

Jean-Paul Baquiast

# Pour un principe matérialiste fort



DécohérenceS

**Compléments du livre**

**"Pour un principe matérialiste fort"**

## **Chapitre 3**

### **Le cerveau et la conscience**

**Chapitre 3, section 2 :**

**Le cerveau des sensations de Antonio Damasio**

**Chapitre 3, section 3 :**

**Modèles du cerveau La TSGN de Gérald Edelman**

**Chapitre 3, section 4 :**

**Modèles du cerveau. Le cerveau cortical associatif**

**Chapitre 3, Section 5 :**

**Modèles du cerveau. Gilbert Chauvet. Le PAAS appliqué au  
cerveau**

## Chapitre 3, section 2 :

### Le cerveau des sensations de Antonio Damasio

Antonio R. Damasio est chef du département de neurologie au Collège de médecine de l'Université de l'Iowa. Il est également professeur adjoint au Salk Institute de La Jolla. Il est aujourd'hui mondialement connu pour ses travaux sur le cerveau humain, dont il explore la complexité, notamment au regard de la mémoire, du langage et des émotions.

Dans ses écrits, traduits en de nombreuses langues, ce chercheur mêle étroitement l'expérience clinique, les études neurologiques, une imagination créatrice et une sensibilité philosophique et humaine remarquables. Antonio Damasio est une référence, non seulement dans le domaine des neurosciences, mais aussi pour tout ce qui concerne la simulation du moi et de la conscience sur des automates.

Son dernier livre, *Looking for Spinoza* (Traduit en français sous le titre *Spinoza avait raison : Joie et tristesse, le cerveau des émotions* . Odile Jacob 2005) développe les résultats expérimentaux mentionnés dans ses précédents ouvrages (Notamment *The Feeling of What Happens..* Traduit en français sous le titre *Le sentiment même de soi, corps, émotions, conscience*. Odile Jacob 2002) ainsi que leur philosophie générale Celle-ci a consacré la prise en considération des sentiments dans la compréhension des relations entre le corps, le cerveau et l'esprit. Mais *Looking for Spinoza* les enrichit encore. L'auteur y présente une synthèse achevée et harmonieuse de ses travaux et hypothèses. Ajoutons que, par ce livre, Antonio Damasio rend un hommage justifié à un philosophe longtemps méconnu, Benedict de Spinoza, étonnamment en avance sur son temps. L'hommage de Damasio à Spinoza est d'autant plus précieux qu'il met en relief ce que ne font pas les biographes ordinaires du philosophe hollandais, c'est-à-dire l'étonnante convergence entre les hypothèses de celui-ci relativement à l'origine de l'esprit, et les études modernes de neurophysiologie rapportées par l'auteur. Cette convergence en elle-même pose problème: comment un esprit qui ne connaissait rien à la science actuelle a-t-il pu si tôt faire œuvre de précurseur ?

La thèse de Damasio sur les origines de l'esprit et de la conscience est remarquablement cohérente et convaincante. Elle confirme et éclaire les théories, de plus en plus fréquentes aujourd'hui, montrant que les formes les plus élaborées de l'esprit et de la conscience humaine sont des acquis de l'évolution ayant émergé, selon les lois simples de la compétition darwinienne, dès l'aube de l'apparition de la vie sur Terre. Damasio, pour sa part, ne se livre pas sur ces questions à des

spéculations philosophiques. Ce qu'il avance s'appuie sur une série impressionnante d'observations cliniques ou permises par l'imagerie cérébrale fonctionnelle. Ces expériences n'apportent évidemment pas de preuves définitives, mais leur convergence permet de donner des fondements solides aux interprétations qu'en propose l'auteur. Ceux que le prétendu réductionnisme de Damasio scandaliserait (faire des sentiments conscients et de l'état du Moi lui-même la conséquence d'états bien définis du corps) doivent se dire que l'auteur n'avance rien sans preuves expérimentales. D'autres expériences à l'avenir nuanceront certainement les premières, mais pour le moment, nous pouvons nous reposer sur elles et ne pas mettre en cause de façon purement idéologique ses propositions.

Essayons de résumer en quelques paragraphes l'essentiel de sa démonstration. Elle semble si cohérente que nous pourrions parler en raccourci d'un véritable "système Damasio"

### **L'homéostasie**

Les organismes vivants se caractérisent d'abord, avant même leur capacité à se reproduire et à muter, par l'existence d'un "corps" assurant la permanence d'un milieu interne protégé de l'extérieur par une barrière. Des plus simples aux plus complexes, ils n'ont pu survivre qu'en maintenant ce milieu interne à l'abri des agressions de l'environnement. C'est ce que l'on appelle couramment l'homéostasie. L'organisme vivant est une "machine homéostatique" dont le métabolisme est assuré par des processus élémentaires acquis génétiquement et présents dès les formes les plus simples de cellules homozygotes. On retrouve ces mécanismes sans changements fondamentaux tout au long de l'échelle des organismes vivants.

Les stimulus et sensations

Les mécanismes assurant la survie et le métabolisme des organismes sont déclenchés par des stimulus externes (réception d'une phéromone provenant d'un partenaire sexuel possible, par exemple) ou internes (sensation de faim provenant de la baisse du dosage du sucre). Une chaîne de déclenchement (trigger) s'engage ensuite, jusqu'au cerveau, mobilisant les différentes ressources de l'organisme. On ne peut séparer conceptuellement le stimulus ou déclencheur et le mécanisme déclenché. L'un et l'autre co-évoluent en inter-relation.

### **Les réflexes**

Les corps ont été dotés progressivement par l'évolution d'organes de plus en plus complexes capables d'assurer les grandes fonctions d'alimentation, d'excrétion,

de reproduction, de fuite devant les prédateurs. Ces organes sont commandés par des réflexes de base (basic reflexes) déclenchés par les stimulus précités. Des dispositifs de contrôle coordonné de la bonne exécution de ces fonctions ont été sélectionnés par l'évolution, y compris chez les organismes les plus simples, notamment sous la forme d'échanges de messages chimiques. Dès le début, un système immunologique s'est développé pour assurer la protection contre les invasions extérieures. Ainsi s'est précisé ce qui était pour ces organismes le Bien (les facteurs leur permettant de se maintenir en vie et en bonne santé) et le Mal (les facteurs les conduisant à dépérir et mourir). Spinoza, rappelle Damasio, a qualifié de "conatus" la tendance, propre à la vie, de chaque organisme à persévérer dans son être, en faisant appel aux ressources nécessaires.

### **Les cartes corporelles cérébrales**

Avec la complexification croissante des organismes, des organes spécialisés dans le contrôle de l'homéostasie et dans le déclenchement des actions réparatrices sont apparus. Ce furent les systèmes nerveux. Une part importante des génomes, chez les organismes simples comme le ver ou la mouche comme chez l'homme, s'est trouvée dédiée à la programmation des processus d'entrée-sortie et de contrôle coordonnés permettant la surveillance des paramètres de bon équilibre et la mise en œuvre des procédures de survie : s'alimenter, se reproduire, élever les descendants, fuir les agresseurs, etc. Pour cela les cerveaux disposent de multiples cartes corporelles (body-map) qui permettent la synthèse des signaux provenant du corps. Le cerveau, en ce sens, élabore une image dynamique du corps analogue aux tableaux de bord des machines complexes. Il s'agit de connaître en temps quasi-réel l'état du système et d'engager immédiatement les actions réparatrices.

### **Les émotions**

Les stimulus permettant la mise en œuvre des différents processus vitaux et leur coordination par le cerveau s'organisent, à partir des organes des sens, en messages sensoriels de plus en plus élaborés (sensations) lesquels donnent naissance, au-delà d'un certain niveau d'évolution, à des tendances et appétits (drives, appetites) puis à des émotions d'appétence ou de rejet produisant des états corporels complexes. Dans la terminologie de Damasio, sensations et émotions ne sont pas nécessairement conscientes. Au contraire, dans la totalité des êtres vivants y compris chez l'homme, elles sont principalement inconscientes.

Les émotions ne sont pas des phénomènes gratuits, mais font partie essentielle de la mise en œuvre des processus vitaux. Elles ont été programmées par l'évolution génétique pour mobiliser le plus efficacement possible les ressources de

l'organisme au service du bon fonctionnement des organes sensoriels et effecteurs. Damasio les désigne du nom de "émotions proper", que l'on pourrait traduire par le terme de "tuteur émotionnel" ou "moteur émotionnel". Il distingue les émotions basiques, énergie, enthousiasme, malaise ; les émotions primaires : faim, plaisir, désir, peur et les émotions "sociales" résultant de l'exercice des précédentes dans la vie en société, laquelle est indispensable à la construction des individus, même des plus simples : orgueil, sympathie, indignation, etc. On a tout lieu de penser que même si les émotions sont difficiles à mettre en évidence chez les organismes relativement simples (insectes, mollusques), elles existent pourtant. En tous cas, on sait maintenant les observer à l'œuvre dans les espèces plus complexes, de la même façon que chez l'homme, même lorsqu'elles ne sont pas entrées dans le champ de la conscience. Dans cette optique, les émotions sont indispensables à la survie.

Les émotions, comme les sensations, mais à un niveau supérieur, se traduisent par diverses modifications corporelles. Celles-ci sont à la fois le signal permettant au cerveau de les enregistrer et le moyen dont dispose l'organisme pour affronter victorieusement les facteurs internes et externes visant à déstabiliser son homéostasie. Ainsi, manifester des signes de colère peut éloigner un adversaire. Là encore, ces modifications corporelles n'ont pas besoin d'être conscientes pour jouer leur rôle protecteur.

Les émotions, facteurs essentiels de la capacité de l'organisme à survivre dans un milieu nécessairement hostile, se déclenchent dès que l'organisme perçoit, sous forme de messages sensoriels simples ou complexes (sensations), les indicateurs internes ou externes signifiant le danger (le Mal) ou au contraire l'obtention d'un état d'équilibre (le Bien). Chaque individu est entouré de stimulus générant des émotions (emotionally competent stimulus, ECS) auxquels il réagit en permanence. L'identification de ces ECS est généralement programmée génétiquement (par exemple la méfiance à l'égard d'un objet non identifié). Mais beaucoup d'ECS sont les produits de l'expérience individuelle, que l'on pourra dire culturelle.

Les processus qui précèdent l'apparition des émotions, et celles-ci elles-mêmes, sont hérités génétiquement, du moins dans leurs grandes lignes. L'évolution individuelle de chacun (sa culture) se borne à individualiser et enrichir ces cadres génétiquement transmis. Les moteurs émotionnels ayant évolué pour optimiser les chances de survie des individus peuvent se révéler mal adaptés ou néfastes dans d'autres circonstances, notamment dans la vie en société moderne. Mais comment espérer que leurs déterminants génétiques puissent cesser d'agir? C'est là tout le problème du contrat social. Le contrat social, ensemble de règles de comportement définies pour que la société démocratique puisse fonctionner harmonieusement, s'oppose le plus souvent aux héritages génétiques, commandant par exemple la méfiance sinon l'agressivité à l'égard du voisin. Il faut alors un grand effort de discussions collectives et d'auto-persuasion pour que les membres du groupe dépassent leurs pulsions primaires.



## Les sentiments et les pensées

Chez les organismes dotés de pré-conscience ou de conscience, notamment chez l'homme, les mécanismes de survie précédemment décrits et générant des émotions, vont plus loin. Certaines émotions deviennent conscientes. On peut les appeler des sentiments (feeling). Ceux-ci, dans leurs formes extrêmes, prendront la forme de passions. Comment définir les sentiments, par rapport aux émotions, outre le fait qu'ils sont conscients et que celles-ci ne le sont pas ? Les sentiments correspondent à la perception d'un certain état du corps à laquelle s'ajoute la perception de l'état d'esprit correspondant, c'est-à-dire des pensées (thought) que le cerveau génère compte tenu de ce qu'il perçoit de l'état du corps. Les sentiments et les pensées ne viennent donc pas de nulle part, mais sont adaptés à la situation où se trouve l'organisme. Damasio rappelle que c'était là le point de vue de William James (1842-1910) (Sur William James, voir <http://www.emory.edu/EDUCATION/mfp/jphotos.html>\*) aussi méconnu en son temps que Spinoza : "le sentiment est la perception du corps réel modifié par l'émotion". C'est donc au sommet seulement de processus empilés (nesting principle) qu'apparaissent les sentiments. Du fait que ceux-ci sont conscients, leur importance a été surestimée, tandis que les mécanismes leur donnant naissance, restant inconscients, ont été ignorés ou peu étudiés.

Quel est alors le rôle des sentiments, en termes de sélection darwinienne ? Poser la question revient à poser la question du rôle de la conscience. La conscience, chez l'homme comme chez les organismes en présentant des formes simples, se construit sur la base d'émotions transformées en sentiments. Sert-elle à quelque chose ? On admet généralement que la conscience n'est pas un simple épiphénomène, mais permet d'organiser les sensations et les émotions du moment en les comparant les unes aux autres, et en les rapprochant de celles constituant la conscience biographique du sujet. La conscience mobilise et regroupe à tout moment dans un espace de travail commun un certain nombre d'informations nécessaires à la définition de stratégies de survie et à la prise de décision. Damasio s'est toujours efforcé de cerner le concept de proto-soi, de soi instantané, de soi biographique (c'est-à-dire capable de se rétrojecter dans le passé et se projeter dans l'avenir). Il reprend aujourd'hui ces hypothèses. Il fait reposer le soi sur une prise de conscience des émotions les plus fortes, c'est-à-dire aussi de certains des facteurs déclencheurs de ces émotions, ainsi que des modifications corporelles qu'elles entraînent. L'état de conscience en ce cas est d'abord une conséquence des émotions qui le précèdent, mais il agit en retour sur celles-ci, en favorisant la prise de décision commandant des comportements d'adaptation et les modifications corporelles qui leurs sont liées. Il peut s'établir à ce niveau une co-évolution ou interaction entre émotions, sentiments et comportements en découlant.

## Les idées

Les sentiments entrant dans le champ conscient génèrent aussi des comportements de type social. La conscience se construit principalement, dans le cerveau conscient, par le jeu des échanges langagiers et symboliques entre individus au sein des groupes. L'interaction entre émotions et sentiments se poursuit à ce niveau. On exprime un sentiment lui-même lié à une émotion, par l'échange d'une information symbolique ayant valeur de langage, signes ou mots. Ceux-ci s'organisent en opinions ou idées dès lors qu'ils respectent un certain formalisme grammatical. Ce faisant on peut communiquer avec les autres sur une base commune, puisque ceux-ci sont organisés génétiquement pour fonctionner d'une façon identique au sein de l'espèce.

A l'intérieur des groupes, les émotions et les sentiments s'expriment sous forme de comportements spécifiques, sélectionnés par l'évolution pour assurer la survie collective. C'est le cas de l'empathie par laquelle on comprend intuitivement (1) ce que ressent autrui. C'est aussi le cas des comportements dits altruistes ou moraux. Les individus y sacrifient un intérêt immédiat au profit d'un avantage plus lointain procuré par la survie du groupe que favorise ce sacrifice. L'établissement et le respect d'un contrat social permettant de sublimer les déterminismes génétiques primaires en découlent aussi.

D'autres comportements collectifs se traduisent par des échanges d'idées. Celles-ci, pour Damasio, ne sont pas inspirées par une rationalité abstraite. Elles expriment directement les émotions et sentiments des individus. Elles ne sont comprises et acceptées par les autres individus que si elles correspondent à leurs propres émotions et sentiments. Sinon, elles sont ignorées ou rejetées.

Toute cette évolution s'est construite par interaction entre les organismes et les milieux de plus en plus étendus auxquels ces organismes se sont trouvés confrontés en conséquence de l'accroissement de leurs possibilités corporelles. On se trouve là dans le paradigme de l'adaptation darwinienne le plus classique, sans qu'il soit nécessaire de faire appel à une quelconque finalité a priori.

1 : L'intuition est un mode de connaissance qui fonctionne sans faire appel à des raisonnements discursifs. Elle fournit des informations immédiates sur le monde. Très souvent, elle sert de précurseur à une véritable recherche rationnelle et scientifique. Mais souvent aussi, si elle n'est pas critiquée, elle peut donner naissance à des erreurs graves. Certains scientifiques pensent qu'elle procède du cerveau droit qui fonctionnerait sur le mode analogique, au contraire du cerveau gauche, le siège des raisonnements logiques.

## Chapitre 3, section 3 :

### Modèles du cerveau La TSGN de Gérald Edelman

Les études consacrées au cerveau humain ne se comptent plus, si bien qu'il est sérieusement envisagé de réaliser un ou plusieurs Atlas du cerveau permettant de rassembler et comparer les données, afin d'en tirer des informations que leur désordre actuel ne permet pas d'obtenir. Les approches et les modes d'observation sont multiples et se situent à tous les niveaux possibles, depuis l'analyse des molécules de liaison intersynaptiques jusqu'au cerveau global. Mais il reste difficile de passer de la description à l'explication, notamment quand il s'agit de comprendre l'origine, les modalités et l'utilité fonctionnelle des grandes propriétés du cerveau, l'intelligence et surtout la conscience. (1)

La conscience est pourtant, elle-aussi, l'objet d'une inflation d'études considérable, comme le montre la lecture des sommaires des revues internationales qui lui sont consacrées. Dans la mesure où on accepte le postulat du matérialisme scientifique selon lequel la conscience est une propriété émergente de l'organisation cérébrale ou neurale, l'étude de la conscience suppose inévitablement celle du cerveau. Le risque est alors de réductionnisme, se centrer sur l'anatomie et la physiologie des aires et des circuits cérébraux supposés impliqués dans les états précurseurs de la conscience, en perdant de vue la nécessité de constituer un modèle d'ensemble dans lequel on pourrait reconnaître ce que l'intuition commune attribue depuis des siècles à l'esprit humain et à la conscience.

Gerald Edelman ne prête pas à cette critique. C'est certainement dans le monde un des bons connaisseurs, aussi bien du cerveau proprement dit que de ses fonctions émergentes, notamment la conscience (2). Il dirige aux Etats-Unis le Neurosciences Institute et préside la Neurosciences Research Foundation. Il est également directeur du département de biologie du Scripps Research Institute. Ses travaux précédents en immunologie lui ont valu le prix Nobel de physiologie/médecine. Depuis 30 ans, il s'est consacré à l'étude des bases neurales des fonctions cérébrales supérieures : comment le cerveau a-t-il évolué pour produire ces fonctions, comment celles-ci se manifestent-elles aujourd'hui.

Ceci l'a d'abord conduit à élaborer une théorie générale, la Théorie de la Sélection des Groupes de Neurones, TSGN (Theory of Neuronal Group Selection; TSGN) , reposant sur le principe que face aux exigences de la survie s'imposant aux organismes, des cellules spécialisées, les neurones, eux-mêmes regroupés en faisceaux, ont été sélectionnés sur le principe de la compétition darwinienne : les groupes de neurones les plus aptes à assurer telle fonction étant retenus et inscrits

dans le patrimoine héréditaire. La compétition entre groupes de neurones règne également au cours du développement du cerveau chez l'individu jeune, l'adulte et même (Des études récentes montrent que le cerveau âgé, convenablement sollicité, garde de grandes capacités d'apprentissage) la personne âgée. Plusieurs groupes de neurones entrent en concurrence (dans certaines limites de spécialité évidemment) pour répondre à tel besoin, sans plan génétique déterminé à l'avance. Ceci explique la diversité des réponses possibles et leur redondance éventuelle ou dégénérescence (entendue au sens de convergence dans les réponses fournies par des organes différents. Voir ci-après Théorie de la sélection des groupes de neurones).

Gerald Edelman, étudiant les propriétés dites supérieures du cerveau, s'est également particulièrement attaché à comprendre le fonctionnement du cortex associatif, depuis longtemps considéré comme le siège de l'intelligence associative et de la conscience. Il a mis en évidence l'existence et le rôle des fibres qu'il a qualifié de réentrantes joignant à partir du système thalamocortical un très grand nombre des aires cérébrales spécialisées. Contrairement aux autres groupes de neurones, organisés soit en boucles fermées soit en faisceaux non remontants, les fibres réentrantes fonctionnent dans les deux sens, émettant vers une zone donnée et transmettant en retour des signaux provenant de cette zone. L'ensemble constitue un réseau dense interactif, couvrant la presque totalité du cerveau supérieur, un peu analogue au réseau Internet (3). On comprend bien qu'un tel maillage soit éminent favorable à l'émergence d'états associatifs plus ou moins volatils et se succédant rapidement tels que ceux identifiés dans la conscience primaire (commune à un grand nombre d'animaux et à l'homme) et dans la conscience supérieure, moins répandue puisqu'apparemment limitée à l'homme.

Tout ceci avait été exposé dans de nombreux articles et ouvrages. Mais il manquait une présentation facilement accessible de la théorie générale du cerveau et de la conscience qui a progressivement émergé de toutes ces études. On la trouve dans le dernier ouvrage de Gerald Edelman (*Wider than the sky, the Phenomenal Gift of consciousness, Plus vaste que le ciel, Une nouvelle théorie du cerveau* [Edelman, op.cit.]) que nous nous proposons ici d'analyser et commenter. L'auteur part du principe qu'il détient aujourd'hui un modèle explicatif global de la conscience permettant de résoudre les difficultés que pose encore la compréhension d'un phénomène paraissant rebelle à la description objective, dans la mesure où nous en sommes nous-mêmes issus (4). Dire que toutes les questions y sont résolues serait sans doute excessif, mais beaucoup de problèmes y sont éclaircis et des pistes pour tenter de résoudre ceux qui ne le sont pas sont offertes.

Il va sans dire que Gerald Edelman est un matérialiste et s'annonce comme tel, ce qui n'est pas sans courage dans un pays comme les Etats-Unis où les fondamentalistes religieux disposent d'une audience accrue, y compris de la part des institutions publiques. Mais ce n'est pas un matérialiste réducteur, puisqu'il postule que si l'esprit n'est rien sans le corps, il admet l'émergence de la complexité à partir

du simple et se donne les outils pour en traiter sans y voir le produit de déterminismes linéaires.

Quelles sont les propositions de Gerald Edelman relatives à la conscience ? Résumons-les ici, en y ajoutant nos commentaires.

### **Le matérialisme scientifique**

L'auteur s'inscrit explicitement dans le postulat du matérialisme scientifique: seule l'étude du fonctionnement des neurones peut donner naissance aux sensations, pensées et émotions subjectives, c'est-à-dire à la conscience. Darwin avait été précurseur sur ce point comme sur beaucoup d'autres, puisque pour lui, contrairement à ce qu'avait affirmé son collègue Alfred Russell Wallace, les facultés de l'esprit devaient être des produits de l'évolution même si initialement ils n'avaient pas contribué directement à l'adaptation. Le langage serait né du développement du cerveau. En retour, il aurait accéléré le développement de celui-ci. La même démarche est proposée concernant les raisons de l'apparition de la conscience. Les bases neurales de la conscience, et non celle-ci, ont été initialement sélectionnées pour leurs contributions à l'adaptation. Ce sont elles qui ont une valeur causale dans le fonctionnement de l'organisme en vue de sa survie. A partir de ces bases s'est construit l'état conscient subjectif. Le livre va s'efforcer de montrer comment ceci s'est fait..

Edelman introduit ainsi une de ses propositions importantes. La conscience n'est pas apparue toute armée chez l'homme. De plus, ses formes primitives elles-mêmes n'avaient pas de valeur adaptative. Le cerveau s'est développé selon certaines structures qui, elles, avaient valeur adaptative. Ces structures ont progressivement servi de bases (les bases neurales) aux premières formes de conscience. C'est alors seulement que la conscience primaire s'est révélée utile à la survie des espèces qui en étaient dotées, ce qui a produit en retour le renforcement des bases neurales.

Pour Edelman, on ne peut progresser dans la compréhension de la conscience qu'en la faisant descendre du piédestal où les philosophies idéalistes l'avaient mise. C'est un peu de la même façon, en détrônant l'homme de la place centrale qu'il s'était attribué dans l'univers, que l'astronomie puis la cosmologie ont pu devenir des sciences.

L'auteur rappelle le fait que la conscience dépend entièrement du cerveau, et non d'autres organes. Quand celui-ci est altéré, elle l'est aussi. Il n'y a pas de survie de la conscience après la mort. Mais les spiritualistes objectent: " qu'en savez-vous ? Vous ne recherchez pas scientifiquement de preuves de l'existence de consciences non liées à des corps, parce que vous postulez que cela est impossible ". Les

matérialistes répondent : “ si de tels faits avaient été scientifiquement observables, ils auraient depuis longtemps été observés”. Il n’y a pas d’ostracisme métaphysique systématique à l’égard de l’hypothèse spiritualiste.

### **La conscience est un processus**

Par ailleurs, pour Edelman, la conscience est un processus et non une chose. On ne peut identifier ni dans le corps ni dans le cerveau de neurones spécifiques qui seraient le siège de la conscience. Ce processus résulte de l’activité de populations de neurones réparties dans de nombreuses aires du cerveau. C’est aussi un processus propre à chaque individu, puisque chaque humain dispose d’un corps et d’un cerveau non partageable et non exactement semblable à celui des autres humains. Par ailleurs, la conscience est continue, intentionnelle (renvoyant en général à des objets ou concepts mémorisés). Elle est aussi partielle et son champ est fonction du degré d’attention. Mais l’attention, qui focalise la conscience, n’est pas la conscience. La conscience enfin est unitaire ou intégrée, sauf en cas de troubles du cerveau. Elle est faite de scènes unitaires se succédant à un rythme rapide mais intégrant des expériences passées (le présent remémoré de Edelman).

Qu’est-ce que l’attention ? L’attention est l’aptitude à sélectionner consciemment certaines caractéristiques dans un large éventail de signaux sensoriels présentés au cerveau. Qu’est-ce qui la provoque ? Un animal peut-il être attentif, sans être conscient de l’être ? Oui, car il s’agit d’un produit des bases neurales de la conscience primaire essentielle pour la survie. Mais comment moi, puis-je sélectionner consciemment les objets de mon attention ? Qui agit en moi pour m’obliger, par exemple, à rester attentif à la poursuite de la rédaction de ce livre ? Il nous semble que ce thème de l’attention est traité trop rapidement par Edelman. Nous pourrions en dire de même du thème de l’heuristique, c’est-à-dire de la recherche attentive de solutions aux problèmes qui se posent à l’homme, et même la recherche de problèmes là où il ne semble pas y en avoir. Ce comportement, chez le chercheur scientifique, est ressenti comme conscient et même volontaire. Mais il s’agit sans doute d’une illusion. Quelles en sont alors les bases neurales sous-jacentes ?

### **Conscience primaire et conscience supérieure**

Edelman rappelle que nous sommes également conscients d’être conscients. Mais il faut distinguer absolument entre la conscience primaire et la conscience supérieure. La conscience primaire consiste à avoir des images mentales dans le présent. Elle ne s’accompagne pas du sentiment d’un soi doté d’un passé et d’un

futur. Elle est faite essentiellement de ce que Edelman a nommé dans ses ouvrages précédents le “ présent remémoré” (rappel d’expériences passées). La conscience primaire est commune à l’homme et à de nombreux animaux (sinon à tous les animaux).

La conscience supérieure est conscience d’être conscient, conscience de ses actes et de ses intentions. Au niveau élémentaire, elle exige l’aptitude sémantique, capacité de créer un symbole en associant une signification à une représentation. Au niveau développé, elle nécessite l’aptitude linguistique. Les primates en ont quelques rudiments mais seuls les humains sont dotés de l’ensemble de ces possibilités. Par la vie en société et le langage, porteurs de connaissances mémorisées au niveau social, ils ont construit des strates successives leur permettant d’accéder à des capacités conscientes de plus en plus élaborées.

L’état conscient, y compris au niveau de la conscience primaire, permet de ressentir les « qualia », c’est-à-dire d’attribuer des qualités subjectives à certaines perceptions. On ne peut décrire les qualia en termes objectifs, précisément parce qu’ils sont subjectifs, c’est-à-dire résultant du fonctionnement de corps et de cerveaux individuels, différents les uns des autres. Les qualia s’enchaînent les uns aux autres. On ne peut être conscient d’un qualia isolé. Les qualia servent à opérer des discriminations d’ordre supérieur, utiles à la survie : par exemple distinguer un son renvoyant à un objet extérieur menaçant d’un autre renvoyant à un objet inoffensif.

Nous pouvons sans difficulté, nous semble-t-il, suivre Edelman quand il impose de distinguer systématiquement la conscience primaire, très répandue dans le monde animal, et la conscience supérieure, limitée aux humains et peut-être à quelques animaux supérieurs. Il y a encore beaucoup de théoriciens de la conscience qui ne font pas cette distinction, peut-être pour des raisons idéologiques, car ils seraient alors obligés de prendre en compte l’existence de la conscience primaire au sein d’espèces réputées très primitives.

En amont de l’une et l’autre de ces consciences se trouvent des bases neurales plus ou moins évoluées qui leur servent de support. Les bases neurales font indiscutablement partie du corps. Evoquer la hiérarchie : - bases neurales, conscience primaire, conscience supérieure - a pour effet de réintroduire la conscience dans les phénomènes biologiques. On pourrait dire qu’il s’agit de la “ réincorporer ”. C’est ce que de leur côté font les roboticiens. Ils ne conçoivent plus de conscience artificielle sans l’implanter dans le corps sensible d’un robot (Voir Chapitre 4).

## **La réétranance**

Edelman montre que c'est l'anatomie du cerveau qui permet de comprendre l'apparition de la conscience : description globale des régions du cerveau, modes de fonctionnement des neurones et des synapses, organisation des trois grands systèmes neuroanatomiques constituant l'architecture générale du cerveau. Il est important de visualiser le mode de fonctionnement de chacun de ces systèmes. Le système thalamocortical, par exemple, assure la connexion entre les différentes aires corticales par l'intermédiaire de fibres " réétranantes " (Voir paragraphe suivant). Par ailleurs, il faut se souvenir de ce que les cerveaux ne sont pas des machines analogues à l'ordinateur, c'est-à-dire construits tous de la même façon et appliquant des programmes identiques. Chaque cerveau est différent des autres et son fonctionnement suppose l'intégration dynamique de nombreuses aires différentes, tant dans le cerveau lui-même que d'un individu à l'autre.

Le terme de « fibres réétranantes » a été forgé par Edelman pour désigner des fibres massivement parallèles fonctionnant dans les deux sens et connectant les cartes construites dans chacune des aires corticales. Ainsi s'établissent des processus synchrones dynamiques mettant en cohérence les contenus de ces aires. Nous verrons que la réétranance est une propriété essentielle à sa théorie du cerveau (Gilbert Chauvet a précisé cette notion de réétranance, qu'il a particulièrement étudiée. Pour lui il ne peut y avoir réétranance stricte en raison de la non-symétrie (voir chapitre 2, section 6)). Depuis longtemps, les anatomistes avaient observé les fibres de liaison dans le cortex dit justement associatif, mais ils n'en avaient pas tiré de conclusions bien précises relativement à la génération des états de conscience. Le terme de réétranance est inspiré de l'informatique, comme quoi, même si les cerveaux ne sont pas selon Edelman analogues à des ordinateurs séquentiels, on peut y retrouver certains traits caractéristiques des traitements informatiques en feed-back.

## **La Théorie de la Sélection des Groupes de neurones**

Le cerveau n'a pas été conçu tout d'une pièce. Il s'est formé au cours d'une longue évolution soumettant ses différentes structures à la sélection naturelle. La théorie du cerveau proposée par Edelman dans ses différents ouvrages repose sur le darwinisme global ou théorie de la sélection de groupes de neurones (TSGN) déjà citée. Seuls les modèles sélectionnistes fondés sur le raisonnement en termes de population lui paraissent pouvoir expliquer que les cerveaux fonctionnent selon des procédures globalement proches malgré l'ampleur considérable des variations individuelles qu'ils présentent. La variabilité neurale ne traduit pas un défaut du système, mais est fondamentale. Elle permet que parmi ces populations de variantes



soient sélectionnés à tous moments les éléments les plus aptes à produire les comportements nécessaires à la survie. Ces éléments n'auraient pu être programmés une fois pour toutes compte tenu de l'ambiguïté des entrées d'informations provenant de l'environnement.

Comment le cerveau produit-il alors des réponses cohérentes et structurées ? La TSGN répond à la question. Cette sélection à partir d'un très grand nombre de variantes s'applique à trois niveaux : celui du développement fœtal, celui de l'expérience acquise au cours du début de vie, celui de la réentrée en cours de vie assurant la coordination de nombreuses aires différentes. L'intégration par réentrée des différentes aires corticales est indispensable pour assurer la liaison (binding) entre ces aires, dont résulte la conscience. Comme elle n'a pu être programmée à l'avance, elle ne peut résulter que de la TSGN, sélection des groupes de neurones les plus aptes à assurer cette liaison.

La TSGN explique pourquoi les réponses du cerveau peuvent à la fois être versatiles et efficaces. C'est qu'elles sont " dégénérées ". La dégénérescence est la propriété qu'ont des éléments différents d'assurer une réponse identique, ou l'inverse. Il s'agit d'une propriété biologique très répandue, assurant la souplesse adaptative des êtres vivants (5).

La liaison entre les neurones ou groupes de neurones participant à l'établissement de ce que Bernard Baars (Baars, A Cognitive Theory of Consciousness, 1988. Le livre est consultable en ligne sur <http://www.nsi.edu/users/baars/BaarsConsciousnessBook1988/index.html>) appelle (d'une expression qui ne dit pas grand chose) l'espace de travail conscient. Celui-ci fait l'objet de diverses hypothèses : synchronisation par médiateurs chimiques ou par les champs électromagnétiques (sinon par des actions mettant en jeu des particules quantiques). Ces hypothèses n'ont jamais pu être vraiment vérifiées. Pour Edelman, il n'y a pas de véritable problème : au sein d'un réseau dense de fibres interconnectées dans les deux sens, des liaisons se produisent et viennent en concurrence jusqu'à ce que les plus efficaces l'emportent, avant d'être à leur tour remplacées. On pourrait peut-être comparer cela à certains effets globaux se produisant (sans que les utilisateurs en soient conscients) au sein du réseau Internet, sous l'influence de contraintes externes ou internes. Les flux s'y regroupent en faisceaux plus ou moins persistants, par exemple en conséquence des requêtes posées aux moteurs de recherche. Ainsi un site très souvent consulté apparaîtra automatiquement en tête de liste.

## La conscience primaire

Différents mécanismes cérébraux ont permis l'apparition de la conscience primaire, c'est-à-dire l'aptitude à construire une scène discriminante (spécifique). Le processus fondamental est l'aptitude à procéder à des catégorisations perceptives (découpages du monde par le cerveau). Celles-ci sont assurées par des interactions entre systèmes sensoriels et moteurs dans des " encartages " globaux assurés par des fibres réentrantes. Un encartage global est une structure dynamique contenant différentes cartes sensorielles liées par réentrées. Cette structure est à la base de la catégorisation. Les entrées sensorielles externes s'y conjuguent avec les entrées proprioceptives (internes). Le monde est ainsi " échantillonné " en fonction des activités de l'animal. Plusieurs encartages permettent de créer un concept, concept du mouvement vers l'avant par exemple.

Ceci, en l'absence de mémoire, ne suffirait pas à assurer l'adaptation. La mémoire est indispensable à la conscience, même primaire. Selon la TSGN, la mémoire est la capacité à répéter ou supprimer un acte spécifique. Elle résulte de modifications dans l'efficacité synaptique de différents groupes neuronaux, modifications qui incitent de façon dégénérée certains circuits à recommencer. Il y a plusieurs sortes de mémoire, à long ou court terme, qui supposent des modifications de force différente (La « vraie » mémoire à long terme impose une biosynthèse protéique donc des mécanismes autres.). Dans cette perspective, la mémoire n'est pas la reproduction d'un comportement à l'identique mais la façon de faire revivre (ou réactiver) non identiquement des comportements antérieurs. D'autres systèmes (systèmes de valeur construits lors du développement) modulent l'étendue des souvenirs. Ainsi, la peur ressentie pendant des bombardements commande, même après plusieurs années, l'importance attribuée à tel souvenir évoquant la guerre (6).

Quel est l'événement décisif de l'évolution ayant donné lieu à l'émergence de la conscience primaire? Ce serait l'apparition des connexions réentrantes du système thalamocortical, à la transition entre reptiles et oiseaux puis mammifères. Deux types de voies réentrantes permettant le traitement des signaux se seraient dégagés, celles distinguant le soi du non-soi. La perception y est associée à la mémoire dans de très courts intervalles de temps (le présent remémoré). Une scène consciente peut alors être créée en une fraction de seconde. Les informations provenant du soi y jouent toujours un rôle clé. L'aptitude à créer des scènes conscientes est utile à la survie (imaginer un prédateur à partir de quelques indices sensoriels).

La conscience supérieure s'est développée chez les primates sur ces bases, par association avec d'autres circuits réentrants permettant l'acquisition de la capacité sémantique et du langage. La conscience supérieure permet d'imaginer le

futur, de se remémorer le passé et d'être conscient d'être conscient. Mais la conscience primaire reste fondamentale. Sans elle, pas de conscience supérieure.

On pourrait penser que Gerald Edelman décrit les mécanismes de la conscience d'une façon arbitraire. Il donne l'impression d'avoir une idée a priori de ce qu'est celle-ci et de chercher à retrouver ensuite dans l'anatomie et la physiologie du cerveau les facteurs pouvant produire les phénomènes qu'il a définis. Mais le livre d'Edelman est un résumé de très nombreux travaux expérimentaux qui ont, semble-t-il, permis de tester les hypothèses initiales et de les organiser en théorie générale du cerveau. Aucune théorie n'étant définitive, celle proposée par Edelman est déjà critiquée dans des articles spécialisés que nous ne citerons pas ici. Mais elle nous offre une base de référence indispensable à connaître.

Nous sommes cependant obligés de constater que Gerald Edelman ne fournit pas d'indications sur les processus évolutionnaires ayant permis l'apparition chez les successeurs des reptiliens du système thalamocortical réentrant qui est décisif, comme il l'indique, pour la discrimination entre le soi et le non soi. Il ne donne pas davantage d'indications concernant l'apparition des autres systèmes neuroanatomiques nécessaires à la conscience primaire. On retrouve la question de fond posée par la théorie darwinienne de l'évolution. Comment des caractères précurseurs de systèmes favorables à une adaptation ultérieure peuvent-ils être sélectionnés alors qu'ils n'apportent aucun avantage dans un premier temps ? Une réponse possible, en termes darwiniens (C'est-à-dire en refusant, à nouveau, la facilité de l'Intelligent Design), serait que des caractères propres aux neurones, par exemple une propension à développer des axones dans diverses directions, auraient conduit à la mise en place au hasard de fibres réentrantes associatives, conservées dans les génomes même si elles n'apportaient pas d'avantages particuliers, jusqu'au jour où de telles fibres auraient rendu suffisamment de services fonctionnels pour être sélectionnées.

Une autre façon d'expliquer l'apparition de cerveaux organisés pour produire de la conscience primaire puis supérieure serait de faire appel aux expériences intéressantes de l'émergence de propriétés linguistiques dans des populations de robots. S'il s'avérait que de tels robots commencent à échanger des contenus sémantiques avant d'avoir, si l'on peut dire, le cerveau pour cela, on devrait pouvoir montrer, dans la suite de ces expériences, que ces échanges, devenant de plus en plus complexes, pourraient favoriser la sélection d'organisations matérielles et logicielles les mieux aptes à les traiter et à en générer d'autres, c'est-à-dire de cerveaux artificiels. On rejoindrait là l'hypothèse selon laquelle ce serait le langage et, avant lui, chez les animaux, les échanges sociaux à base de symboles, qui auraient entraîné le développement des cerveaux (Voir sur ces points les chapitres 4 et 5).

## Le noyau dynamique

Selon la TSGN étendue, toutes les expériences conscientes sont des qualia, autrement dit des discriminations personnelles dans des scènes complexes. Mais comment expliquer la richesse de chaque état de conscience et son unité ? Pour Edelman, comme pour beaucoup de neurologues, il faut faire appel aux propriétés des systèmes complexes (Voir chapitre 2). Un système complexe peut à la fois intégrer ses parties et prendre beaucoup d'états différenciés combinant les propriétés de ces parties. C'est le cas du cerveau. Ses réseaux interactifs manifestent une intégration fonctionnelle poussée (par exemple l'aire corticale responsable de l'orientation) puis grâce aux liaisons réentrantes ils deviennent intégrés au niveau supérieur, c'est-à-dire qu'ils acquièrent davantage de propriétés unitaires quand ils sont liés que quand ils ne le sont pas.

Cette description peut être appliquée au système thalamocortical. Il est dynamique et grâce au nombre considérable de ses connexions neurales, il change d'état en quelques fractions de seconde. Par ailleurs il est constitué d'un plus grand nombre d'interactions internes que d'interactions avec les autres parties du cerveau. Il " se parle principalement à lui-même ". On peut dire qu'il s'agit d'un noyau fonctionnel au service de la conscience. Edelman l'a nommé le " noyau dynamique ". C'est l'outil nécessaire aux propriétés unitaires et pourtant différenciées du processus conscient. Ses réponses peuvent aussi stimuler des systèmes non conscients donc moduler le comportement de l'organisme entier.

Les premières discriminations influençant le noyau dynamique proviennent des signaux du corps puis, au cours de la vie, du soi corporel. Mais il n'y a nulle part dans le noyau dynamique un observateur interne (un homoncule) qui pourrait apprécier son état instantané, même si nous avons nous-même l'impression d'être cet observateur. Cette impression est une construction culturelle qui s'est superposée, dans les sociétés occidentales, au mécanisme d'intégration des différents signaux endogènes et exogènes recueillis et traités par le cerveau (7).

Aussi convaincante que soit l'hypothèse d'un noyau dynamique qui soit le moteur principal de la fabrication du soi dans la conscience primaire puis supérieure, il faut bien avouer que l'hypothèse laisse en grande partie insatisfait. Comment un tel processus peut-il générer la conviction qu'éprouve le sujet conscient d'être un Je observateur mais aussi un Je acteur, doté de volonté ? Edelman répondra sans doute que cette sensation d'être un Je est un qualia créé au niveau de la conscience supérieure qui ne peut être décrit en terme de processus neural. On pourra dire aussi que le Je est une création récente de certaines sociétés, et qu'il en existe d'autres où les individus ne se perçoivent pas comme des Je, mais comme baignant dans une sorte de conscience diffuse, telle que la méditation peut en donner l'exemple.

Il reste que, les qualia en général et l'intuition d'être un Je au sein du champ conscient en particulier étant des phénomènes fondamentaux dans notre appréhension du monde, l'impossibilité précise de décrire la façon dont ils émergent et se manifestent à l'intérieur d'un système de réseaux d'informations tel que le noyau dynamique décrit par Edelman est très frustrante. Le problème, bien souligné par les spiritualistes, mais qui devrait trouver des réponses matérialistes, est le suivant : comment un mécanisme physique est-il transformé en sensation.

Pourrait-on espérer que les spécialistes des systèmes cognitifs artificiels puissent un jour proposer des modèles de traitement d'information (sans doute en réseaux multi-agents) plus convaincants que la description du noyau dynamique donnée par Edelman, avec une représentation du soi " vu de l'intérieur " dans laquelle nous pourrions entrer ? Avec un peu d'optimisme au regard des possibilités de la conscience artificielle, on pourrait envisager en effet de voir opérer un robot doté d'un Je artificiel dont nous comprendrions mieux les processus que ceux des cerveaux biologiques. Nous présenterons quelques perspectives en ce sens dans le prochain chapitre.

### **La " transformation phénoménale " n'est pas causale**

L'impossibilité de localiser dans le cerveau le Je conscient pose la question du rôle causal de la conscience. C'est pour Edelman le nœud de sa théorie. On a vu comment le processus conscient peut être causé par des processus neuraux, grâce aux interactions réentrantes centrées sur un soi servant de référence pour la mémoire, y compris dans la conscience primaire. L'activité du noyau dynamique convertit les signaux reçus du milieu extérieur ou provenant de l'intérieur du corps en ce que Edelman appelle une " transformation phénoménale " : qu'est ce qu'il en est d'être tel animal conscient doté de ce qualia spécifique qu'est la conscience de soi. Elle n'est pas causée par les processus neuraux mais l'accompagne. Quel est alors son rôle?

Pour Edelman, la transformation phénoménale (conscience de soi) ne peut être causale. Ce sont les processus neuraux qui la génèrent qui le sont. Mais elle est un indicateur plus ou moins fiable de la façon dont ces processus se déroulent. Cet indicateur sert d'abord au sujet lui-même. Il lui donne si l'on peut dire l'état d'un certain nombre de ses compteurs corporels. De plus, si cet état peut être communiqué aux autres par un langage quelconque, le groupe tout entier en bénéficie. Prenons un exemple simple. Si je glisse sur une pente verticale, le fait d'être conscient de ma chute ne m'empêchera pas de tomber. Mais prendre conscience de cet accident me permettra peut-être une autre fois de ne pas me mettre dans la situation de tomber à nouveau. Par ailleurs, je pourrai signaler à mes compagnons que je suis en train de tomber, ce qui leur évitera de s'avancer comme

je l'avais fait sur le bord de la falaise. Compte tenu de tels avantages, les processus neuraux produisant la conscience de soi, conscience primaire puis conscience supérieure, ont été sélectionnés par l'évolution.

Edelman fait l'hypothèse que c'est principalement du fait de leurs bénéfices en termes de communication que les processus générant la transformation phénoménale (notamment le noyau dynamique) se sont développés au cours de l'évolution. C'est ce qu'illustre l'exemple que nous venons de donner. Tout ce qui est causal provient de l'état du système thalamocortical et des autres systèmes neuroanatomiques. Sans être directement causale, la conscience de soi produit par le noyau dynamique, sous-ensemble du système thalamocortical, serait avant tout un véhicule de communication faisant connaître à l'extérieur, notamment aux semblables du sujet, ce qui se passe au sein de l'organisme de celui-ci.

Les processus décrits sont conformes semble-t-il aux observations neurologiques, y compris en ce qui concerne le temps de retard entre une action (ou décision) engagée au niveau de l'organisme et la conscience que l'on peut en avoir. Ces observations comme on le sait, donnent des arguments à ceux pour qui le libre-arbitre est une illusion. Comment faire coexister le déterminisme au niveau des décisions prises par le corps et la liberté au niveau d'une prise de conscience survenant après que la décision du corps ait déjà été mise en œuvre ? (8)

Nous pouvons ainsi dire que Edelman "réincorpore" la conscience. Sans en faire un processus sans influence, un simple épiphénomène, il en fait une des modalités par lesquelles le corps manifeste ses décisions et amplifie leurs effets. Ceci notamment dans le monde des informations communicables par le langage. De la même façon, quand le corps prend une décision, cette décision s'accompagne généralement d'une action musculaire qui entraîne une conséquence sur le monde.

Dire que le corps décide ne signifie évidemment pas que le corps jouit d'un libre-arbitre quelconque. Le corps est "décidé" par de nombreux déterminismes non linéaires, bien décrits par Alain Berthoz, comme nous l'avons vu. Il en résulte que, dans une très large mesure, la conscience réincorporée est décidée par ces mêmes déterminismes, avec un niveau de complexité supplémentaire apportée par ce qui se passe au niveau de l'individu conscient.

Le corps est évidemment aussi le siège d'un grand nombre de comportements inconscients médiatisés par le cerveau. Edelman étudie notamment le système des ganglions de la base et du cervelet, responsables d'automatismes inconscients, comme le contrôle du mouvement. Ils ne comportent pas de fibres réentrantes. Ils sont reliés au cortex mais celui-ci ne les commande que lorsque l'attention prend le relais des automatismes, par exemple à certains moments difficiles de la conduite automobile. Il existe plusieurs niveaux d'attention, sous commande du cortex. Dans certains cas, les ganglions de la base peuvent agir sur le

cortex. Cela pourrait être une façon d'expliquer l'inconscient freudien et le refoulement, notamment en ce qui concerne le retour du refoulé.

## **Langage et conscience**

La conscience primaire ne permet pas de se représenter le passé, le futur et soi-même comme conscient d'être conscient. La conscience supérieure le peut. Les animaux en semblent dépourvus, ce qui ne veut pas dire qu'ils n'ont pas de soi, ni d'image du passé dans le présent remémoré ni de mémoire à long terme. Ce qui leur fait défaut, ce sont les aptitudes sémantiques, c'est-à-dire l'utilisation de symboles pour donner du sens aux événements et raisonner sur eux en leur absence. Ceci montre l'importance essentielle du langage dans l'émergence de la conscience. On peut supposer en fait que les deux propriétés, aptitudes langagières et aptitudes à la conscience, ont co-émergé et co-évolué chez les hominiens. D'où l'intérêt de l'étude de l'émergence du langage chez les robots, évoquée dans le prochain chapitre.

Le langage ne se limite pas à l'utilisation de symboles, car il suppose l'aptitude syntaxique, c'est-à-dire la possibilité de former des phrases. Certains animaux comme les chimpanzés sont capables de certaines postures reposant sur la recherche de sens (par exemple se reconnaître dans un miroir) ce qui prouve qu'ils ont un début de conscience supérieure. Les humains ont des aptitudes sémantiques et syntaxiques étendues, se superposant à leur conscience primaire. On a identifié depuis longtemps les zones cérébrales y jouant un rôle, notamment l'hippocampe nécessaire à la mémoire épisodique, ainsi que les aires du langage dites de Broca et de Wernicke. Mais comment les aptitudes sémantiques et syntaxiques sont-elles apparues au cours de l'évolution ? Comment a-t-on pu découvrir qu'un geste, un son ou un objet pouvait tenir lieu d'une chose ? Sans doute par l'apparition de nouvelles voies et nouveaux circuits réentrants se superposant à ceux déjà existants du cortex, et donnant notamment un accès étendu à la mémoire. Par ailleurs nombreuses sont les hypothèses anatomiques et comportementales pouvant expliquer l'explosion du langage chez l'homme.

Edelman ne pose pas la question de savoir qui, dans la naissance du langage, a précédé et induit l'autre, la modification neuronale ou la modification anatomique et comportementale - étant entendu aussi que le langage n'existant pas encore n'a pu apporter ses bénéfices, initialement, à ceux qui en étaient dépourvus. On pourrait là de nouveau imaginer qu'une petite modification, soit neuronale, soit anatomique, soit même comportementale, ait apporté des bénéfices très réduits mais suffisamment significatifs en terme d'aptitudes sémantiques, le processus s'accélérait ensuite. Les paléanthropologues considèrent ainsi que la capacité des hominiens à créer et se transmettre des outils révèle l'existence chez eux du langage,

ou précède de peu l'apparition de celui-ci (9). Dans les expériences déjà citées d'émergence du langage chez les robots, c'est la pression de sélection en faveur de la communication qui a permis l'exaptation en faveur de la création d'un langage d'aptitudes sensori-motrices pré-existantes. Certaines hypothèses relatives à l'apparition du langage chez les hominiens vont dans le même sens. Des aptitudes sémantiques rudimentaires présentes chez les anthropoïdes, peut-être à la suite de mutations favorables dans les circuits réentrants, se seraient trouvées brutalement encouragées par un changement d'habitat rendant la communication linguistique indispensable à la survie. Ainsi, la conscience supérieure et ses bases neurales seraient un résultat de l'émergence du langage, lui-même étant un résultat de l'apparition d'une vie en société exigeant pour sa survie la communication symbolique d'individu à individu.

Mais peut-on être conscient sans langage ? Autrement dit, peut-on penser sans les mots ? Les théoriciens du " déterminisme linguistique " répondent par la négative. Edelman semble partager ce point de vue. On sait que la question est controversée.

Il est important pour comprendre les cultures d'admettre que la conscience supérieure est le produit du langage, lui-même produit d'une société obligée à communiquer et maîtrisant la production et l'usage d'outils. La vie sociale structurée par le langage donne naissance en effet aux grandes constructions symboliques dépassant largement les consciences individuelles (mythologies puis théories scientifiques). Ces constructions modèlent la façon dont se construisent les contenus conscients individuels, puis par l'intermédiaire des individus qu'elles mobilisent, elles modèlent le monde tout entier. Les processus impliqués paraissent très proches de ceux décrits par Edelman au niveau des cerveaux individuels. On pourrait parler d'une théorie de la sélection des groupes de neurones transposée au champ des méta-circuits et méta-représentations sociales.

Les considérations qui précèdent montrent, à notre sens, que l'étude de la conscience limitée à ce qui se passe au plan du cerveau individuel n'est pas suffisante. Il faut absolument étendre l'étude aux processus se déroulant au plan des super-organismes sociaux. Mais ceci oblige à bien d'autres considérations, notamment celles concernant le rôle éventuel des entités informationnelles autonomes, les mèmes, circulant sur les réseaux et susceptibles de formater les contenus de la conscience supérieure (Voir Chapitre 5. Susan Blackmore, théoricienne des mèmes, suggère que le Je serait en fait un mème qui se serait emparé du cerveau des hommes modernes.).

L'élargissement de l'étude de la conscience au niveau du super-organisme humain n'enlève évidemment rien à l'intérêt de comprendre ce qui se passe au plan des individus humains, de leur corps, de leur cerveau et de leurs contenus cognitifs. L'individu demeure en effet un agent essentiel de l'évolution du super-organisme, du fait qu'il est équipé pour générer de l'émergence à un rythme rapide et avec une efficacité d'action physique sur le monde considérable.



## La représentation

Pour Edelman, la représentation est le résultat de discriminations et catégorisations effectuées par le sujet conscient. Par exemple, je me représente la table que je regarde comme distincte du reste de la pièce. L'auteur ne veut pas en faire l'équivalent, trop souvent utilisé par ce qu'il nomme les psychologues cognitivistes, des structures neuronales (équivalentes à des informations dans un ordinateur) induites par les signaux venus de l'extérieur. Il s'agit là, dit-il, d'une description vue de l'extérieur ou objective qui perd de vue les sens et les intentionnalités qu'ont ces structures pour celui qui les héberge. De plus, pour lui, le substrat neural de la conscience n'est pas représentationnel. Des formes de représentation se produisent dans la conscience mais elles n'évoquent pas les états neuraux sous-jacents, mémoire, cartes perceptives par exemple. Ceci permet de ne pas lier les représentations, terme à terme, avec les états du cerveau ou les états de l'environnement. Des formes diverses de représentations, par exemple des images mentales, sont liées à des états divers de la conscience primaire et de la conscience supérieure, mais ne les déterminent pas. La cognition et l'intentionnalité de la conscience supérieure ne déclenchent pas nécessairement des images. Ceci permet d'éviter de traiter les représentations comme les données nécessaires au fonctionnement de l'ordinateur cérébral, lequel se livrerait sur elles à des calculs.

La signification, essentielle à l'intentionnalité, résulte du jeu de nombreux processus convergents qui enferment la représentation dans les circuits " dégénérés " résultant du fonctionnement des fibres réentrantes support de la conscience supérieure. Il n'y a pas une fonction s'appliquant à une représentation, comme dans l'ordinateur, mais des interactions multiples et changeantes dont beaucoup se passent de représentations. N'importe quelle représentation peut correspondre à de nombreux états neuraux sous-jacents et à de nombreux signaux différents reçus de l'extérieur. Edelman présente des expériences de magnéto-encéphalographie qui démontrent cette affirmation. C'est la diversité des faisceaux réentrants qui permet une telle convergence, nouvelle preuve apportée selon lui à la TSGN étendue.

Ceci montrerait qu'une grande partie de la psychologie cognitive perd de son intérêt, quand elle prétend attribuer des états fonctionnels équivalents à des informations de même nature codées dans les cerveaux et traités par des programmes computationnels identiques (10). Une très grande diversité et variabilité est la règle, non seulement au niveau des représentations et de leur rôle, mais au niveau des états neuraux sous-jacents. L'intentionnalité et la volonté dépendent de l'interaction des contextes locaux du milieu environnant, du corps et du cerveau.

D'une façon générale, Edelman diminue l'importance que l'on attribue généralement aux représentations. Il n'en fait certes pas des épiphénomènes mais des productions non immédiatement significatives. En d'autres termes, selon lui, il ne faudrait pas attribuer trop d'importance aux images du monde que nous hébergeons. C'est là un nouvel aspect de sa démarche générale : réincorporer la conscience dans les mécanismes neuraux sous-jacents, afin d'aller directement à ces derniers. Mais alors se pose la question du rôle de ces représentations, qu'il faudra bien expliquer pour comprendre leur apparition et leur survie au cours de l'évolution. Y a-t-il là quelque chose à voir avec l'imaginaire ou même avec le rêve ? Quel est le lien entre la représentation susceptible de prise de conscience et la catégorisation qui constitue la façon élémentaire dont l'organisme vivant s'inscrit dans le monde, ceci avant même qu'il ne dispose d'une conscience primaire ?

On ne doit pas non plus oublier que les représentations ne sont pas seulement des constructions individuelles. Elles sont aussi construites au cours des relations entre individus. Par exemple, la représentation d'un prédateur est construite au cours d'expériences vécues par le groupe ou transmise par lui. Elles jouent donc un rôle plus important que ne semble le penser l'auteur.

Ceci nous conduit au concept de concept, dont Edelman ne parle pas, sauf à dire que le concept désigne l'aptitude du cerveau à catégoriser ses propres activités et à construire un universel, ce qui ne nous paraît pas suffisant. Le concept est une des briques de base avec lesquelles se construisent les échanges langagiers au sein d'une collectivité. Pour simplifier, on pourrait dire que les concepts correspondent aux mots du langage verbal. Ils ne sont pas construits par des individus particuliers, à partir de leurs représentations. Ils émergent sur le mode darwinien des interactions multiples entre locuteurs ayant acquis des représentations globalement comparables et vérifiées par l'expérience collective. Ainsi, à force d'être exposés à la pluie, comme tous les êtres vivants, certains d'entre eux y ont associé une représentation individuelle que, dans leurs échanges, ils ont fini par nommer de la même façon, rassemblant sous ce nom un certain nombre de caractères statistiquement significatifs (humidité, froid, utilité pour l'agriculture, etc. ). Les concepts, en retour, contribuent à formater les représentations individuelles en les enrichissant de tous les sens donnés par la collectivité au phénomène désigné par le concept.

Le concept de concept, entraînant celui de loi reliant les concepts (loi scientifique, par exemple) nous conduit à la question de la construction des connaissances. Même si celle-ci est un phénomène collectif, les cerveaux individuels y contribuent directement. On trouve toujours, semble-t-il, un individu à l'origine de la qualification d'une entité observée. Il ne semble pas qu'aux origines, ce processus soit très différent de ceux intéressant la catégorisation, notamment au sein de la conscience primaire. Mais, dans la conscience supérieure, comment le cerveau observant les images qu'il reçoit du monde extérieur à partir de ses organes

sensoriels et de ses instruments, en fait-il des “ objets ” de connaissance scientifique ? Autrement dit, le processus décrit par Mme Mugur-Schächter sous le nom de MRC s’applique-t-il, consciemment ou inconsciemment ? (Voir chapitre 1 sur le rôle de l’observateur dans la construction de la « réalité ».).

Plus généralement, ne faut-il pas s’interroger sur le processus épistémologique d’acquisition et de contrôle des connaissances, incluant notamment l’induction et l’abduction (11), quand on étudie la conscience et les bases neurales de celle-ci ?)

1 : Pour ceux qui voudraient approfondir l’architecture générale du cerveau et ses grandes fonctions, ce que nous conseillons vivement, les ouvrages ne manquent pas. Voir *Le cerveau et la pensée : La révolution des sciences cognitives*, collectif coordonné par Jean-François Dortier Editions sciences humaines 2003. Pour les détails anatomiques, l’*Atlas du cerveau*, De Boeck, de Joseph Hanaway et al, 2000 reste une bonne référence, même s’il n’intègre pas les dernières observations permises par l’imagerie fonctionnelle cérébrale.

2 : Nous pensons néanmoins que les travaux de Gilbert Chauvet portant sur le cerveau et la conscience permettront d’aller plus loin que ceux de Gerald Edelman, car ils reposent sur une base théorique mathématique que n’ont pas les seconds.

3 : Les logiques de construction et de fonctionnement du réseau Internet dont la complexité s’accroît quotidiennement, fournissent de nombreux éléments de comparaison pouvant aider à comprendre le fonctionnement des réseaux biologiques complexes, que ce soit dans le cerveau, dans les échanges évolutifs entre espèces ou dans les sociétés humaines non encore informatisées. Est-ce à dire qu’Internet est copié de la nature. Internet évolue seul et selon ses lois propres, du fait de millions d’initiatives individuelles non planifiées. Il s’agit là encore d’une convergence de solutions permettant de résoudre des problèmes fonctionnels identiques à partir de substrats très différents.

4 : Il s’agit d’une question difficile à résoudre par des raisonnements linéaires : comment un cerveau qui constitue un système fermé peut-il générer, par la conscience ou autrement, des descriptions de lui-même nécessitant de sortir du cerveau pour le décrire de l’extérieur ? La cosmologie se heurte à la même difficulté. Comment sortir de l’univers pour en donner des descriptions objectives alors que nous sommes enfermés sans recours dans cet univers. La question se résout si l’on admet que les descriptions ne sont pas à ambition réaliste (voir chapitre 1) mais constituent des constructions relatives à l’objet décrit, à l’observateur/agent causal et à ses instruments.

5 : Le mot « dégénérescence » est évidemment un faux ami. Ici, il est employé dans une acception inspirée aux mathématiques. Ainsi une courbe qui « dégénère » est une courbe qui se décompose en courbes distinctes plus simples, convenant à des besoins de description particuliers.

6 : Gilbert Chauvet, dans son ouvrage, a traité de l'intelligence collective (Chauvet, Comprendre l'organisation du vivant, p. 111). Celle-ci met en œuvre une suite d'opérations qui suppose la mémoire, puisqu'elle permet d'apprendre en comparant le présent aux expériences passées par une mise en relation temporelle. L'auteur met en évidence quatre grandes étapes : la mémorisation préalable de faits à partir d'expériences appartenant à un ensemble d'évènements – la stimulation cognitive par un nouvel évènement E – la mise en situation de cet évènement E par rapport au contexte des évènements appris mémorisés – la mise en relation (explication) avec d'autres contextes, qui crée un nouvel évènement dans la contrainte d'environnement de E, appartenant à une autre série d'évènements. A partir de cette quatrième étape, il y a satisfaction ou non, autrement dit l'explication donnée apparaît ou non satisfaisante. L'américain R.C. Schank, par des voies différentes, a obtenu des conclusions similaires (Goal-based scenarios, dans Leake Case-based reasoning : experiences, lessons and future directions. MIT Press, 1996).

7 : Nous reviendrons dans le chapitre 5 sur la question de la construction du Je, ignoré par d'autres sociétés, notamment extrême-orientales. Lire à ce sujet The Rise and Fall of Soul and Self, de R. Martin et J. Barrest, Columbia University Press. 2006.

8 : De nombreux neuroscientifiques actuels persistent à tenter de démontrer l'irréductibilité de la question du Je, c'est-à-dire du Hard problem. Citons Christof Koch, "The Quest for Consciousness", Roberts and Co ou Jeffrey Gray, "Consciousness, creeping up on the Hard problem", Oxford University Press. Mais tous finissent par nier le fait que le Hard problem se pose vraiment. Jeffrey Gray envisage même l'hypothèse selon laquelle le monde perçu par la conscience ne soit pas le monde réel.

9 : C'est ainsi que ceux qui se demandent si l'Homo Neanderthalensis espèce humaine antérieure (ou parallèle) à l'espèce Homo et ayant été sa contemporaine, disposait de langage, on répond par l'affirmative. Les Néanderthaliens en effet maîtrisaient le feu et des techniques complexes de fabrication d'outils, qui n'auraient pas été concevables sans l'existence d'une culture basée sur des échanges langagiers. Sur ce point, on lira le récent et très intéressant ouvrage de la préhistorienne Marylène Patou-Mathis, Néanderthal, une autre humanité, Perrin, 2006. Voir aussi le numéro spécial de La Recherche, Néandertal, août-octobre

2006. Sur le Néanderthal ou Néandertal, l'orthographe diffère selon les auteurs. Nous ne pouvons malheureusement demander leur avis aux intéressés.

10 : Ainsi, il est soit naïf soit volontairement tendancieux d'imputer aux actualités télévisuelles souvent violentes le développement de la violence dans les sociétés urbaines. Chaque cerveau interprète et réutilise des images identiques d'une façon qui lui est propre. On peut cependant proposer des relations statistiques générales grossièrement exactes entre violences symboliques et violences vécues. La question réfère à la mémétique.

11 : L'abduction (selon Pierce) consiste à passer de déductions puis d'inductions locales à l'élaboration de théories plus générales, de type paradigmatique. On considère généralement qu'il s'agit d'une aptitude dont seul l'homme serait capable.

## Chapitre 3, section 4 :

### Modèles du cerveau. Le cerveau cortical associatif

L'ingénieur informaticien Jeff Hawkins, dans un ouvrage récent intitulé *Intelligence* (1), propose une vision un peu différente, mais en fait complémentaire, de la façon dont fonctionne le cerveau humain. Il montre que celui-ci n'est pas un ordinateur (ce dont tout le monde aujourd'hui est convaincu) mais un vaste système de mémoires qui engrange les expériences accumulées par chaque personne depuis sa naissance voire depuis le stade foetal. Ces expériences accumulées reflètent la structure du monde tel qu'il est apparu au sujet possesseur du cerveau tout au long de sa vie, sous la forme de séquences d'événements et des relations entre ces séquences. A partir de ces contenus de mémoire, le cerveau fait à tout instant des prédictions qui sont confrontées aux nouvelles expériences et mémorisées à leur tour après modifications éventuelles. C'est ce système de mémoire-prédiction qui constitue l'essentiel de l'intelligence humaine, en organisant les perceptions, la créativité et même la conscience.

Le siège de ces propriétés se trouverait dans le néocortex (2). Grâce à sa structure hiérarchique, le néocortex reçoit en parallèle les messages envoyés par les organes sensoriels et par les couches plus profondes du cerveau, et les organise sous forme de «patterns». Par pattern, il faut entendre des ensembles d'informations présentant une cohérence au plan géographique et au plan temporel (séquences chronologiques) (On pourrait traduire ce mot par celui de « blocs ou cadres organisés de représentations », mais nous préférons garder le terme anglais bien plus simple.). Ces patterns correspondent, après diverses opérations destinées à éliminer l'accessoire pour garder le permanent, aux représentations que nous nous faisons du monde (3). Le point essentiel, souvent négligé par les modèles courants du cerveau, est que le néocortex ne fonctionne pas seulement du bas vers le haut. Les patterns se forment dans chacune des six couches de neurones et sont restitués à la couche inférieure en même temps qu'ils sont adressés à la couche supérieure. Ce rétro-feedback prend la forme d'une prédiction renvoyée à la couche inférieure et instantanément comparée et modifiée si besoin est compte tenu des nouvelles informations provenant de l'extérieur. Dans ce modèle d'architecture, on pourrait dire que chaque couche de neurones du néocortex se comporte comme un néocortex à elle seule, à la différence qu'intégrée dans un système hiérarchique, ce qu'elle reçoit et émet influence en parallèle ce que reçoivent et émettent les autres couches.

C'est cette architecture qui permet notamment de ne pas confondre le cerveau avec un ordinateur classique. Contrairement à celui-ci qui produit une

information en sortie à chaque entrée d'information venant de l'extérieur, le cerveau fait, à diverses échelles de complexité, des prédictions basées sur les expériences et les séquences préalablement enregistrées. Ces prédictions provoquent des sorties motrices qui modifient le monde environnant et génèrent en retour de nouvelles informations au niveau des entrées sensorielles. L'interaction avec l'extérieur résulte du fait que le système global est doté d'organes sensoriels et d'organes effecteurs. Mais ces organes ne fonctionnent pas isolément. Ils s'auto-influencent à tous les niveaux de complexité du cortex, du fait des nombreuses liaisons synaptiques qui les relient.

Deux autres caractères différencient le néocortex des autres fibres et structures neuronales. L'une est l'organisation en colonnes verticales de quelques millimètres de diamètres qui mettent les neurones de chaque couche, à l'intérieur de ces colonnes, en communication avec ceux situés au dessus et au dessous. Ces colonnes (ou groupes de colonnes, car rien n'est unique dans le cerveau) sont en principe dédiées, du fait d'une architecture acquise par l'évolution et commandée à la naissance par les gènes de structure, à des types spécifiques de messages, par exemple transmettre via le cortex visuel, l'information relative à la perception d'une ligne verticale. Mais si nécessaire, en cas d'accident, elles peuvent se remplacer l'une l'autre. Le cerveau n'est donc pas au départ un amas indifférencié de neurones, un «tas de nouilles», selon l'expression, traduite de l'anglais, consacrée.

Le deuxième caractère est opposé au précédent, ou plutôt complémentaire de celui-ci. Dans toutes les couches, avec une densité de plus en plus grande lorsqu'on s'élève dans la hiérarchie de ces couches, il existe des fibres de liaison qui, grâce aux synapses, permettent d'associer les mémoires et donc les prédictions formulées à chaque niveau hiérarchique. Ce sont ces fibres horizontales qui ont donné d'ailleurs au cortex son qualificatif d'associatif. L'existence de liaisons horizontales était connue depuis longtemps, mais Hawkins montre qu'elles jouent un rôle permanent dans la modulation tant des informations émises en sortie que des informations reçues en entrée, ceci quelle que soit la complexité des patterns et des séquences transitant au sein des couches et à travers elles.

Un autre point important, qui permet au cerveau, malgré la lenteur de ses composants primaires, de réagir vite et de façon régulière, est l'invariance des patterns stockés à chaque niveau du cortex. Il s'agit d'une invariance relative, puisque les séquences mémorisées peuvent être modifiées si elles sont contredites par de nouvelles expériences. Mais lorsque ce n'est pas le cas, elles peuvent être immédiatement mobilisées pour produire des prédictions et entraîner des actions s'appuyant sur elles. On savait depuis longtemps, en ce sens, que le cerveau commande de nombreux comportements sur le mode automatique, par exemple dans le cas de la conduite automobile, l'appel à des solutions plus complexes ne survenant qu'en cas de difficulté inattendue.

Cette invariance des patterns ne doit pas être confondue avec les boucles sensori-motrices automatiques caractérisant l'ensemble de la vie organique et ne faisant pas intervenir le cortex, même lorsqu'elles transitent par le cerveau. Mais le point sur lequel Hawkins insiste est que ce phénomène de l'invariance des patterns au niveau du néocortex est absolument général. C'est pourquoi il caractérise le cortex comme un ensemble de mémoires prédictives. Dans chacune des couches et des colonnes, le néocortex stocke des séquences de patterns. Il s'agit d'un stockage auto-associatif, tel élément de séquence pouvant suffire à faire apparaître la séquence entière ou des séquences différentes dans lesquelles il figure. Là encore, c'est l'extrême connectivité synaptique du cerveau qui rend possible ces associations.

Enfin, les patterns sont mémorisés sous une forme invariante et hiérarchique. C'est ce caractère qui assure la permanence (relative) de la mémoire et le fait que les représentations primaires que nous nous donnons du monde s'articulent dans notre esprit en représentations de plus en plus complexes. Plus on s'élève dans la hiérarchie, plus les détails, importants dans les niveaux inférieurs, s'atténuent au profit des lignes générales. On voit ainsi apparaître, au sommet des couches néocorticales, des représentations correspondant à ce que l'on appellera en linguistique des concepts ou des noms. Les concepts sont seulement des abstractions épurées des détails. Ils n'ont pas besoin d'être nommés par le langage social pour exister et servir à orienter le comportement intelligent supérieur. Au plus haut de la pyramide, c'est le concept de « moi » qui synthétisera l'ensemble des expériences passées et actuelles enregistrées par le sujet. Mais de nouveau, on rappellera que la permanence et la hiérarchie ne sont que relatives. Elles peuvent laisser place à des variantes de représentations ou de hiérarchies si de nouvelles expériences imposent ces changements et si la plasticité d'ensemble du système permet d'en tenir compte pour assurer la réadaptation du système à un milieu profondément changé.

Comment s'élaborent les abstractions, au fur et à mesure que les patterns s'élèvent dans la hiérarchie corticale ? Hawkins évoque des processus statistiques (hebbiens), le niveau supérieur ne mémorisant que les données les plus fréquentes et ne tenant pas compte des données plus occasionnelles. Ceci ne paraît pas poser de difficulté conceptuelle, sauf qu'il faudra pour bien faire préciser dans chaque cas le processus neuronal à l'oeuvre dans la fabrication de l'abstraction. Ce pourra être le renforcement des liaisons synaptiques les plus souvent ou les plus fortement sollicitées.

Finalement, on pourra se représenter l'organisation du cortex comme un modèle du monde acquis par le sujet au fur de son développement et tout au long de sa vie. Par organisation, on entendra les patterns et associations entre patterns qui ont été mémorisés dans les couches et colonnes corticales sous forme d'associations de plus en plus stables et hiérarchisées au fur et à mesure que l'on



s'élève vers les couches corticales supérieures. Il s'agit de représentations du monde et du sujet lui-même qui sont propres à ce dernier. Un autre sujet, disposant d'un cerveau au départ identique (supposition en fait absurde, car en biologie aucun organe même simple n'est jamais identique à un autre) mais ayant vécu une histoire différente, disposerait de représentations différentes.

On n'oubliera pas cependant que l'architecture de base du cerveau est transmise par les gènes. Chaque individu en hérite sans rien pouvoir y changer. Elle correspond à un modèle du monde acquis non plus par l'individu mais par l'espèce au cours de son évolution darwinienne. Les cortex des dauphins ne sont pas analogues à ceux des humains car ces espèces vivent dans des mondes différents. Nous dirions presque, en nous appuyant sur l'hypothèse constructiviste, qu'elles ont construit des mondes différents. Pour prendre une comparaison informatique, on dira que l'individu reçoit dès le stade foetal, comme tous les membres de son espèce, une machine à traiter l'information –qui n'est pas un ordinateur, rappelons le, ou du moins pas un ordinateur numérique tel que nous le connaissons - et un système d'exploitation vides. Il les organisera en systèmes d'informations tout au long de sa vie. Ces contenus lui seront propres. Ils matérialiseront le monde particulier dans lequel vit le sujet. Mais il s'agit d'un modèle du monde qui est aussi un modèle du Je du sujet, ou plus exactement qui constitue très exactement le Je du sujet (Comme indiqué plus haut, nous reviendrons au chapitre 5 sur la question de la construction du Je). On retrouve là le "rêve" des nouvelles générations de moteurs de recherche sur Internet: représenter extensivement une personne par la collection de tous les messages émis et reçus par lui pendant une période déterminée.

### **Comparons le cortex avec une division d'infanterie**

L'architecture neuronale décrite par Hawkins rappelle, en infiniment plus complexe, celle des réseaux de neurones formels ou réseaux neuronaux artificiels (Voir ci-dessus, section 1, note ...). Mais l'auteur réfute cette comparaison, tout au moins dans l'acception donnée aujourd'hui à de tels réseaux par les informaticiens. Les réseaux de neurones formels produisent des patterns mais ils n'ont que peu de couches et fonctionnent essentiellement dans un sens, soit en entrée, soit en sortie. De plus ils ne sont pas conçus pour enregistrer des séquences temporelles. Une comparaison plus parlante, bien qu'éloignée de la neurologie, est proposée par l'auteur. C'est celle que l'on peut tirer de l'architecture et du fonctionnement d'une entreprise hiérarchique ou mieux, d'une unité militaire, par exemple une division d'infanterie, manoeuvrant sur le terrain.

On peut regretter que Hawkins n'ait abordé cette comparaison qu'en quelques lignes, au lieu de la développer. D'une part, elle peut rendre plus parlantes les descriptions de l'architecture neuronale qu'il nous propose. D'autre part, la similitude entre le cortex et les images de structures sociales organisées que l'on

peut en donner conduisent inévitablement à des réflexions d'ordre épistémologiques que l'auteur n'a pas faites. .

Développons la comparaison du cortex avec une unité militaire manoeuvrant sur le champ de bataille. Chaque soldat individuel peut être considéré comme un neurone. Il dispose de nombreux organes sensoriels spécialisés qui lui permettent de collecter des informations sur ce qui se passe autour de lui et, dans un premier temps, de les traiter à son niveau. Ainsi il confirmera une image visuelle par la recherche d'un son ou d'une odeur. S'avançant vers l'ennemi, il fait constamment des prédictions sur ce qu'il va trouver, en s'appuyant sur son expérience antérieure. Il confronte ces prédictions à ce qu'il constate. S'il observe des phénomènes correspondants à des séquences classées dans sa mémoire comme non significatives (par exemple l'envol d'un oiseau qu'il dérange), il ne leur donne pas suite. S'il remarque un événement important, auquel il s'attendait, comme la vue d'une mine isolée, mais s'il peut traiter lui-même cet événement, par exemple en neutralisant la mine, il n'alertera pas, en principe, l'échelon hiérarchique supérieur. Il pourra par contre prévenir ses voisins immédiats en utilisant les moyens de communications à courte portée dont il dispose. Si enfin il observe un fait auquel il s'attendait également mais auquel il ne pourra pas faire face seul, comme l'apparition d'un char d'assaut, il alertera aussitôt son chef de section. Il fera de même, a fortiori, s'il note un événement grave auquel il ne s'attendait pas, par exemple une incapacité physique subite pouvant laisser penser qu'il est soumis à une agression chimique.

Le chef de section, à son niveau, fait de même. Tant que le soldat en avant-garde ne signale rien, il ne prend pas de mesures particulières. Mais il s'attend, en fait, à ce que son groupe rencontre une résistance ennemie. Dès que le compte-rendu du soldat lui parvient, il peut préciser l'image théorique de l'ennemi qu'il avait formée dans sa mémoire suite à des expériences précédentes et donner des ordres en conséquence. Ces ordres impliquent l'ensemble des hommes de la section et circulent entre eux immédiatement, de façon horizontale, grâce aux moyens de communication propres à cette dernière. Si la section, en tant qu'unité, peut traiter l'événement, le chef de section n'alertera pas le capitaine. Il le fera au contraire si l'occurrence se révèle dépasser les dimensions de l'événement qu'il avait mémorisées et auquel il était préparé à faire face.

Le processus se répète à tous les niveaux de la hiérarchie, jusqu'au général commandant la division. Celui-ci avait prédit le fait, parmi d'autres occurrences possibles déjà mémorisées elles aussi, que sa division rencontrerait une forte résistance. Si les différents niveaux inférieurs lui signalent qu'une résistance effective est observée, il prend un ordre global en conséquence. Cet ordre pourra être de stopper la progression et de s'abriter. L'ordre redescend tous les échelons et est traité, de façon différente, en fonction du terrain, c'est-à-dire des retours d'expérience, par chaque niveau. Le fantassin de base l'exécute à son tour et rend

compte, en signalant au besoin les événements nouveaux pouvant rendre l'ordre inopportun.

On voit que dans une telle architecture, chaque niveau est en principe spécialisé. Mais la spécialisation peut s'adapter aux retours d'expérience. Cette possibilité d'adaptation résulte du fait que chaque niveau émet à la fois vers le haut (rendre compte, voire prendre des initiatives) et vers le bas (transmettre les ordres). Elle résulte aussi du fait qu'il existe des communications horizontales entre niveaux permettant aux unités composant ceux-ci de coopérer entre elles et se remplacer si nécessaire. Ces communications sont assurées par des liaisons radio ou autres spécialisées à chaque niveau. Si les soldats étaient dotés d'émetteurs Internet personnels, ils pourraient en principe, comme dans l'Internet civil, correspondre avec tous les autres abonnés au réseau, quel que soit leur niveau hiérarchique. Mais ils ne le feraient que pour des raisons exceptionnelles, signalant l'émergence d'occurrences profondément inattendues (En bonne règle, on ne dérange pas plus le général pour signaler un trou de mulot qu'on n'envoie un courrier électronique au premier ministre pour demander la pose d'un sens unique dans sa rue.).

Si nous poursuivons la comparaison entre le cerveau et une unité militaire opérant sur le champ de bataille, que pouvons-nous ajouter ? La constatation essentielle qui s'impose est la suivante : l'unité, à chaque instant, constitue une mémoire d'ensemble représentant le champ de bataille tel qu'il est à cet instant. D'autres points de vue sur le champ de bataille pourraient être obtenus par d'autres observateurs situés à l'extérieur, mais ces points de vue ne seraient pas plus «vrais» que ceux obtenus par les différents éléments de l'unité, depuis le fantassin à la base jusqu'au commandant de division au sommet, qui interagissent en temps réel avec les événements du monde survenant dans le cours de la progression de l'unité sur le terrain, au contact de l'ennemi.

Ceci ne veut pas dire que l'unité obtient une description objective du champ de bataille, c'est-à-dire du monde en soi hypothétique qui correspond à ce terme. Les patterns ou représentations stables obtenues par l'unité ne sont pas imaginaires, puisqu'elles sont testées en permanence par les différents capteurs sensoriels et actuateurs physiques dont dispose l'unité. Mais ils restent relatifs à ces capteurs et actuateurs, c'est-à-dire finalement à l'observateur et à ses instruments observant le monde externe. Ils n'ont donc qu'une vérité relative, mais ceci est suffisant pour permettre à l'unité de naviguer dans le champ de bataille en optimisant ses chances de survie grâce aux échanges permanents entre ces divers composants.

Une autre question intéresse le niveau de connaissance du champ de bataille, c'est-à-dire du monde extérieur, que peuvent obtenir les différents échelons hiérarchiques. Le soldat individuel ne pourra produire qu'une connaissance locale, celle qu'il obtient personnellement ou en communiquant avec ses voisins immédiats. Mais chaque niveau hiérarchique supérieur produira une connaissance plus étendue, dont la combinaison au sommet de la division aboutira à une

connaissance globale - aussi globale que possible, tout au moins, compte tenu des moyens mis en oeuvre. Dans le schéma du cerveau proposé par Hawkins, les patterns ou représentations dotés d'une certaine permanence et produits aux différents niveaux hiérarchiques du cortex servent de base aux prédictions qui déterminent les actions et qui sont elles-mêmes soumises à la sanction de l'expérience découlant de la mise en oeuvre de ces actions.

Il en sera de même dans notre exemple. Au reçu des informations provenant des niveaux subordonnés, le commandant de la division donnera un ordre global qui sera exécuté par chaque échelon subordonné pour ce qui le concerne. Mais celui-ci n'obéira pas en aveugle, si l'on peut dire. Il adaptera l'exécution de l'ordre compte tenu des représentations locales du milieu dont il est porteur, et que le commandant de division ne pouvait pas connaître. Tout problème non prévu par le commandement et pouvant remettre en question l'exécution de l'ordre remontera immédiatement vers les échelons supérieurs, si son importance le mérite. Ainsi la division se comportera comme un ensemble d'agents auto-adaptatifs capables de prendre en compte les éléments les plus fins de la réalité du monde. Autrement dit, elle aura un comportement intelligent. Les conseils en organisation civils et militaires ont compris cela depuis longtemps, ce qui a condamné les structures hiérarchiques à la prussienne, où le soldat ne doit être qu'un exécutant aveugle.

### **Une architecture homogène**

L'observation selon laquelle, malgré les différents rôles fonctionnels des aires du néocortex, celui-ci présenterait partout la même structure en 6 couches neuronales elles mêmes organisées en colonnes verticales mérite un commentaire. En découle l'hypothèse que les différentes parties du néocortex fonctionnent de la même façon, selon un algorithme de base unique. Les différences entre les fonctions remplies par chacune d'elles tiendraient non à des différences de structure mais à la façon dont elles sont reliées aux autres parties du cerveau. Et de quel type d'algorithme s'agit-il ? Produire à partir des scènes et séquences préalablement mémorisées des prédictions sur ce que l'action en cours va faire apparaître, afin de mettre en évidence les différences entre ce qui était prévu et ce qui se produit réellement. Une fois ces différences reconnues, de nouvelles mémorisations les intégrant remplacent les précédentes, ce qui permet au cerveau et plus exactement au néocortex d'optimiser en permanence son adaptation aux changements du monde. Or à la base des mécanismes permettant ces prédictions, ne retrouvent-t-on pas le Principe d'Auto-Association Stabilisé (PASS) décrit par Gilbert Chauvet. Nous reviendrons sur ce point dans la dernière section.

La plupart de ces processus de prédiction-mémorisation sont inconscients car le néocortex traite l'information à la milliseconde et ce sont le plus souvent des

détails infimes qui font l'objet du mécanisme d'adaptation. De plus, et c'est un point essentiel intéressant la question de la conscience de soi, le cerveau n'est pas organisé pour s'auto-informer de la façon dont ses différents composants fonctionnent - ou cessent de fonctionner (alors que, dans de nombreux cas, il peut être informé de la façon dont fonctionnent un certain nombre de composants du corps, organes sensoriels et moteurs notamment. Ainsi lorsque les sens font apparaître des éléments discordants, par exemple un oeil manquant dans un visage, nous éprouvons inconsciemment un choc qui nous oblige à rechercher la cause du conflit et à lever l'incertitude : est-ce le système visuel ou l'objet perçu qui comporte le défaut ?

Les neurologues modernes ont confirmé que plus de la moitié des échanges d'information entre les couches supérieures du cortex, siège des fonctions nobles, et les couches inférieures en relation avec les entrées-sorties endogènes et exogènes sont dans le sens descendant et non, comme on le croyait jusque là dans le sens ascendant (les sens informant la conscience). Pour illustrer le rôle de ces communications descendantes, Hawkins prend l'exemple de l'audition. Pourquoi parvenons à comprendre un interlocuteur dans une foule dont le bruit de fond dévore la plupart des paroles. Si notre compréhension se construisait uniquement de bas en haut, à partir des messages brouillés que nous recevons, nous ne comprendrions pas grand-chose. En réalité, notre cortex supérieur anticipe, compte tenu de la connaissance que nous avons de notre interlocuteur, ce que sera l'essentiel de son discours. L'information est communiquée au système auditif, de haut en bas. Ce dernier n'aura plus qu'à enregistrer, même à partir de réceptions brouillées, les quelques différences qui permettront de restituer le sens complet du nouveau discours.

## **Commentaires**

Il est intéressant de voir grâce à ce livre une nouvelle illustration possible de la façon dont les entités biologiques, à commencer par les plus simples, ont progressivement élaboré sous la pression de sélection darwinienne des représentations ou modèles symboliques du monde. Ces représentations leur ont permis de créer des réponses de plus en plus complexes aux contraintes du milieu et donc d'envahir celui-ci à une vitesse de plus en plus accélérée. Cette évolution a produit l'homme doté non seulement d'un cerveau mais d'un néocortex. Elle n'est sans doute pas près de s'arrêter, d'une part parce que les cerveaux humains ne sont que des unités au sein d'une société de milliards d'équivalents reliés par les langages et les voies de communication des sociétés et cultures modernes. Il n'est d'ailleurs pas impossible que sous la pression de la sélection culturelle, l'héritage génétique commandant l'organisation du cerveau humain, dès le stade embryonnaire, puisse encore évoluer. Mais d'autre part aussi, comme le suggère fort opportunément

l'auteur, parce que la société humaine est sur le point de créer des cerveaux artificiels qui augmenteront de façon sans doute illimitée les capacités des cerveaux biologiques. Il s'agit de phénomènes qui n'intéressent pas seulement l'avenir de l'homme mais celui de l'intelligence consciente au sein de l'univers global.

Cependant le livre de Hawkins ne se borne pas à répéter ce qui n'est plus pour nous aujourd'hui qu'une constatation triviale. Il essaye de montrer avec précision, en descendant au niveau des neurones individuels, comment l'architecture du cerveau, sous la forme récemment apparue du néocortex, permet de générer, dès le plus jeune âge, les bases de l'intelligence culminant dans la représentation de soi dans le monde et dans l'élaboration de stratégies très larges de survie individuelle et collective. Certes, comme l'auteur le rappelle avec justesse, ces bases existent chez tous les animaux, même dotés de simples amorces de systèmes nerveux centraux. L'intelligence n'est pas le monopole de l'homme. Elle prend des formes très différentes et très subtiles dans l'ensemble des espèces. Mais c'est le néocortex humain qui en permet le déploiement le plus élaboré, du moins tel que nous pouvons en juger à ce jour.

On lira donc avec intérêt les descriptions anatomiques et fonctionnelles du néocortex proposées par Hawkins, lequel bien que non neurologue, s'est appuyé sur les constations les plus récentes des neurosciences (4). Mais il nous semble cependant que l'enthousiasme de l'auteur vis-à-vis de la pertinence de ses propres hypothèses mérite un certain recul. On ne peut pas ne pas poser des questions qui ne contrediront certainement pas ces hypothèses, mais qui devraient pouvoir au contraire les enrichir et les prolonger. L'auteur d'ailleurs nous invite à le faire, sur le site qu'il a consacré à son livre et qui est référencé au début de cet article.

### **Que pourraient être ces questions ?**

La première concerne la validité des descriptions du cortex présentées dans le livre, tant au plan de l'anatomie que du fonctionnement. Sans les contester radicalement, ce dont d'ailleurs nous n'aurions pas la compétence, on se bornera à remarquer qu'il ne s'agit encore que d'hypothèses, restant à vérifier par des études fines menées in vivo. On constate par exemple qu'à toute difficulté intéressant par exemple les capacités de la mémoire et des associations, l'auteur fait appel à la richesse quasi infinie du cerveau en neurones et plus encore en connexions synaptiques. Mais ceci ne suffit pas à montrer comment les associations s'organisent effectivement, certaines pouvant subsister indéfiniment et d'autres disparaître. Ce n'est pas parce que notre galaxie comporte cent milliards d'étoiles, autant que de neurones dans le cerveau, qu'elle est intelligente. Or de telles études sont encore très difficiles, compte tenu du manque de précision des méthodes actuelles d'imageries conduites sur des personnes vivantes en action. L'auteur est

conscient de ce problème, il suggère d'ailleurs des expériences qui pourraient être menées dans un proche avenir pour vérifier ses hypothèses. Mais c'est plus généralement sur la possibilité pour un observateur confronté à un observé susceptible de multiples interprétations qu'il faut s'interroger. On se trouve un peu dans la situation du physicien quantique face au monde microscopique. La conscience de l'observateur crée dans une certaine mesure l'observé. Tout au moins, elle le qualifie d'une certaine façon, qui est fonction des instruments dont il dispose sur le moment. Ce n'est donc pas finalement le monde quantique en soi qui est observé et décrit, mais la construction résultant de l'interaction de l'observateur/acteur humain et de ses instruments avec un continuum quantique non descriptible. Nous renvoyons sur ce point aux développements du chapitre 1 de ce livre.

Que conclure de cette première remarque ? C'est que la description des mécanismes du cerveau, de l'intelligence et de la conscience que nous propose Jeff Hawkins est nécessairement le reflet de la façon dont il voit le monde en tant qu'observateur/acteur situé dans un temps et dans un espace bien définis. Ceci ne veut pas dire qu'il faudrait les rejeter, mais les replacer dans la démarche constructiviste de l'ensemble des sciences occidentales contemporaines. Il ne décrit certainement pas ni l'intelligence ni le néocortex en soi, tâche impossible voire sans objet, mais quelque chose en train de se construire à son insu et dont il est un agent particulièrement bien informé. Que voulons-nous dire ? Les entités conscientes (dont nous faisons partie) sont sans doute engagées dans un processus général d'extension au sein du cosmos. Tout au moins c'est ainsi qu'elles se conçoivent. Se concevant ainsi, elles font en sorte, sans le vouloir et donc inconsciemment, de transformer le monde naturel au profit du type d'organisation qu'elles représentent. Hawkins, comme nous qui le lisons, sommes donc en train de nous construire de façon à augmenter en nous les capacités de traitement intelligent et conscient telles que, par émergence, nous avons appris à nous les représenter. Nous projetons donc dans ce que nous croyons observer du monde des processus visant à assurer cette augmentation, processus qui eux-mêmes résultent d'une longue évolution biologique dont nous sommes les produits. Les observations du cortex que formulent Hawkins sont à cet égard des prédictions visant à être prescriptives de ce que devrait être le cortex pour répondre à nos ambitions d'êtres en marche vers plus d'intelligence. En d'autres termes, à travers les yeux d'Hawkins, nous ne voyons pas le cortex tel qu'il est (concept d'ailleurs non acceptable aux yeux du non-réalisme) mais tel qu'il devrait être pour répondre à nos exigences actuelles en matière de système intelligent.

On peut avoir un soupçon de cela en constatant que Hawkins, ne sachant pas dans le détail comment se font les échanges d'informations au sein des couches corticales, se comporte comme s'il était le responsable d'une organisation sociale (une division d'infanterie pour reprendre l'exemple ci-dessus) dont il aurait mission d'améliorer l'adaptabilité. Il disposerait pour ce faire d'une grande quantité de

ressources, les neurones ou soldats, les liaisons synaptiques ou radio entre ceux-ci, et il devrait organiser au mieux l'émergence d'une entité intelligente globale, le sujet conscient d'une part, la division d'infanterie d'autre part. A cette fin, il imagine les meilleures solutions possibles, en l'état actuel de ses connaissances, pour améliorer le fonctionnement du cortex. Nous ne voulons pas dire que ces solutions ne correspondent pas à ce qui se passe dans le cortex ou ne sont pas de bonnes solutions en elles-mêmes. Mais le lecteur doit toujours se demander si le cortex, et plus généralement le cerveau, et plus généralement encore le corps humain tout entier immergé dans les relations sociales, ne font pas appel à d'autres solutions pour générer de l'intelligence et de la conscience. En d'autres termes, pour parler simplement, ne prenons pas les perspectives ouvertes par Jeff Hawkins au pied de la lettre, ce qui risquerait d'empêcher de voir des solutions profondément différentes susceptibles, elles aussi, d'explicitier les faits de conscience et notre propre conscience...à plus forte raison des types d'intelligences non humaines. A fortiori, ne nous appuyons pas exclusivement sur les schémas de Hawkins pour réaliser des automates intelligents.

Nous pouvons faire d'autres observations à l'auteur, dont certaines ont d'ailleurs été esquissées dans le forum consacré au livre. En voici quelques-unes dans le désordre :

- La question de la conscience de soi, du Je tel que nous le percevons, n'est pas véritablement traitée. L'auteur s'en tire par une pirouette. Dans l'architecture proposée, on en voit pas comment s'organise ce que Bernard Baars appelle l'espace de travail conscient. Même si celui-ci n'est pas localisé de façon permanente dans le cortex, il doit bien correspondre à des échanges neuronaux spécifiques, qui ne se produisent pas dans les traitements inconscients, fussent-ils «intelligents». Comment dans ce cas distingue-t-on le conscient de l'inconscient et du préconscient. Quel apport spécifique au système décrit par l'auteur représente l'intervention du moi conscient et de ce que l'on appelle encore le "libre-arbitre". S'agit-il d'illusions ou non?

- De la même façon, l'auteur évacue bien trop cavalièrement la question des sentiments et affects. Pour la plupart des spécialistes de la conscience, notamment Damasio et Edelman, étudiés dans les sections précédentes, les sentiments constituent un élément essentiel de la construction des architectures neuronales. Ils modulent souvent de façon très profonde les inputs ou entrées responsables des connexions synaptiques, venant soit de l'extérieur soit de l'intérieur du corps. Ils modulent également les expressions émises en sortie. Aussi bien les concepteurs de conscience artificielle leur font jouer un rôle essentiel dans la réponse à une question fondamentale : qu'est-ce qui peut faire penser une machine consciente ? (Voir chapitre 4.)



- Pourquoi ne pas évoquer le rôle des mèmes (Voir chapitre 5) dans la construction de la capacité associative du cortex ? Les mèmes de type langagier, selon les méméticiens, représentent une bonne explication à la croissance extraordinaire du cortex associatif observée depuis 1 ou 2 millions d'années chez l'homme, seule créature capable d'échanges sociaux systématiques à base de messages symboliques. La colonisation du cortex et de ses ressources en connexions synaptiques par des mèmes en compétition darwinienne pourrait aussi expliquer pourquoi les connexions internes au cortex se sont produites si efficacement, en apparence, bien qu'aucun ingénieur n'ait prévu à l'avance le moindre schémas de réseaux. Ne nous trouvons nous pas en présence de phénomènes d'auto-organisation entraînant leur propre optimisation, analogues d'ailleurs - et ce ne serait pas une simple coïncidence - à ce qui se passe au sein du réseau Internet. Lorsqu'en quelques secondes l'internaute trouve en utilisant un moteur de recherche puissant l'adresse d'un nom précis, fut-il présent en un seul exemplaire au sein du web mondial, il est aussi surpris que lorsque il constate que son cerveau peut se remémorer subitement un fait de son passé oublié depuis longtemps et qui était pourtant mémorisé quelque part dans les cent milliards de neurones dudit cerveau.

1 : Jeff Hawkins (inventeur du Palm Pilot). n'est pas un neurobiologiste mais un informaticien s'intéressant aux neurosciences. Son livre, écrit conjointement avec Sandra Blakeslee, journaliste scientifique, a été traduit en français sous le titre « Intelligence », Campus Press. 2005. Nous nous sommes demandé s'il n'était pas hasardeux, du point de vue scientifique, de consacrer une section entière à Jeff Hawkins, considéré par beaucoup de neuro-scientifiques comme un amateur. Mais il nous a semblé que les « images » qu'il propose du fonctionnement du cortex (images qu'il n'a d'ailleurs pas inventées seul) sont éclairantes pour nous, dans un domaine encore très obscur, où les études spécialisées sont ininterprétables par un profane.

2 : Le néocortex est fait de six couches empilées de neurones aux fonctions spécifiques, sur une épaisseur qui ne dépasse pas 2 mm, mais qui recouvre, chez l'homme, l'ensemble de la surface de l'encéphale. Le néocortex ne regroupe qu'un petit nombre des 100 milliards de neurones dont dispose le cerveau et qui tous contribuent (en association avec les cellules gliales ou astrocytes, comme on l'a montré récemment), au bon fonctionnement du corps et de l'esprit.

3 : (Dans *The Wisdom Paradox: how your mind can grow stronger as your brain grows older* (Free Press 2005) le professeur de neurologie clinique à la New York University School of Medicine Elkhonon Goldberg montre que le cerveau des personnes vieillissantes conserve ou augmente sa capacité à résoudre des problèmes, ceci même si l'âge inflige d'autres pertes de cognition, notamment en termes de mémoire immédiate. Cette propriété découle du fait que la fonction dite "pattern recognition" ou aptitude à reconnaître des patterns ne fait que s'accroître

au fur et à mesure que s'accroît l'expérience des personnes âgées. La fonction est complexe et mal comprise, que ce soit chez l'homme ou dans les robots. Cependant, elle apparaît comme générique à la plupart des espèces dotées d'un encéphale et par conséquent ne subirait pas la dégradation des autres fonctions du cerveau humain. Encore faut-il l'exercer, aussi bien par des activités intellectuelles que par des activités physiques obligeant à résoudre des problèmes. Acceptons en l'augure (voir New Scientist 13 Août 2005, p. 51).

4 : A l'époque tout au moins où le livre fut écrit. Or les connaissances évoluent vite en la matière. Dans ses dernières publications (voir la bibliographie de son livre Comprendre l'organisation du vivant , Gilbert Chauvet a montré que les neurones ne sont pas les seuls à intervenir dans la mémorisation. Le rôle des astrocytes (cellules gliales) et capillaires est considérable. Le couplage entre les trois réseaux neurones, astrocytes et capillaires (sans oublier les neurohormones ) est nécessaire à une explication de la mémoire et de l'apprentissage. Il reste que, pour les non-spécialistes, des livres comme ceux de Hawkins ont l'intérêt de proposer des interprétations aisément compréhensibles. Leur danger, par contre, est de laisser croire que tout les problèmes sont résolus. Ils empêchent le lecteur de comprendre que de véritables théories (mathématiques) du fonctionnement du cerveau restent nécessaires à une compréhension en profondeur. Voir à ce sujet la section 5.

## Chapitre 3, Section 5 :

### Modèles du cerveau. Gilbert Chauvet.

#### Le PAAS appliqué au cerveau

Comment l'évolution a-t-elle pu conduire au cerveau humain ? La théorie de la vie présentée par Gilbert Chauvet (Voir Chapitre 2, section La présente section, comme la précédente, s'inspire directement du travail de Gilbert Chauvet, à qui elle a été soumise.) fondée sur la double organisation structurale et fonctionnelle des êtres vivants et de leurs inter-relations, permet-elle d'obtenir une interprétation des concepts d'apprentissage, de mémoire, d'intelligence, et en particulier de celui de conscience ? L'organisation fonctionnelle, de nature mathématique, est définie par l'auteur comme la distribution des puits selon les interactions fonctionnelles élémentaires à un niveau donné de la hiérarchie. Il s'agit de dénombrer les puits qui correspondent à une interaction fonctionnelle, en d'autres termes, de savoir dans combien de puits chaque produit émis par une source va « tomber » ! La relation structure-fonction est alors clairement établie comme la relation organisation structurale - organisation fonctionnelle.

Comment une propriété particulière peut-elle « émerger » de l'organisation fonctionnelle ? Dans le cas particulier du cerveau, le problème est de déterminer comment une telle propriété peut émerger de réseaux de neurones réels, autrement dit, comment l'intelligence du mouvement émerge du cervelet et autres organes du système sensorimoteur, la mémoire cognitive émerge de l'hippocampe, et finalement la conscience du cerveau.

#### **L'unité fonctionnelle, résultat de l'évolution**

L'évolution peut s'expliquer par une série d'auto-associations stabilisatrices à l'origine de la double hiérarchie du système, structurelle et fonctionnelle. C'est aussi le cas du système « cerveau » constitué de réseaux de neurones réels. Un réseau de neurones réels, comme tout système d'un organisme vivant, dispose d'une organisation doublement hiérarchique, à la fois structurale et fonctionnelle. Comment de cette hiérarchie peut émerger une unité fonctionnelle ? On peut très schématiquement représenter les réseaux de neurones réels par des réseaux de neurones formels. On y voit apparaître des hiérarchies mettant en évidence des propriétés « émergentes », c'est-à-dire des propriétés qui apparaissent au niveau supérieur dans une nouvelle structure. L'un des avantages de la représentation

hiérarchique est de conduire à une définition rigoureuse de ce que Gilbert Chauvet a nommé l'unité fonctionnelle, c'est-à-dire une unité structurale ayant sa propre fonction à un niveau supérieur de l'organisation.

L'unité fonctionnelle assure donc une nouvelle fonction physiologique déduite mathématiquement des niveaux inférieurs. Elle a sa propre échelle de temps. La difficulté consistera à la reconnaître dans le système extrêmement complexe qu'est le cerveau. Pour ce faire, il faudra identifier les nombreux mécanismes qui permettent à l'information électrique de passer d'un neurone à l'autre (1).

### **Les mécanismes**

Nous ne reprendrons pas ici la description de la route suivie par un influx nerveux, ou potentiel d'action (2). Ce sont les lois de la physique qui expliquent comment s'accomplit cette propagation, analogue à celle qui existerait dans un câble électrique. Mais il existe des milliards de neurones et des milliers de synapses par neurone avec une très grande possibilité de connectivité interneurones.

Les mécanismes physiologiques agissant dans les neurones n'ont qu'un but : laisser passer rapidement un potentiel (en quelques millisecondes) et moduler ce transfert en en conservant une trace. Cette seconde opération, intervient sur une échelle de temps beaucoup plus longue que la première, de l'ordre de la minute à l'heure. On voit apparaître ici la hiérarchie fonctionnelle due à ce découplage temporel des dynamiques (G. Chauvet 1993). Chaque neurone réalise une intégration des signaux qui lui arrivent en provenance de diverses autres régions et pendant une certaine fenêtre de temps. On conçoit alors le nombre énorme de possibilités d'intégration grâce à la connectivité. L'intégration est à la fois spatiale et temporelle. A cause de leur complexité, on pourrait croire à l'impossibilité de démontrer une quelconque propriété de tels réseaux. Or, depuis une vingtaine d'années, les chercheurs ont trouvé les ingrédients à l'origine de l'apprentissage et de la mémorisation par ces neurones (J.A. Anderson et E. Rosenfeld, 1990). C'est là qu'interviennent les mathématiques.

Considérons un réseau de neurones formels. Chaque neurone possède seulement deux propriétés : il somme les signaux d'entrée puis transforme cette somme ; il modifie l'entrée selon certaines règles dites « d'apprentissage » (D.O. Hebb, 1949). Ces deux seules propriétés d'un objet-neurone conduisent à un réseau de ces objets capable d'apprendre et de mémoriser. C'est une propriété d'essence mathématique, qui intègre les facteurs de transmission entre neurones à la suite d'une modification de leur structure. Ceci est appréciable pour le neurobiologiste théoricien qui obtient une propriété globale du réseau à partir de ses éléments. Un

autre avantage la propriété de généralisation de tels réseaux dits auto-adaptatifs. Si on stimule le réseau avec un signal assez proche du signal appris, on obtient la même réponse ! D'où une flexibilité jamais obtenue avec un système physique, point de départ de nouvelles technologies comme la reconnaissance de la parole ou de l'image, et la naissance de la mémoire distribuée.

Or cela se révèle vrai avec un neurone réel, pourvu de ses mécanismes physiologiques. Les deux propriétés macroscopiques du comportement de l'objet-neurone, qui sont suffisantes pour obtenir le comportement global du réseau de neurones formels, se retrouvent dans les mécanismes neuronaux réels et font de ce système un réseau réel capable d'apprendre et de mémoriser sous contraintes moléculaires ? Cependant les réseaux de neurones réels sont beaucoup plus complexes que les réseaux de neurones formels car ils respectent une architecture hiérarchique. Ils sont constitués de types cellulaires différents, avec des règles d'apprentissage résultant de toute une cinétique biochimique cellulaire, donc non imposées au système comme le sont les règles des réseaux artificiels (règles de Hebb). L'intégration de la dynamique physique à chacun des niveaux d'organisation, subcellulaire, neuronal et populationnel est ici nécessaire.

La théorie du champ neural proposée par Gilbert Chauvet, adaptée de sa théorie des champs biologiques, intègre les mécanismes physiologiques. En chaque point, groupe de neurones, neurone, synapse, complexe canal-récepteur, existe une source locale qui, parce qu'elle est aussi un puits, a reçu les effets des autres sources par l'intermédiaire du champ. Cela correspond à des effets non locaux. Puis, la source a transformé ces effets pour les "projeter" ailleurs. Dans le cerveau, l'organisation fonctionnelle des neurones intègre le temps par la mémoire, l'espace par la non-localité. Un potentiel d'action émis en un endroit donné du réseau continu en produit un autre en un autre endroit à un instant ultérieur en passant par les niveaux d'organisation inférieurs que sont les canaux et les synapses (émission des neurotransmetteurs, diffusion, liaison au récepteur, conduction dendritique, sommation). La théorie mathématique du champ (3) permet de calculer la dynamique des potentiels, donc l'activité, en tenant compte de la traversée des niveaux d'organisation structurale.

Ayant vu comment sont réalisés l'apprentissage puis la mémorisation, Gilbert Chauvet peut passer au niveau global du comportement cognitif à partir des mécanismes moléculaires, c'est-à-dire traiter de l'intelligence et même de la conscience. Pour un physiologiste théoricien comme lui, la question essentielle est de savoir si l'intelligence est une fonction physiologique. Autrement dit, si on peut l'expliquer à partir de l'anatomie et de la physiologie.

Pour ne pas s'engager dans des discussions sans fins sur ce qu'est l'intelligence, il propose d'analyser l'« intelligence du mouvement ». Le mouvement semble quelque chose de tellement commun qu'il apparaît comme inné. Mais

lorsqu'on analyse ce qu'est le mouvement, on s'aperçoit de sa complexité et du nombre considérable d'éléments qu'il met en œuvre. Ainsi, le cervelet agit pour coordonner les mouvements fins, par exemple ceux du violoniste utilisant son archet. Ce sont les observations médicales qui l'ont enseigné. Placé en parallèle sur les grandes voies motrices volontaires et automatiques, le cervelet assure les meilleures performances possibles dans les réponses motrices à l'environnement.

Pourquoi peut-on parler d'intelligence à propos du mouvement ? Quelles en sont les bases neurobiologiques ? Comment une explication plausible peut-elle être déduite des mécanismes physiologiques et des données anatomiques ? Plus directement, comment donner une explication à partir de l'intégration des mécanismes ?

Le cervelet, comme son nom l'indique, est un « petit cerveau ». Il est situé à la partie postérieure de celui-ci et est constitué d'un cortex (le cortex cérébelleux) et de noyaux (les noyaux du cervelet). Les cellules du cortex cérébelleux sont disposées de façon très régulière, elles reçoivent des signaux de la périphérie par l'intermédiaire de « senseurs », elles les transforment et les envoient vers les noyaux centraux du cervelet, ces derniers les envoyant à leur tour vers le système moteur. Le cortex cérébelleux constitue un réseau de réseaux de neurones, et s'organise, de ce fait, en un système structurellement hiérarchisé. A chaque cellule de Purkinje est associé un réseau de neurones appelés cellules des grains. A la taille imposante des premières, en particulier de leur majestueux arbre dendritique, s'oppose la petitesse des innombrables cellules des grains dont les axones qui sont les fibres parallèles courent sur plusieurs millimètres le long du cortex en « faisant synapse » avec plusieurs cellules de Purkinje.

Peut-on dans ces conditions démontrer les propriétés d'apprentissage et de mémorisation de trajectoires, et par suite, de la coordination des mouvements ? Peut-on ainsi expliquer pourquoi le cervelet permet de coordonner des mouvements fins et rapides ? Il est clair que le résultat observé, à savoir la coordination, doit résulter de l'intégration des mécanismes. Comme l'a montré Gilbert Chauvet (G. Chauvet, 1995), il faut mathématiquement découvrir l'unité fonctionnelle siège de la (ou des) propriété (s) émergente (s). Dans ce domaine, seule l'approche mathématique permet l'intégration, au sens strict de ce terme.

### **Gilbert Chauvet La coordination du mouvement par le cervelet**

Interpréter la coordination du mouvement par le cervelet fut pour moi un très long travail, car il a fallu passer du simple au compliqué, de la modélisation classique à l'introduction de la théorie générale. Pour cela, je suis parti de travaux antérieurs remarquables, comme ceux de D. Marr (Marr, 1969)) et de J.S. Albus

(Albus, 1971), pour arriver à la théorie du champ hiérarchique. Une question importante était de savoir comment l'organisation structurale et l'organisation fonctionnelle se combinent pour produire cette propriété de coordination du mouvement, ceci imposant en premier lieu d'identifier les interactions fonctionnelles et les unités structurales.

J'ai d'abord conduit une première série de travaux sur le comportement d'un modèle mathématique semi-réel du cortex cérébelleux, c'est-à-dire possédant les mécanismes physiologiques d'excitation ou d'inhibition des divers neurones et la disposition anatomique du réseau, mais où les règles d'apprentissage lui sont imposées, au lieu de provenir de la chimie du système lui-même. Ceci est une simplification importante car, comme je l'ai dit plus haut, les règles d'apprentissage sont indispensables à l'apprentissage en permettant aux synapses de moduler localement la transmission inter neuronale. Les réseaux dits artificiels ont des règles imposées qui, d'ailleurs, ont fait l'objet de recherches intensives pour conduire à des réseaux efficaces. Mais pour le biologiste, le problème est différent : il s'agit de découvrir les mécanismes physiologiques de ces règles, et donc d'identifier le réseau biochimique correspondant. Le réseau semi-réel que j'ai d'abord considéré n'incluait pas le niveau moléculaire, mais il m'a permis de mener l'étude mathématique poussée de la stabilité du réseau hiérarchique. La théorie du champ et les S-propagateurs autorisent ensuite l'intégration complète des mécanismes.

Les principaux résultats (mathématiques) du modèle semi-réel ont montré que la hiérarchie fonctionnelle crée les règles émergentes, que l'organisation structurale crée la stabilité de la fonction, et que le couplage entre unités accroît le domaine de stabilité de la fonction.

En d'autres termes, les nouvelles règles d'apprentissage que j'ai appelées règles variationnelles et qui assurent la coordination des mouvements sont déduites mathématiquement du niveau neuronal. Elles émergent au niveau supérieur du réseau des unités de Purkinje sous certaines conditions toujours satisfaites, car elles sont liées à la différence des échelles de temps entre la fonction « efficacité synaptique » et la fonction « activité ». C'est donc la hiérarchie fonctionnelle qui « crée » les règles variationnelles ; les nouvelles règles « émergent » de l'organisation temporelle.

En outre, la stabilité mathématique de la fonction d'apprentissage et de mémorisation (au sens habituel d'une perturbation qui laisse le système dans son domaine de stabilité) qui doit être réalisée à l'intérieur d'une unité entraîne une première condition entre les paramètres géométriques de l'unité. Mais elle doit être aussi assurée pour le réseau. On obtient donc une seconde condition de stabilité qui dépend aussi de paramètres géométriques, en particulier de la distance entre les unités (G. Chauvet, 1993). C'est donc finalement l'organisation structurale selon les échelles d'espace qui est à l'origine de la stabilité de la fonction.

Il est rassurant de voir que le fonctionnement stable de ce réseau compliqué repose sur la double organisation hiérarchique structurale et fonctionnelle. On peut dire que ces propriétés, parce qu'elles sont de nature mathématique, résultent de hiérarchie 3-D du système biologique. Mais il y a plus intéressant encore, à savoir que la condition de stabilité inter-unités entraîne la condition de stabilité pour l'unité. On en déduit que le couplage entre unités augmente la stabilité du système global (ce qu'exprime le PAAS). La théorie générale est par conséquent vérifiée pour le couplage entre deux unités de Purkinje. Cet exemple étant considérablement plus complexe que celui des voies biochimiques, il constitue une avancée significative pour la réalité de ce paradigme.

Mais la théorie du champ à  $n$  niveaux avec ses S-propagateurs apporte encore davantage à une interprétation de la fonction physiologique assurée par le réseau de neurones réels. D'abord, le réseau inclut le niveau moléculaire, ce qui nous permet d'analyser l'effet de molécules sur les mécanismes physiologiques connus. Ensuite, Il tient compte des délais de propagation correspondant à l'anatomie, c'est-à-dire à la géométrie du système. Par conséquent, on peut étudier un effet physiologique majeur comme l'induction de la mémorisation (

Ce qui s'appelle LTP ou potentialisation à long terme (Long-Term Potentiation).

non seulement au niveau cellulaire, mais aussi au niveau du réseau anatomique. Ceci est une grande avancée que nous exploitons désormais pour la recherche de médicaments efficaces. A la lumière de ces propriétés, comment pouvons-nous « voir » l'intelligence du mouvement ?

Considérons un mouvement principal, disons  $M$ , coordonné avec d'autres mouvements auxiliaires, disons  $M_i$ . Le mouvement « intelligent » correspondra à la suite des étapes suivantes (on se référera à la Erreur ! Source du renvoi introuvable. pour les détails anatomiques cités) :

- 1. Mémorisation (par le réseau des domaines de Purkinje) de mouvements  $M_i$  à partir d'expériences (c'est le cas de l'apprentissage de la marche, par exemple chez l'enfant). Il y a création de contextes ;

- 2. Stimulation (par les fibres moussues) par un mouvement donné  $M$  (par exemple l'obligation de l'exécuter face à un contexte imprévu, comme un évitement) ;

- 3. Situation du mouvement  $M$  (par les fibres grimpantes) par rapport aux contextes appris  $M_i$  ;

- 4. Prise de décision (par le réseau des domaines de Purkinje) : choix du contexte correct dans la contrainte d'environnement de  $M$ .



Ces quatre étapes de l'intelligence du mouvement sont identifiées mathématiquement et, bien sûr, anatomiquement. Évidemment, des circuits autres que le cervelet interviennent ici, comme le cortex cérébral qui déclenche le mouvement, disant d'aller du point A au point B. Mais dans les mouvements fins, le cervelet dit comment aller de A à B, car il a non seulement mémorisé une trajectoire à la suite d'un apprentissage, mais plus précisément les durées qui correspondent à tous les contextes appris, ce qui lui permet de restituer le « meilleur » mouvement, meilleur au sens d'une situation donnée d'environnement. Voilà ce que l'on peut dire sur l'intelligence du mouvement qui peut être « vue », imaginée, ce qui n'est pas le cas de l'intelligence cognitive, qui est plus intuitive. On peut cependant s'inspirer de l'interprétation précédente en quatre étapes pour formuler une définition générale de l'intelligence.

### **Comment définir la conscience ?**

L'intelligence pouvant être abordée dans le cadre de la théorie des interactions fonctionnelles de Gilbert Chauvet, est-il possible d'avoir une interprétation de la conscience dans ce même cadre ? Pour cela, il faut d'abord une définition mathématiquement opérationnelle de la conscience. Quatre axes paraissent nécessaires pour la définir :

- la conscience est réflexive, en ce sens qu'elle permet de se reconnaître dans l'autre ;

- la conscience doit émerger des processus cérébraux. Comme l'intelligence, elle est un phénomène biologique, une caractéristique du cerveau ;

- la conscience doit résoudre le problème du passage du phénomène objectif physique au phénomène qualitatif subjectif privé (ce qui est appelé « problème des qualia »).

- la conscience ne peut être un programme d'ordinateur. Il ne s'agit pas de l'exécution toujours identique d'un programme écrit à l'avance par un programmeur extérieur au système.

La première propriété paraît la plus importante. En effet, la reconnaissance de soi est la faculté de se prendre comme son propre objet d'observation, et par conséquent d'avoir connaissance de ses propres états mentaux. Elle permet donc de se mettre à la place de l'autre, de faire cette distinction entre le soi et le non-soi. On peut alors se représenter (et non savoir) ce que l'autre pense. C'est la « conscience-

miroir ». Elle apparaît à des degrés divers chez les organismes supérieurs, et en particulier chez les singes (Voir ci après, section 6.).

Dans le cadre de la théorie proposée par Gilbert Chauvet, on se trouve là face à une rupture. On a ici une interaction avec soi-même par l'intermédiaire des autres. Il existe une interaction fonctionnelle entre une unité, le cerveau, avec lui-même, à travers d'autres unités, sans transformation dans ces autres unités. L'interaction fonctionnelle est donc symétrique, à ce niveau supérieur de l'organisation. Elle possède une propriété très différente de ce que l'on a observé partout chez l'être vivant, et devrait par conséquent être à l'origine d'un comportement très particulier.

Des champs neuro-hormonaux couplent les sécrétions hormonales dans le cerveau à l'activité électrique (G. Chauvet, 1996). Ces champs dépendent de la topologie du cerveau et de sa géométrie. Ils interviennent donc nécessairement dans la genèse de la conscience. Comme ils représentent la dynamique d'interactions fonctionnelles non locales et non symétriques, ils contiennent un opérateur de champ (4) non local et non symétrique. Mais alors, comment peut être générée une solution symétrique sachant que les processus cérébraux sont des solutions d'équations de champ non symétriques (et non locales) ? Comment est réalisée cette rupture de la non-symétrie lorsque la conscience apparaît ?

La théorie proposée par Gilbert Chauvet est une théorie causale événementielle, car elle repose sur une représentation du système biologique en termes d'interactions fonctionnelles Source ? Puits (transformation), c'est-à-dire un ensemble d'événements dont l'existence des retards est une manifestation. Toute transformation dans un puits est le résultat de l'émission par une source qui est donc la cause du phénomène observé dans le puits. C'est là le propre de la causalité événementielle.

Or d'une part, au niveau du comportement collectif créé par l'unité fonctionnelle « cerveau », la non-symétrie de l'interaction fonctionnelle agit de façon réflexive par rapport à une autre unité « cerveau » en créant la distinction soi/non-soi. D'autre part, la causalité événementielle devient par rupture de la non-symétrie une causalité non-événementielle. La conscience apparaît donc comme une caractéristique du cerveau, décrite par une théorie physique. La conscience n'est plus une propriété physiologique. Elle devient – ou redevient – une propriété non-physiologique, appartenant au monde physique.

Évidemment, une question découle immédiatement de cette analyse. Comment la conscience, caractéristique du cerveau satisfaisant à des propriétés causales non événementielles, peut-elle être déduite dans le cadre d'une théorie causale événementielle ? Il faut bien que la conscience, en raison de sa nature ontologique à la première personne, présente un problème théorique fondamental.

Ce problème fondamental sera en partie résolu lorsque le problème de la rupture de non-symétrie sera résolu.

Dans ces conditions, quelle pourrait être la solution au problème de la conscience, qui est celui de la rupture de la non-symétrie dans l'organisme ? Autrement dit, comment déduire une propriété symétrique des processus issus d'interactions fonctionnelles non symétriques du cerveau ? Donnons quelques idées de solution. Gilbert Chauvet suggère une première voie, qui fait actuellement de sa part l'objet d'une formalisation mathématique non encore publiée (voir encadré ci-après)

### **Le texte en italiques est de Gilbert Chauvet**

Une première voie serait de rechercher les conditions conduisant les équations du champ neuro-hormonal à avoir des solutions de « type physique », en déterminant l'unité fonctionnelle siège du processus émergent non symétrique qu'est la conscience. Expliquons-nous par une digression mathématique. Une rupture de la non-symétrie correspond à une unité-puits dans laquelle il n'y a plus de transformation de sorte que le signal est ré-émis sans être modifié. Une solution symétrique du champ neuro-hormonal, qui serait dans ce cas une « représentation mentale », autrement dit un ensemble caractérisé d'activités neuronales, proviendrait d'un ensemble de réseaux de neurones en interaction rapide. Ce qui signifie que le retard dû à la propagation entre les réseaux tend vers zéro. On aurait alors une limite (au sens mathématique) de l'équation du champ définie par un opérateur symétrique déduit du S-Propagateur, celui obtenu pour un retard nul. Dans l'unité fonctionnelle « conscience », la propagation n'existe plus, mais la traversée des niveaux existe toujours.

Je vois cela comme une remémoration extrêmement rapide des souvenirs, une situation dans laquelle les patterns mémorisés sont rappelés tellement vite que l'on est capable de comparer instantanément la situation présente aux souvenirs « actuels ». Ce pourrait être cela la conscience, une performance du réseau nerveux réalisée seulement par des êtres doués de cette rapidité de remémoration. A cette limite de propagation nulle, le temps serait uniforme, non physiologique, puisque les effets de temps antérieurs n'existent plus. Effectivement, dans notre expérience, la conscience semble bien ne pas avoir de référence au temps. Une conséquence immédiate serait l'existence d'un temps uniforme pour la conscience, celui de la physique, ce qui est un argument supplémentaire (voir « Le temps en Biologie », G. Chauvet, à paraître) ».

Une autre voie serait de rattacher la conscience directement à la matière physique (quantique) comme le propose Roger Penrose (Roger Penrose. Les deux infinis et l'esprit humain. Flammarion 1999). Mais dans ce cas, il existerait sans doute une relation avec la solution précédente, relation qui lèverait l'impossibilité de la décrire dans notre représentation des fonctions physiologiques. Là encore, le temps de la conscience serait le temps uniforme de la physique newtonienne et non celui, non-uniforme, de la physiologie, avec ses échelles de temps différentes associées à chaque fonction.

Certains lecteurs seront sans doute déçus par cette explication, qui semble exclure la conscience de l'application de la théorie de la vie proposée par Gilbert Chauvet. Ne sera-t-on pas obligé, sauf à en revenir aux explications spiritualistes, de chercher dans le monde physique des causes « exotiques » au « hard problem » de la conscience, comme l'ont fait et continue à le faire de nombreux chercheurs. Mais pour Gilbert Chauvet, il n'en est rien. Pour lui, la conscience s'inscrit dans sa théorie, mais elle exige de la prendre à l'envers, si l'on peut dire. Cela traduit bien le fait que la conscience reste pour les scientifiques une propriété exceptionnelle. A propriété exceptionnelle, solution exceptionnelle. Bien évidemment, Gilbert Chauvet est le premier à reconnaître qu'il lui faudra préciser ses intuitions – ce qu'il est d'ailleurs en train de faire.

1 : Ces mécanismes sont construits sur des structures révélées par les recherches anatomiques, différentes les unes des autres. On perçoit la difficulté de mener ces recherches neuroanatomiques et neurophysiologiques avec des dimensions de l'ordre du micron ( $10^{-6}$  m) pour le neurone, alors qu'il existe des milliers de synapses par neurone, et que les canaux membranaires ont la taille de grosses molécules.

2 : Ce phénomène peut être facilement simulé numériquement en utilisant les équations qui le décrivent, et dont A. Hodgkin et A. Huxley (Hodgkin and Huxley 1952) ont donné l'interprétation jusqu'à maintenant toujours vérifiée. On connaît désormais la raison de ce comportement dit en « tout-ou-rien », donné par les non-linéarités mathématiques de ce système d'équations.

3 : Théorie mathématique du champ à n niveaux avec son formalisme des S-Propagateurs. Avec une théorie mathématique, un formalisme (mathématique) est nécessaire pour manipuler les concepts qui sont à la base de cette théorie. Ce formalisme associé à la théorie du champ dans une organisation hiérarchique manipule des opérateurs qui représentent le transport d'ions ou de molécules à travers les niveaux de l'organisation structurale. D'où le nom qu'il a donné à ces opérateurs : les S-propagateurs.

4 : Un opérateur de champ est un algorithme permettant de calculer les valeurs définies par un champ, ici le champ neuro-hormonal.