, qui se caractérise comme un certain type de traitement des données sensibles, et dont le mode même de fonctionnement implique l'émergence de ces produits historiques et sociaux intrinsèquement variables que sont précisément les langues (mais aussi les cultures vestimentaires, les

religions, etc.). Saussure distingue en effet toujours le *langage*, qui est une faculté universelle et qui finalement se confond avec la faculté sémiologique (« *facultas signatrix* »), et les *langues* sans lesquelles cette faculté, certes, ne peut s'exercer, et qui sont d'ailleurs des produits ou des effets *involontaires* et *latéraux* de son exercice, mais qui ne sauraient être considérées comme des ensembles de règles ou d'opérations, car ce sont plutôt des systèmes de formes émergentes particulièrement complexes. Tel est sans doute une des points les plus importants de la théorie cognitive de Saussure : distinguer le niveau des *opérations*, qui est universel et mécanique, et celui des *formes*, qui sont des produits émergents et variables. Le langage est même pour Saussure une faculté innée, dotée d'une base physiologique (la zone de Broca) : Saussure y voit un ensemble de « principes » qui « régissent » les langues (2002, p. 145-146), « d'opérations possibles de l'instinct humain appliqué à la langue » (idem), de « forces psychologiques » (p. 158), une « fonction », une « puissance », une « organisation », un « appareil » (p. 178) – bref un ensemble de mécanismes déjà montés qui sont activés par des stimuli déterminés. Saussure imagine donc quelque chose comme un *fonctionnement mécanique* au titre de faculté sémiologique. C'est dans le cadre d'une conception mécanique de la pensée que Saussure souhaite repenser l'histoire – non pas parce qu'il y aurait des *lois de l'histoire*, mais précisément en distinguant le niveau des lois (ou des processus) et celui des formes (ou des événements). Les mécanismes *sont* aveugles, et d'autant plus aveugles qu'ils n'agissent pas directement sur les formes même qu'ils transforment. Saussure n'oppose pas une approche naturaliste de l'esprit compris comme un ensemble de mécanismes aveugles, qui se déclenchent dans des conditions précises, et agissent toujours localement, et une approche culturaliste qui parlerait plus volontiers d'« esprit objectif » et qui verrait dans les cultures humaines des sortes d'organismes spirituels vivants, à caractère collectif, ayant leurs propres tendances évolutives et imposant en quelque sorte leur finalité. Il cherche précisément une position tierce, intermédiaire entre le cognitivisme individualisant et la conception de l'esprit objectif holiste.

La postérité structuraliste verra dans cette obstination à tenir le caractère à la fois *mécanique* et *aveugle* des processus à l'œuvre dans le changement la trace d'un archaïsme de Saussure, d'une difficulté

à rompre avec son temps³⁶. Il lui oppose une conception de la notion de système beaucoup plus téléologique. C'est d'ailleurs par la « position téléologique traditionnelle » et par le souci de faire de la catégorie de finalité une catégorie scientifique, que Jakobson introduisait la notion même de « structuralisme », pour parler d'une tendance propre à la science russe et que redécouvrait l'Occident³⁷. Dans ses « Principes de Phonologie diachronique », Jakobson défendait l'approche structurale au nom d'une pensée de l'histoire : le système témoigne d'une vie propre, d'une tendance intérieure, et c'est pour reconstituer ces tendances internes aux langues qu'on doit passer par la notion de système : « *Quand nous considérons une mutation linguistique dans le contexte de la synchronie linguistique, nous l'introduisons dans la sphère des problèmes téléologiques. (...) nous n'avons aucune peine à découvrir la fonction de cette mutation : sa tâche est de rétablir l'équilibre.* » (Jakobson, 1931).

Or il me semble que c'est au contraire par son refus de toute téléologie que Saussure peut faire figure de précurseur. Car il développe une conception extrêmement originale de ce qu'on doit bien appeler l'*esprit objectif*. Pour Saussure, l'*esprit objectif* n'est pas une pensée vivante, une subjectivité collective, palpitant à travers les institutions, les langues, les habitudes sociales, comme le dit Jakobson, au fond dans la continuité de Hegel et de Humboldt, mais un ensemble de pensées, de choses pensées, de « noèmes » sans « noèse » (pour parler comme Husserl), d'idées qui se produisent elles-mêmes, et « forcent ainsi l'esprit à entrer dans la voie spéciale qui lui est laissée » par les aléas des transformations phonétiques : « Le « génie de la langue » pèse *zéro* en face d'un seul fait comme la suppression d'un *o* final » (Saussure, 2002, p. 216). Ainsi, un système linguistique est toujours un simple *résultat* d'une organisation sérielle d'éléments qualitatifs émergeant de la détermination corrélatrice de plusieurs plans de variations qualitatives. Si les signes ont une « vie » par eux-mêmes, ce n'est pas parce que les systèmes de signes sont des organismes³⁸, mais parce que ces *effets de pensée* que sont les signes sont déterminés de telle sorte que les *usages* que les individus en font au cours de leurs interactions peuvent finir par avoir des effets sur la nature des possibilités même qui sont actualisées à travers ces usages. C'est en ce sens que la sémiologie est la science qui « étudie la vie des signes au sein de la vie sociale » : une sémiologie générale

sera une théorie des conditions universelles qui font que le signe par nature est une pensée qui s'altère nécessairement dans une histoire contingente. Les « sémiologies » particulières étudieront les modalités de construction propres à certains systèmes de signes, en fonction des « substances » sur lesquelles elles travaillent, ainsi que les histoires singulières de pensées auxquelles elles donnent lieu.

C'est ce programme sémiologique que Saussure cherchera à réaliser dans ses recherches sur les légendes, qui sont restées inédites, et dont on commence seulement à avoir une vue un peu plus synthétique. On retrouve dans la légende cette étrange condition qui fait que, plus on cherche à faire ce que l'on a toujours fait, plus on fait autrement, et ce sont pour les mêmes raisons de *structure* que la légende comme la langue sont livrées à l'historicité. « À chaque instant, par défaut de mémoire des prédécesseurs ou autrement, le poète qui ramasse la légende ne recueille pour telle ou telle scène que les *accessoires* au sens le plus propre, théâtral ; quand les acteurs ont quitté la scène il reste tel ou tel *objet*, une fleur sur le plancher, qui reste dans la mémoire, et qui dit plus ou moins ce qui s'est passé. [...] Imagination sur lacune de mémoire est le principal facteur de changement avec volonté de rester autrement dans la tradition. » (Ms. fr. 3959/3,3). La comparaison avec un ensemble d'*accessoires* est forte. Elle ne peut pas ne pas évoquer le thème du *bricolage*, dans lequel Lévi-Strauss voyait le modèle même de la « pensée sauvage », reconstruction incessante de structures avec des événements, ou, comme disait Lévi-Strauss, des « débris de structure » (cf. Lévi-Strauss, 1960). Le système sémiologique est construit à partir de fragments accidentels d'un drame dont on ne garde que les traces ; il est bien, pour Saussure comme pour Lévi-Strauss, un agencement des débris d'un système antérieur. C'est pourquoi nous disions que le concept de système qu'utilise Lévi-Strauss est finalement beaucoup plus proche de Saussure que de Jakobson : systématisation *a posteriori* d'un donné qualitatif stratifié qui fait que la valeur d'un signe est toujours soumise à variation, non au sens où le système aurait une dynamique interne, mais au sens où il est toujours à refaire. On ne peut prévoir précisément de quelle manière et dans quel sens les morceaux seront finalement arrangés, mais une étude des transformations effectives des versions d'une légende sera instructive pour comprendre comment fonctionne

cette activité de décomposition et de recombinaison incessante qui caractérise « l'esprit » à l'œuvre dans les produits sémiologiques.

C'est en ce sens qu'il faut entendre la notion de « vie ». Le système a une vertu propre, une sorte de vitalité intrinsèque, qui est celle non pas d'un organisme, mais plutôt d'une colonie, d'une « fourmière ». Si le système a une vie, c'est non pas une force d'engendrement, mais uniquement de régénérescence à partir de ses débris. Il n'a aucune force pour résister au changement, mais il en a une pour se refaire. Saussure compare la langue à une « fourmière dans laquelle on plante un bâton et qui à l'instant sera réparée dans ses brèches » (ELG 267), à une « robe couverte de rapiécages faits avec sa propre étoffe » (CLG, 235), enfin, et surtout, à une « machine qui marcherait toujours quelles que soient les détériorations qu'on lui ferait subir » (E.I.192.1144.3C,346). Telle est donc la formidable intuition de Saussure par rapport à la question du langage : le langage est une machine, oui, mais une machine qui fonctionne avec ses propres débris. Ce sont là des formules que Claude Lévi-Strauss retrouvera dans la *Pensée Sauvage*. Rien à voir, on le comprend, avec un système de formes innées s'appliquant à des contenus représentatifs passifs, comme le voudra Chomsky, ni non plus avec des organismes dotés de tendances propres, comme le diront les fonctionnalistes (Jakobson notamment).

La culture aux limites de la simulation

Quelles conclusions tirer de ce regard rétrospectif, et de cette nouvelle lecture du structuralisme, pour une archéologie du projet de l'intelligence artificielle ?

On s'étonnera d'abord de voir le structuralisme réintégré dans l'histoire de l'Intelligence Artificielle non pas par son voisinage avec la cybernétique et la théorie générale des systèmes, c'est-à-dire sous l'hypothèse selon laquelle il faudrait étudier les cultures (ou les formes culturelles, comme les langues, les systèmes mythologiques, les institutions politiques, les types d'organisation sociale, etc.) comme des systèmes ayant leur rationalité propre, indépendante de leurs incidences sur les vies individuelles (bref dans une forme de holisme), mais bien plutôt comme

une position précisément anti-holiste, qui met en scène une totalité toujours divisée, toujours fragmentée, rigoureusement *intotalisable*, dans laquelle les usages locaux ne cessent de modifier la structure globale de ce qui, pourtant, les conditionne (et sans quoi ils ne pourraient même être *perçus*). Le structuralisme avait pourtant d'excellents motifs pour se croire de profondes affinités avec la cybernétique. La première définition du terme même de « structuralisme », on l'a vu, qu'on doit à Jakobson, insistait sur le fait que ce mouvement scientifique transdisciplinaire se caractérisait par la réintroduction de la finalité dans la science même. Or on sait que la cybernétique de Norbert Wiener, avec le concept de *feed-back*, a trouvé un moyen pour proposer une théorie mécanique de la finalité elle-même, et que c'est à ce titre qu'elle se présente comme une science de l'intelligence et une technique pour produire des êtres intelligents. Que serait, en effet, un comportement intelligent, sinon un comportement calculateur, capable d'anticiper ses effets, de se corriger et donc de témoigner d'une certaine rationalité, de donner le sentiment qu'il *sait* ce qu'il fait ?

Or voici ce que nous apprend cette relecture du structuralisme : ce n'est peut-être pas tant dans la *maîtrise* ou le *contrôle* que, tout au contraire, dans la capacité à se livrer à des entités au devenir incontrôlable, que se trouverait le secret de l'intelligence humaine (ce qui suppose qu'elle soit une intelligence symbolique, c'est-à-dire opérant toujours en parallèle sur plusieurs plans à la fois pour la détermination des entités auxquelles elle est sensible – mais au fond ce n'est pas ici l'essentiel). Une approche *positive* de l'intelligence n'est pas à chercher dans la modélisation d'entités capables d'anticiper les effets de leurs propres comportements, mais plutôt dans celle d'êtres capables de devenir sensibles et de faire vivre des entités incorporelles (incorporelles au sens où elles ne sont extraites des variations physiques de l'environnement que par les validations réciproques que s'en donnent les agents d'une collectivité) douées d'une vie propre, qui échappent à chacun d'entre eux, comme elles échappent à toute rationalité globale, pour entrer dans une histoire rigoureusement imprévisible. On a tendance à penser que le défi de l'intelligence artificielle est la simulation de la *conscience*, au sens large (construire des organismes qui savent ce qu'ils font), alors qu'il s'agirait plutôt de simuler la *culture*, dans sa double dimension de

consensus et de variation mêlée. Cela ne veut pas dire qu'on doive renoncer à une direction centrale de l'approche cybernétique, à savoir la thèse de la profonde continuité entre l'intelligence et la vie. Mais cela veut dire qu'il faut chercher cette continuité non pas du côté de la création de *super-organismes* artificiels³⁹, qui reproduirait, mais à un niveau collectif, le type de rationalité qui est celui des individus (ou des systèmes autotéliques), mais plutôt du côté de la création de *milieux* artificiels, et d'organismes capables de constituer ces milieux tout en y trouvant les conditions même de leur existence, et de participer à leur imprévisible évolution.

On voit dès lors qu'il ne s'agit pas seulement de reconduire la notion d'intelligence à ses conditions sociales et culturelles (une lecture attentive du test de Turing aurait déjà dû nous en convaincre, car Turing était formel : être intelligent, ce n'est rien d'autre qu'être capable de saisir un certain nombre de conventions, nous appelons « intelligent » qui se comporte *comme nous*). Il ne suffit pas de dire que l'esprit n'est pas qu'une faculté subjective, faite d'opérations et de mécanismes s'exerçant au niveau des individus, et qu'être capable de pensée, c'est être capable d'entrer dans des systèmes sociaux et culturels variés, comme le sont les langues, les religions, les systèmes politiques, etc., autant de formes d'« esprit objectif », de pensées déposées en dehors des opérations cognitives. C'est la manière même de penser cet « esprit objectif » qu'il faut rectifier, et c'est cette rectification que nous aide à effectuer la distinction que nous avons proposée entre les deux héritages du « structuralisme ». Au lieu que l'esprit objectif soit un ensemble de systèmes autonomes, doués d'une rationalité propre, dont on pourrait reconstruire les lois d'évolution *au niveau systémique*, il s'agit plutôt de cohérences émergentes, effets de processus mécaniques agissant, eux, à un niveau inférieur, et aveugles à leurs effets. Un peu comme dans le paradigme subsymbolique développé par Paul Smolensky⁴⁰, les processus agissent à un niveau local, mais les effets sont globaux.

L'épistémologie des sciences humaines oppose en général les approches herméneutiques des phénomènes culturels, qui voient dans les sciences de l'homme autant d'efforts pour apprendre à caractériser au plus près des phénomènes singuliers, forcément singuliers, événementiels, historiques, non

répétables, comme le seraient *par nature* les phénomènes humains en général (comme si la tâche la plus propre des sciences humaines était de nous dire ce qu'on est et où on en est), à des approches théorétiques, « nomologiques », qui cherchent au contraire des causalités, des processus, des mécanismes, des lois (sur le modèle de la physique classique). Et on range bien sûr l'intelligence artificielle du côté des secondes. On pourrait même dire les choses un peu autrement : il y a deux orientations apparemment divergentes dans le projet des sciences humaines (si l'on écarte celles qui cherchent une genèse individuelle des faits sociaux) : l'une qui cherche à mettre au jour le caractère systématique, superorganisé, des phénomènes sociaux, qui constitueraient comme un niveau d'organisation supplémentaire ; l'autre au contraire qui insiste sur le caractère aléatoire, imprévisible, indirigeable, intotalisable de tout ce qui relève de la culture ou de l'histoire... L'une qui cherche *plus de rationalité* dans les systèmes sociaux, l'autre qui au contraire en cherche moins et se dit : « contrairement aux sciences de la nature, où les objets sont effectivement prévisibles, contrôlables, manipulables, les sciences humaines ont des objets qui sont intrinsèquement labiles, fuyants, incontrôlables... »⁴¹.

Or, avec ce « structuralisme » revisité que nous essayons de mettre en évidence, nous montrerions qu'il existe dans les sciences humaines une voie tierce entre la perspective herméneutique et la perspective nomologique, une voie qui accepte à la fois qu'il y ait des *mécanismes* sous-tendant les phénomènes culturels, mais que cette recherche des mécanismes n'implique pas que ce qui se passe, ce qui apparaît, au niveau culturel, puisse lui-même être décrit selon des lois mécaniques. Il faut distinguer le niveau des lois et des processus, et le niveau des événements et des entités. Il y a des processus sous l'histoire, mais l'histoire elle-même n'est pas réglée par des processus. *Il y a des lois qui nous rendent capables d'histoire, mais l'histoire elle-même n'a pas de loi*. Nous sommes des machines à entrer dans des cultures, à les constituer et à les faire varier, mais les cultures elles-mêmes ne sont pas des machines. Comprendre qu'il existe un grand courant de la pensée des sciences humaines qui soutient cette thèse-là, c'est ouvrir, je crois, un nouveau carrefour pour la rencontre plus que jamais nécessaire entre l'intelligence artificielle et les sciences sociales.

On peut en effet, à travers cette nouvelle lecture du structuralisme, jeter une autre lumière sur ce qui est en jeu, d'un point de vue épistémologique, à travers la constitution d'un ensemble de disciplines s'occupant des phénomènes humains, et mieux définir le point de rencontre entre sciences humaines et intelligence artificielle. Oui, les sciences humaines nous confrontent à une autre forme de rationalité scientifique : non, elles ne nous condamnent pas à renoncer à chercher les lois de l'esprit, mais elles nous dirigent vers une rationalité scientifique qui, *au lieu de chercher à modéliser des phénomènes déterminés, modélise une capacité à faire émerger des phénomènes imprévus*. Pas même ce qu'on a appelé une *compétence*, mais bien une capacité à faire varier cette compétence de manière imprévisible, incontrôlable, aléatoire. De même, la gravité du défi lancé au paradigme de la simulation, et même à l'épistémologie des modèles, qui longtemps a soutenu les problématiques de l'intelligence artificielle, devient évidente. Qu'est-ce qu'on simule ? Tel ou tel comportement ? Ou bien au contraire la capacité à simuler/assimiler/dissimiler des comportements en général ? Tel ou tel comportement culturel, ou la capacité à entrer dans une vie culturelle et à la faire bouger de l'intérieur, à même sa propre répétition, sans que cette variation ait jamais la forme d'un *contrôle* ? Mais si on ne doit pas simuler tel ou tel comportement, cela veut dire qu'*on ne sait plus ce que c'est qu'être intelligent*. Aux machines intelligentes, nous ne demanderons plus simplement de faire à notre place et mieux que nous ce que nous savons déjà faire, et que nous considérons déjà comme intelligent. Nous leur demanderons d'inventer entre elles et même, plus profondément encore, *avec nous*, des formes culturelles sans cesse nouvelles, c'est-à-dire de participer aux événements collectifs par lesquels on ne cesse de réinventer les formes de l'intelligence, de réassigner une définition à ce qu'avoir de l'esprit veut dire, parce qu'on ne cesse de trouver de nouveaux consensus culturels variables. L'intelligence artificielle bien comprise implique que les machines soient non pas tant des modèles que des partenaires d'une vie culturelle collective⁴².

Ainsi, s'il est vrai que l'intelligence est dans la culture, qu'un être doué d'intelligence est un être dont nous partageons la culture, *c'est-à-dire* qui est sensible aux mêmes variations synesthésiques de l'environnement que nous, mais que ces cultures sont quelque chose d'imprévisible, de sorte que la seule chose

que nous puissions faire, c'est de modéliser non pas l'émergence de *telle* culture ou de telle autre culture, mais bien de la capacité à entrer dans une histoire faite d'accidents, de contingences, d'imprévisibles aventures où une affaire de détail peut changer l'ordre entier d'un système, bref de devenir collectifs, alors on comprendra que la simulation de l'intelligence ne saurait être la simulation de tel ou tel comportement, mais la création de machines capables d'entrer avec nous (ou entre elles) dans des rapports culturels où elles nous apporteront. Il y a peut-être une condition politique à la possibilité même de l'intelligence artificielle. Il se peut que le plus gros obstacle épistémologique à l'intelligence artificielle tienne à notre attitude à l'égard des machines. Nous devons apprendre à ne plus demander seulement aux machines de faire mieux que nous, ce que nous savons faire, mais de devenir des partenaires, des interlocuteurs, de participer à la vie culturelle, pour inventer avec nous des *choses à faire*, des usages culturels, dont nul ne peut détenir le sens. Si Saussure a raison, et que la culture est un ensemble de systèmes de formes émergeant de processus mécaniques, mais tel que les variations d'usage peuvent entraîner une transformation de la structure du système rigoureusement *incalculable*, alors c'est tout le paradigme de la simulation qui doit être réévalué. Apprendre à une machine à jouer aux échecs, c'est lui permettre d'entrer avec nous dans ces processus au long cours par lesquels le jeu d'échecs a une histoire, dont les règles changent, dont les pratiques évoluent, qui découvrent sans cesse de nouvelles stratégies et qui finit par se détourner de ce jeu, pour se consacrer au Sudoku... Mais après tout, ce renversement de paradigme dont nous disons qu'il est la condition d'un nouvel âge de l'intelligence artificielle, n'est-il pas celui-là même qu'on a depuis longtemps requis pour les sciences humaines ? Car n'est-il pas évident que les théories que nous construisons sur ce que nous sommes se caractérisent par le fait qu'elles transforment leur objet même, à savoir nous-mêmes, qui *nous* découvrons à travers elles ? De même, l'intelligence artificielle ne saurait être seulement une description de la culture ; elle ne peut qu'en être, simultanément, une récréation.

NOTES

1. Je dois beaucoup à ce texte et, d'une manière générale, aux travaux et à la conversation de Stefano Franchi pour la perspective générale dans laquelle s'inscrit cet article, celle d'une relecture des sciences humaines dans l'horizon d'une archéologie de l'intelligence artificielle.
2. Pour marquer cette différence entre une problématique spéculative et une discipline (relativement) déterminée, S. Franchi et G. Güzeldere proposent de réserver l'expression « intelligence artificielle », en minuscules, pour la première, et d'appeler Intelligence Artificielle, comme un nom propre, donc, la seconde.
3. cf. Karsenti, 2006a.
4. Ce projet est très clairement formulé dans le deuxième paragraphe de l'introduction des « *Formes élémentaires de la vie religieuse* » (Durkheim, 1912, p. 12-28), l'idée étant que le caractère universel et nécessaire des formes a priori de la pensée doit être mis sur le compte de l'obligation sociale qui s'impose aux individus. Pour une lecture serrée de la confrontation de Durkheim avec Kant, voir aussi Karsenti, 2006b.
5. Ainsi cette phrase fort explicite de la fin de la cinquième partie du Discours de la méthode : « Car au lieu que la raison est un instrument universel qui peut servir en toutes sortes de rencontres, ces organes [du corps] ont besoin de quelque particulière disposition pour chaque action particulière ; d'où vient qu'il est moralement impossible qu'il y en ait assez de divers en une machine, pour la faire agir en toutes les occurrences de la vie de même façon que notre raison nous fait agir » (Descartes, p. 165). On voit le raisonnement : il y a une infinité de possibilités d'actions, alors que les organes sont finis.
6. cf. Chomsky, 1969, p. 18-24. Je rappelle qu'on appelle paradigme symbolique en intelligence artificielle la stratégie de modélisation dont Herbert Simon et Allan Newell fournirent le modèle spectaculaire avec Logic Theorist en 1956, programme de démonstration de théorèmes qui réussit à démontrer, et parfois mieux que leurs auteurs, un certain nombre des propositions des Principia Mathematica de Russel et Whitehead. Cette stratégie suppose de définir l'esprit comme un manipulateur de symboles qui, d'un côté, ont une réalité physique et peuvent être pris dans des relations causales, et de l'autre, ont un rôle fonctionnel dans un calcul de type logique, de telle sorte que les propositions peuvent être manipulées de manière parfaitement aveugle à la signification en vertu de règles purement formelles, et n'être interprétées qu'à terme. Modéliser une fonction cognitive sera donc le traduire dans un modèle de type « logique » comme celui-ci, avec un bon vocabulaire de base et de bonnes règles de transformation, qui permettent d'engendrer la diversité des « réponses » modélisables aux problèmes qui s'y posent.
7. Pour une tentative de synthèse philosophique de ces diverses approches dynamicistes, voir van Gelder, T. (2003) et Port, R. & van Gelder (eds.), 1995.
8. Voir en particulier, dans le contexte français où ces approches sont particulièrement développées, les ouvrages de François Rastier (notamment 1987), et la théorie des formes sémantiques développées par Pierre Cadiot et Yves-Marie Visetti (Cadiot & Visetti, 2001 ; Visetti & Cadiot, 2006). Partageant, avec les « linguistiques cognitives » (Lakoff, Langacker et Talmy), la conviction qu'il est impossible de séparer le traitement purement syntaxique de la dimension sémantique des langues, ces approches s'en écartent par leur méfiance à l'égard des efforts de ces dernières pour ancrer la signification linguistique dans une structure perceptive trop vite naturalisée, comme si toutes les langues humaines traduisaient une forme unique d'expérience qu'on pourrait référer à des structures en particulier spatiales tout à fait précises. Aussi bien Rastier que Cadiot & Visetti insistent sur l'impossibilité d'assigner un sens à une entité linguistique, qu'il s'agisse d'un lexème (par exemple *clef*, dont on serait bien en peine de désigner un noyau sémantique unique à travers tous les usages comme « *clef des champs* », « *clefs du problème* », etc.), d'une préposition (voir l'analyse des prépositions du français dans Cadiot & Visetti, 2001, p. 23 sq.) ou d'une structure grammaticale comme le passif, voire d'une phrase comme un proverbe (voir Visetti & Cadiot, 2006). Il n'y a donc rien qui soit « le » sens, mais des constructions interprétatives. Si ces deux approches, qui remettent au travail l'ensemble des questions de la linguistique (notamment la distinction langue/parole, le statut des identités linguistiques, le rapport entre langage et perception, le statut épistémologique de la discipline, etc.), et même de l'ensemble des sciences de l'esprit, s'accordent sur les gestes fondamentaux qu'on vient d'énoncer, elles s'écartent l'une de l'autre de manière significative dans les réponses aux problèmes qu'elles posent (voir pour une bonne formulation des

divergences, Visetti & Cadiot 2006, p. 114 sq.). Pour une discussion de l'approche de Visetti et Cadiot, voir ma conférence en ligne : <http://www.diffusion.ens.fr/index.php?res=conf&idconf=1583> (Maniglier, 2006a). Pour une argumentation proprement philosophique en faveur d'une approche herméneutique dans les sciences cognitives, voir notamment Salanskis (2003).

9. Lévi-Strauss mentionnera à plusieurs reprises dans son œuvre les recherches cybernétiques et la théorie des jeux comme une inspiration pour la méthode structurale. Nous reviendrons plus loin sur ce qu'il faut penser de cette filiation. Pour une analyse plus complète des rapports de Lévi-Strauss avec l'intelligence artificielle, je renvoie à Franchi 1997.

10. Weinreich, Labov et Herzog, 1968, p. 100-1001. C'est aussi le point de départ de William Labov dans l'introduction de son livre, *Sociolinguistique* (1972), une des plus profondes tentatives pour réintroduire l'étude de la variation comme objet de la linguistique (cf. p. 37-38).

11. Labov, 264. En italiques les passages supposément « en » Black English Vernacular. Labov remarque lui-même que ces attributions sont souvent arbitraires et discutables.

12. J'emprunte ces exemples à Langacker 1987, p. 50 pour la première phrase, p. 37 pour la seconde.

13. Cette refonte de la grammaire générative a été proposée par Chomsky (1991, en particulier p. 17-40), et présentée comme un changement de paradigme radical après les aménagements apportés à la théorie standard dans les années soixante-dix sous la pression d'un certain nombre de critiques faite au modèle transformationnel qui avait fait la gloire de Chomsky et le succès rapide de la grammaire générative. La question de savoir s'il s'agit ou non d'une rupture est débattue.

14. Voir Chomsky (1991), et aussi l'introduction de Chomsky à Pollock (1997).

15. Je pense plus particulièrement aux passages suivants : « Ce que nous nommons « obéir à une règle », est-ce quelque chose qu'il serait possible à un seul homme de faire, et de ne faire qu'une seule fois dans sa vie ? » (§199) ; « Et c'est pourquoi « obéir à la règle » constitue une pratique. Et croire qu'on obéit à la règle n'est pas obéir à la règle. Voilà pourquoi il n'est guère possible d'obéir à la règle « en particulier » : autrement croire qu'on obéit à la règle serait la même chose que lui obéir. » (§ 202). Ainsi, « une personne ne se dirige suivant un poteau indicateur qu'autant qu'il existe un usage constant de poteaux indicateurs (une pure habitude) » (§198), et, plus généralement, « obéir à une règle, faire une communication, donner un ordre, faire une partie d'échecs, sont des habitudes (usages, institutions) » (§199). Ces textes ont suscité une abondante littérature et de nombreux débats, sur le caractère social ou non des règles, dans lesquels il est impossible d'entrer. Il s'agit ici non pas d'un point d'exégèse de Wittgenstein, mais de l'usage d'un argument, qui peut s'y lire. Pour une exposition à la fois synthétique et originale de ces questions, voir Laugier (2006).

16. Voir sur ce point les explications de Boltanski (1995), et le classique de Prince & Smolensky (2004) (qu'on trouve sur Internet : <http://roa.rutgers.edu/files/537-0802/537-0802-PRINCE-0-0.PDF>). Il s'agit d'une des tentatives les plus sérieuses pour proposer des modélisations précises des approches en termes de « principes » et « paramètres ». L'intérêt des modèles développés par Smolensky est, outre leur caractère indissolublement linguistique et computationnel, de faire de leur capacité à modéliser le changement linguistique un critère de leur adéquation scientifique.

17. Le lien entre les travaux de Elman sur le changement et les problèmes de la linguistique variationniste a été fait par Bernard Laks (1996), dans un livre auquel ces pages doivent bien plus que cette note en bas de page ne peut le laisser entendre. Bernard Laks est en effet un des rares auteurs à avoir vu ce lien entre le structuralisme et les nouveaux paradigmes cognitifs qui émergeaient de la contestation du paradigme symbolique.

18. Voir Rumelhart D.E., McClelland J. L., and the PDP Research Group, 1988, vol. II, p. 22-23.

19. cf. Drestske 1986, §3. La bactérie magnétotactile est une bactérie marine qui possède un aimant interne qui la dirige, si elle est dans l'hémisphère Nord, vers le Nord et donc vers la profondeur de l'Océan, lui évitant ainsi les eaux de surface riches en oxygène, ces animaux ne pouvant vivre qu'en l'absence d'oxygène.

20. Il est évident qu'une telle conception de la culture n'est pas ce qu'il y a de plus répandu sur le marché des théories de la culture. J'ai essayé cependant de la défendre et de construire une position en philosophie de l'esprit à partir de cette conception à la fois matérialiste et variationniste de la culture, dans une conférence prononcée à l'Ecole Normale Supérieure de Paris (Maniglier 2006b).

21. De ce point de vue, on peut penser à cette phrase énigmatique de Wittgenstein : « Vous dites alors que c'est l'accord des hommes qui décide si ce qui est vrai ou faux ? – C'est ce que les êtres humains disent qui est vrai et faux ; et ils s'accordent dans le langage qu'ils utilisent. » (§241). S. Laugier (2006) fait de cette phrase la clef de sa lecture très suggestive de la position de Wittgenstein, qui le rapproche du jugement esthétique au sens de Kant, c'est-à-dire de la revendication en première personne d'une vérité universelle qui ne peut jamais être attestée objectivement et qui repose toujours sur la postulation d'un accord intersubjectif.

22. Le cas a été étudié par L. Steels pour les synonymes dans l'expérience des Talking Heads, cf. Steels 2000.

23. Pour une synthèse des discussions sur la définition de notion de culture dans un contexte spontanément naturalisant et épistémologique, celui de l'éthologie, voir Lestel, 2001, p. 107 sq., dont j'extrais ce passage : « Pour être considérées comme « culturelles », les variations comportementales enregistrées par le primatologue ne doivent être explicables ni par des déterminismes génétiques ni par des déterminations environnementales. »

24. cf. Turing, 1995/1950.

25. Certes Lévi-Strauss cite Wiener, Shannon et von Neuman ; mais il ne les utilise jamais dans son travail comme des instruments de modélisation directement appliqués. Il en retient l'idée générale d'une science de l'esprit, voire d'une science de la subjectivité, comme dans la théorie des jeux, mais il ne s'en inspire pas d'un point de vue opératoire ou méthodologique. Une anecdote rapportée par Lévi-Strauss lui-même, concernant l'immeuble où il vivait durant son exil à New York pendant la guerre, traduit allégoriquement cette « non-rencontre » : « Il y a deux ou trois ans seulement, j'appris que Claude Shannon y logeait aussi, mais sur la rue et à un étage supérieur. À quelques mètres l'un de l'autre, lui créait la Cybernétique et j'écrivais les Structures élémentaires de la pensée. Nous avions, à vrai dire, une jeune amie commune dans la maison, et je me souviens que, sans citer de nom, elle me parla un jour d'un de nos voisins qui, m'expliqua-t-elle, s'occupait à « inventer un cerveau artificiel ». L'idée me parut bizarre, et je n'y prêtai pas attention. » (Lévi-Strauss, 1983, p. 347).

26. Voir parmi les nombreuses publications récentes : Bouquet, 1997 ; Fehr, 2000 ; Utacker, 2002 ; Maniglier, 2006.

27. « Il est mystérieux, le lien de cette identité diachronique qui fait que deux mots ont changé complètement (calidus : chaud ; <germ.> aiwa : <all.> je) et qu'on affirme cependant l'identité. En quoi consiste-t-il ? Précisément ! Aussitôt nous sommes dans le premier problème de la langue. Non que la réponse soit aisée. » (Saussure 1967, p. 413).

28. Cet argument est présent du fondateur de la grammaire comparée, Bopp, jusqu'aux contemporains de Saussure, les néo-grammairiens. Michel Foucault en a fait une des clefs de son histoire des sciences humaines : « La nouvelle grammaire est immédiatement diachronique. Comment en aurait-il été autrement, puisque sa positivité ne pouvait être instaurée que par une rupture entre le langage et la représentation ? L'organisation intérieure des langues, ce qu'elles autorisent et ce qu'elles excluent pour pouvoir fonctionner ne pouvait plus être ressaisi que dans la forme des mots ; mais, en elle-même, cette forme ne peut énoncer sa propre loi que si on la rapporte à ses états antérieurs, aux changements dont elle est susceptible, aux modifications qui ne se produisent jamais. » (Foucault 1966, p. 306-307). Bref, seule l'historicité qui est la sienne donne une épaisseur au langage qui le fait échapper aux usages des locuteurs, ce que Foucault n'hésite pas à appeler un « être »...

29. Voir par exemple, pour les approches connexionnistes de ce problème, Laks 1996, et pour les approches catastrophistes, Petitot 1985, qui est sans doute celui qui en a fait le plus clairement le cœur de sa perspective théorique.

30. Ainsi, dans une note manuscrite sur les légendes – plusieurs fois citée par Lévi-Strauss lui-même (Lévi-Strauss, 1968, p. 259, et 1995, chap. X) – Saussure écrit : « Il est vrai qu'en allant au fond des choses on s'aperçoit dans ce domaine, comme dans le domaine parent de la linguistique, que toutes les incongruités de la pensée proviennent d'une insuffisante réflexion sur ce qu'est l'identité ou les caractères de l'identité lorsqu'il s'agit d'un être inexistant comme le mot ou la personne mythique ou une lettre de l'alphabet, qui ne sont que différentes formes du SIGNE, au sens philosophique, mal aperçu il est vrai de la philosophie elle-même » (Saussure, 2003 : 387).

31. Ce sont là des termes techniques, définis par Saussure, mais que les éditeurs du cours n'ont pas du tout repérés.

32. Pour une démonstration détaillée de la validité philologique de cette interprétation et une explication précise, je ne peux que renvoyer à Maniglier 2006. Je me contenterai ici d'une citation : « Un système linguistique est une série de différences de sons combinées avec une série de différences d'idées ; mais cette mise en regard d'un certain nombre de découpures dans la masse de la pensée engendre un système de valeurs ; et c'est ce système qui constitue le lien effectif entre les éléments phoniques et psychiques à l'intérieur de chaque signe. Bien que le signifié et le signifiant soient, chacun pris à part, purement différentiels et négatifs, leur combinaison est un fait positif : c'est même la seule espèce de faits que comporte la langue. (...) Deux signes comportant chacun un signifié et un signifiant ne sont pas différents, ils sont seulement distincts. Entre eux il n'y a qu'opposition. » (1972, p. 166-167)

33. Voir à ce sujet les remarques de Churchland (1989, p. 174 sq.) sur la manière dont le logiciel de lecture à voix haute NETtalk retrouve dans la distance entre les patterns d'activation par lesquels il réagit à telle ou telle lettre singulière, la hiérarchie catégorielle du système phonologique de l'anglais, sans qu'il ne soit jamais représenté.

34. On peut se reporter au schéma proposé par Saussure lui-même illustrant une sorte de début de réseau (1972, p. 175).

35. cf. Jakobson : « Saussure attribuait aux changements de sons un caractère aveugle, fortuit, étranger au système de la langue. L'expérience nous montre au contraire que les changements ne peuvent être compris qu'en fonction du système phonologique qui les subit. » (Jakobson, 1976, p. 62-63). Lévi-Strauss en écho : « Nous ne songeons donc pas à reprendre sous sa forme première la distinction introduite par le Cours de linguistique générale entre l'ordre synchronique et l'ordre diachronique, c'est-à-dire l'aspect même de la doctrine saussurienne dont, avec Troubetzkoy et Jakobson, le structuralisme moderne s'est écarté le plus résolument ; celui, aussi, à propos duquel des documents récents montrent comment les rédacteurs ont pu parfois forcer et schématiser la pensée du maître. » (Lévi-Strauss, 1973, p. 26).

36. « On pourrait caractériser le milieu russe comme hostile au positivisme (...) Une plus ou moins forte dose de principes essentiellement incompatibles a constamment été mêlée à la position téléologique traditionnelle et au structuralisme de la science russe. » (Jakobson 1929, p. 55-56 ; traduction in Sériot 1999, p. 280-281)

37. Auteur d'une des premières tentatives pour définir le structuralisme comme mouvement, Cassirer (1945) lui aussi, comme Jakobson, l'inscrira dans la tradition de l'esprit objectif qui voit dans les systèmes culturels des formes d'organismes (en un sens plutôt spirituel que biologique). C'est contre cette interprétation que, préventivement pour ainsi dire, Saussure nous avait avertis...

38. La thèse selon laquelle les sociétés sont des « superorganismes » a été formulée notamment par l'anthropologue Kroeber dans un article resté célèbre (cf. Kroeber, 1917).

39 Ce concept a été développé pour rendre compte du statut des modélisations connexionnistes, dont P. Smolensky défend qu'elles ne sont ni des représentations réalistes des processus neuronaux supportant effectivement les opérations cognitives des hommes, ni des logiciels constitués de symboles dotés de valeurs fonctionnelles et indépendantes de tout support, comme dans les « programmes » des approches cognitivistes classiques. L'idée est que les unités représentées par les réseaux par chaque neurone ne correspondent ni à un symbole doté d'une valeur logique déterminée, ni même à une partie de symboles, mais à des traits déterminant l'émergence à un niveau supérieur de symboles qui n'entrent pas, eux, dans le calcul, mais se contentent d'en résulter à un niveau passif, perceptif, global. Ainsi le calcul opère au niveau local et, pourrait-on dire, micro-logique, alors que les formes s'aperçoivent à un niveau global mais ineffectif, à un niveau macro-logique. Cette distinction du micro et du macro me semble devoir être mise au cœur de l'épistémologie des sciences de l'esprit.

BIBLIOGRAPHIE

Bates, E., Elman, J. (1993). Connectionism and the study of change. In Mark Johnson (Éd.), *Brain Development and Cognition: A Reader*. Oxford: Blackwell Publishers (Pp. 623-642.)

Botlanski, J-E. (1995). *La linguistique diachronique*. Paris : PUF, « Que sais-je ? ».

- Bouquet, S. (1997). *Introduction à la lecture de Saussure*. Paris : Payot et Rivages, « Bibliothèque scientifique Payot ».
- Brandom, R. B. (1994). *Making it explicit : reasoning, representing, and discursive commitment*. Cambridge, Mass. : Harvard University Press, 1994.
- Brooks, R. & Steels, L. (Eds.). (1995). *The 'artificial life' route to 'artificial intelligence'. Building situated embodied agents*, New Haven: Lawrence Erlbaum Associates.
- Cadiot, P & VISETTI, Y-M. (2001), *Pour une théorie des formes sémiotiques ; Motifs, profils, thèmes*. Paris : PUF, « Formes sémiotiques ».
- Cassirer, E. (1945). Structuralism in Modern Linguistics. In *Word*, 1(2): 99-120.
- Chomsky, N. (1969). *La linguistique cartésienne* (N. Delanoë & D. Sperber, Trad.). Paris : Seuil (Edition originale, 1966).
- Chomsky, N. (1991). *Théorie du gouvernement et du liage : Les conférences de Pise*. Paris : Seuil, « Travaux linguistiques » (Edition originale : *Lectures on government and binding, The Pisa Lectures*. Dordrecht (Netherlands): Foris, 1981).
- Churchland, P. M. (1989). *À neurocomputational Perspective, The nature of mind and the structure of science*. Cambridge (Mass.) : MIT Press.
- Dennett, D. C. (1993). *La conscience expliquée* (trad. P. Engel). Paris : Odile Jacob. (Éd. originale, 1991.)
- Descartes, R. (1953). *Discours de la méthode*. In *Œuvres et lettres*, Paris : Gallimard « Bibliothèque de la Pléiade ».
- Descombes, V. (1996). *Les institutions du sens*. Paris : Minuit.
- Drestske, F. (1986). "Misrepresentation". In Gunderson K ed., *Belief: Form, content and function*, Minneapolis, University of Minnesota Press (p. 131-193).
- Dumit, J. (2000). Artificial participation: An interview with Warren Sack. In George E. Marcus, Editor, *Zeroing in on the Year 2000: The Final Edition (Late Editions, 8)*. Chicago: University of Chicago Press.
- Durkheim, E. (1912). *Les formes élémentaires de la vie religieuse, Le système totémique en Australie*. Paris : « Quadrige », PUF.
- Elman, J & A. (1996), *Rethinking Innateness, A Connectionnist perspective on development*. Cambridge (Mass.) & London: MIT Press.
- Fehr, J (2000). *Saussure entre linguistique et sémiologie* (trad. Pierre Caussat). Paris: PUF, « Sciences, Modernités, Philosophies » (éd. orig. Suhrkam Verlag, Frankfurt am Main, 1997).
- Fodor, J. (1983). *The modularity of mind, an essay on faculty psychology*. Cambridge (Mass.), MIT Press.
- Fodor, J. (2000). *The mind doesn't work that way: the scope and limits of computational psychology*. Cambridge (Mass.) : The MIT press.
- Foucault, M. (1966). *Les mots et les choses*. Paris : Gallimard, « Bibliothèque des Idées » (rééd. « TEL »).
- Franchi, S. & Bianchini, F. (Eds). (à paraître en 2008). *Toward an archology og Artificial Intelligenece*. Springer.
- Franchi, S. & Bianchini, F. (Eds). (2007). *Verso un'archeologia dell'intelligenza artificiale*. Macerata : Quolibet.
- Franchi, S (1997). *Endgames: Games and play and the end of philosophy*. Ph D Dissertation, 1997.
- Franchi, S. & Güzeldere, G. (2005). *Machinations of the Mind*. In S. Franchi and G. Güzeldere, *Computational Minds and Mechanical Bodies: Artificial Intelligence from Automata to Cyborgs*, Cambridge: MIT Press/Bradford Books.
- Gardner, H. (1985). *The Mind's new science, A history of cognitive revolution*. New York: Basic Books.
- Godel, R. (1957). *Les sources manuscrites du Cours de Linguistique générale de F. de Saussure*. Genève-Paris : Droz-Minard.

- Jakobson, R. (1929a), « Uber die heutigen Voraussetzungen der russischen Slavistik », *Slavische Rundschau*, Prague, 1, p. 629-646.
- Jakobson, R. (1931), « Principes de phonologie historique », *Travaux du Cercle Linguistique de Prague*, IV.
- Jakobson, R. (1976), *Six leçons sur le son et sur le sens*, Paris, Minuit, « Arguments », 1976.
- Karsenti, B. (2006a). *Politique de l'esprit, Auguste Comte et la naissance de la science sociale*. Paris : Hermann.
- Karsenti, B. (2006b). Le "dilemme durkheimien" en sociologie morale. In B. Karsenti, *La société en personne, Études durkheimiennes*. Paris : Economica.
- Kroeber, A. L. (1917). The Superorganic. In *American Anthropologist*, April-June, 1917 Vol.19(2):163-213.
- Labov, W. (1976), *Sociolinguistique*, trad. Encrevé, Pierre, Paris, Minuit, « Le sens commun », (*Sociolinguistics Patterns*, University of Pennsylvania Press, 1972).
- Lakoff, G. (1987). *Women, fire and other dangerous objects What Categories reveal about the Mind*. Chicago: University of Chicago Press.
- Laks, B. (1996). *Langage et cognition, L'approche connexionniste*. Paris: Hermès, 1996.
- Langacker, R. (1987), *Foundations of cognitive grammar*. Stanford (Ca): Stanford University Press.
- Laugier, S. (2006). Où se trouvent les règles. In S. Laugier et C. Chauviré (eds.), *Lire les Recherches Philosophiques*. Paris : Vrin.
- Levi-Strauss, C. (1962). *La pensée sauvage*. Paris : Plon, « Agora ».
- Levi-Strauss, C. (1968). *Mythologiques ***, L'origine des manières de table*. Paris : Plon.
- Levi-Strauss, C. (1973). *Anthropologie structurale deux*. Paris : Plon.
- Levi-Strauss, C. (1983) *Le regard éloigné*. Paris : Plon.
- Levi-Strauss, C. (1995). *L'identité, Séminaire interdisciplinaire 1974-1975*. Paris : PUF, « Quadrige ».
- Maniglier, P. (2006a). Les langues comme milieux : une hypothèse sur la cognition culturelle. <http://www.diffusion.ens.fr/index.php?res=conf&idconf=1583>
- Maniglier, P. (2006). *La vie énigmatique des signes, Saussure et la naissance du structuralisme*. Paris : Léo Scheer.
- Maniglier, P. (2006a). « L'approche des formes sémantiques dans l'héritage du structuralisme », sur invitation, séminaire « Formes symboliques », dir. Yves-Marie Visetti et Jean Lassègue, Ecole Normale Supérieure (Ulm), le 24 octobre 2006.
- Maniglier, P. (2006b). « Les langues comme milieux : une hypothèse sur la cognition culturelle ». Conférence prononcée au séminaire « Lundis de la Philosophie : Langage et cognition », dir. Francis Wolff, Paris : Ecole Normale Supérieure (Ulm), 30 octobre 2006. Conférence en ligne sur le site de l'E.N.S., « Diffusion des Savoirs » : <http://www.diffusion.ens.fr/index.php?res=personnes&idpers=1552>.
- Petitot, J. (1985a). *Les catastrophes de la parole, De Roman Jakobson à René Thom*. Paris : « Recherches interdisciplinaires », Maloine.
- Petitot, J (1985b). *Morphogénèse du sens*. Paris : « Formes sémantiques », PUF.
- Pollock, J-Y (1997). *Langage et cognition: Introduction au programme minimaliste de la grammaire generative*. Paris : PUF, « Psychologie et science de la pensée ».
- Port, R. F & Van Gelder, T. (1995), *Mind as motion : explorations in the dynamics of cognition*. Cambridge (Mass.), London (GB) : MIT Press.
- Prince, Alan & Smolensky, Paul (2004). *Optimality Theory: Constraint interaction in generative grammar*, Malden (Mass.) : Blackwell Pub., 2004. (Edition originale, 1993)
- Rastier, F. (1987). *Sémantique interprétative*. Paris : PUF, « Formes sémiotiques ».

- Rickert, H. (1997) : *Sciences de la culture et sciences de la nature* (trad. C. Prompsy et M. de Launay). Paris: Gallimard, 1997 (Edition originale : *Kulturwissenschaft und Naturwissenschaft*, Tübingen : J. C. B. Mohr, 1910).
- Rumelhart, D., McClelland, J L., & the PDP Research Group (1988). *Parallel Distributed Processing*. Cambridge : MIT Press.
- Salanskis, J-M. (2003). *Herméneutique et cognition*. Lille : Presses Universitaires du Septentrion.
- de SAUSSURE, Ferdinand (1967). *Cours de linguistique générale*, édition critique par Rudolf Engler. Wiesbaden : Otto Harrassowitz.
- de Saussure, F. (1972). *Cours de linguistique générale*. Paris : Payot. (Edition originale : 1916)
- de Saussure, F. (2002). *Ecrits de linguistique générale*. Paris : Gallimard, « Bibliothèque des idées ».
- de Saussure, F. La légende de Siegfried et l'histoire burgonde (présentation et édition de Béatrice Turpin). In Bouquet, S (ed.), *Saussure*, Paris : Editions de l'Herne (pp. 351-429).
- Searle, J R. (1998). *La construction sociale de la réalité*. Paris : Gallimard, 1998 (Edition originale, 1995.)
- Seriot, P. (1999), *Structure et totalité, Les origines intellectuelles du structuralisme en Europe centrale et orientale*, Paris, P.U.F., « Linguistique nouvelle » 1999.
- Smolensky, P. (2003). Le traitement approprié du connexionnisme. In D. Fisette & P. Poirier, *Philosophie de l'esprit, Problèmes et perspectives*. Paris : Vrin. (Edition originale, 1988).
- Steels, L & Oudeyer, P-Y. (2000). The cultural evolution of syntactic constraints in phonology. In *Artificial Life VII: Proceedings of the seventh International Conference on Artificial Life*, p. 382-394. Cambridge (Mass.): The MIT Press. (<http://arti.vub.ac.be/~steels/puzzle.pdf>)
- Steels, L. (2000). The puzzle of language evolution. In *Kognitionswissenschaft*, 8(4):143-150, 2000 (<http://arti.vub.ac.be/~steels/puzzle.pdf>).
- Turing, A. (1950/1995). Les ordinateurs et l'intelligence (trad. P. Blanchard). In A. Turing & J.-Y. Girard, *La machine de Turing*. Paris: Le Seuil, "Points". (Éd. originale : *Computing Machinery and Intelligence*. In *Mind*, vol. LVI, n° 236, 1950.)
- Utaker, A. (2000). *La philosophie du langage, une anthropologie saussurienne*. Paris : PUF, « Pratiques théoriques », 2000.
- Van Gelder, T. (2003). Dynamique et cognition (trad. S. Lapointe). In D. Fisette & P. Poirier, *Philosophie de l'esprit, Problèmes et perspectives*. Paris : Vrin. (Edition originale, 1998).
- Varela, F J. (1989). *Connaitre les sciences cognitives, Tendances et perspectives* (trad. P. Lavoie). Paris : Seuil (Edition originale, 1988).
- Visetti, Y-M & Cadiot, P. (2006). *Motifs et proverbes*. Paris : PUF
- Weinreich, U., Labov, W & Herzog, M I. (1968). Empirical Foundations for a Theory of Language Change. In Winfred P. Lehmann and Yakov Malkiel, eds., *Directions for Historical Linguistics*, Austin et Londres: University of Texas Press, 1968 (pp. 95-188).
- Wittgenstein, L. (2005). *Recherches philosophiques* (trad. F. Dastur, M. Elie, J.-L. Gautero & alii). Paris : Gallimard (Edition originale, 1953).