

A p e r ç u d e l a t h é o r i e  
s é l e c t i o n n i s t e

d e G e r a l d E d e l m a n

B e r n a r d P a s o b r o l a

*Gerald M. Edelman est directeur du Neurosciences Institute, à La Jolla, en Californie, président de la Neurosciences Research Foundation et chef du département de neurobiologie du Scripps Research Institute.*

P r i n c i p a u x o u v r a g e s

• e n a n g l a i s

- *The Mindful Brain*, avec V. Mountcastle, Cambridge-Mass., MIT Press, 1978
- *Dynamic Aspects of Neocortical Functions*, avec E. Gall et M. Cowan, New York, Wiley 1984
- *Neurobiology. An introduction to Molecular Embriology*, New York, Basic Books, 1988
- *Neural Darwinism. Selection Neural Groups Theory*, New York, Wiley, 1987
- *Bright Air, Brilliant Fire: On the Matter of the Mind*, Basic Books, 1992
- *A Universe of Conciousness: How matter becomes imagination*, avec Giulio Tononi, Basic Books, 2000

• e n f r a n ç a i s

- *Biologie de la conscience, Comment la matière devient conscience*, avec Giulio Tononi, éd. Odile Jacob, 2000
- *Plus vaste que le ciel : Une nouvelle théorie générale du cerveau*, éd. Odile Jacob, 2004
- *La Science du cerveau et la connaissance*, éd. Odile Jacob, 2007.

A p e r ç u d e l a t h é o r i e  
s é l e c t i o n n i s t e

d e G e r a l d E d e l m a n

Les travaux du neurobiologiste Gerald Edelman ont l'ambitieux projet de définir les propriétés clés de la conscience qui sont pour lui fondamentalement : 1) l'intégration car chaque expérience consciente est unique et indivisible ; 2) la différenciation car le nombre d'états de conscience que l'on peut éprouver en quelques millisecondes est immense. Partant de cela, la question est de savoir quels peuvent en être les substrats neuronaux.

Pour Edelman, la conscience est un *processus non causal* alors que sa base cérébrale possède une efficacité fonctionnelle. À chaque moment, la conscience est d'une seule pièce et toute l'expérience passée est engagée pour former la conscience intégrée de ce moment singulier – c'est ce qu'il nomme le « présent remémoré ». Le fait qu'une aire cérébrale puisse être essentielle ou nécessaire à la conscience ne signifie pas qu'elle est suffisante – tel ou tel groupe de neurones dans telle ou telle aire corticale peut d'ailleurs contribuer à l'activité consciente à un moment donné et pas au suivant. « Mon hypothèse, écrit-il, est qu'il s'agit d'un "noyau dynamique", rassemblement de groupes neuronaux qui à un moment donné interagissent plus entre eux qu'avec les autres groupes neuronaux. »

Edelman rejette donc l'idée selon laquelle le cerveau s'apparenterait à un quelconque modèle instructionniste reposant sur des programmes et des algorithmes. Il propose un modèle sélectionniste qui raisonne en

termes de *populations de neurones* et s'appuie sur la sélection d'éléments ou états particuliers tirés d'un vaste répertoire de possibilités. Pour lui, le processus de la conscience est donc une manifestation dynamique de l'activité des cartes neuronales réparties dans de nombreuses aires différentes du cerveau.

Comme le système immunitaire, ce système de reconnaissance réagit aux événements se produisant à l'extérieur du corps en fournissant une réponse adaptative. On dit que le système immunitaire est un système de « reconnaissance » car il est capable de distinguer au niveau moléculaire entre soi et non soi ; il agit par la mise en correspondance adaptative et continue des éléments d'un domaine physique aux nouveautés survenant dans les éléments d'un autre domaine physique, plus ou moins indépendant – ces deux domaines, en immunologie, sont les anticorps et les antigènes. Les ajustements en question ont lieu en l'absence de toute instruction préalable, par sélection au sein d'un ensemble hypervariable.

En neurosciences, l'interaction entre des structures neuronales dynamiques et des groupes d'informations et de stimuli provenant du monde extérieur conduit à la sélection de groupes de neurones et à l'amplification de la force de leurs synapses. Les trois grands principes de la théorie édelmanienne de la sélection des groupes de neurones, ou darwinisme neural, sont les suivants : (1) la sélection développementale donne lieu à un ensemble

extrêmement divers de circuits ; (2) la sélection par l'expérience donne lieu à des changements dans la force de connexion des synapses, ce qui favorise certaines voies et en affaiblit d'autres ; (3) le processus dynamique entraîne la synchronisation dispersée de l'activité de groupes de neurones et relie leur activité isolée d'un point de vue fonctionnel pour former des circuits capables de produire des sorties cohérentes.

Prenons l'exemple de deux cartes neuronales, dont l'une devrait spécialisée dans la détection de la couleur et l'autre dans celles des contours des objets extérieurs. Elles reçoivent des signaux du monde extérieur indépendamment l'une de l'autre. Mais il existe simultanément des interactions entre les cartes neuronales grâce à un processus nommé « réentrée ». Les signaux réentrants sont ceux qui relient fortement certaines combinaisons actives de groupes neuronaux appartenant à l'une des cartes à différentes combinaisons appartenant à l'autre. Les fonctions et les activités des deux cartes sont donc reliées, mais il n'y a pas de « carte de supervision » ou d'homunculus dans le cerveau dont le rôle serait de trier les signaux. En revanche, des liaisons s'établissent entre les cartes par sélection en parallèle et mise en corrélation des différents groupes neuronaux.

Dans un système sensoriel comme celui de la vision, de multiples régions corticales sont isolées d'un point de vue fonctionnel. Les aires spécialisées dans la couleur, le

mouvement, l'orientation peuvent dépasser la trentaine et sont réparties dans tout le cerveau. Et pourtant, comme on l'a dit, il n'existe pour Edelman aucune aire de supervision ou de système d'exploitation reliant la couleur, le bord, la forme et le mouvement d'un objet pour former un percept cohérent. Il n'existe pas d'« exécutif » instancié dans telle ou telle partie du cerveau. Pour extrapoler, disons qu'il n'y a pas d'État à l'intérieur du cerveau – qu'il soit « démocratique » ou « despotique » –, qui se chargerait de coordonner ces milliards d'opérations/seconde dont la finalité est d'adapter la vie d'un individu à son environnement. Mais comment se fait-il que, malgré l'absence de « carte de supervision », jusqu'à trente-trois cartes visuelles isolées d'un point de vue fonctionnel et très dispersées puissent déclencher une perception qui relie de façon cohérente des bords, des orientations, des couleurs et du mouvement en une seule image perceptuelle ?

Le modèle d'Edelman est dit sélectionniste, et non pas instructionniste, car il raisonne sur la base de populations de neurones – les diverses cartes sélectionnent et renforcent leurs liaisons utiles pour communiquer directement les unes avec les autres par réentrée. La réentrée n'est pas un feed-back correcteur comme dans le modèle connexionniste, mais une liaison entre des processus extrêmement dispersés dans les aires corticales. « La réen-

trée, c'est l'échange récursif permanent de signaux parallèles entre les aires cérébrales qui sert à coordonner l'activité des différentes zones dans l'espace et dans le temps. » (Edelman, *Plus vaste que le ciel*)

Des groupes variés de neurones de chaque carte sont donc reliés à ceux des autres cartes pour former un circuit fonctionnel. Mais ces liaisons ne sont pas fixes dans le temps. Durant la période suivante, différents neurones et groupes neuronaux peuvent former un circuit différent d'un point de vue fonctionnel, qui a cependant la même sortie. Et, de nouveau, au cours de la période suivante, un autre circuit se forme et utilise certains des mêmes neurones, ainsi que d'autres qui sont entièrement nouveaux dans des groupes différents. Ces circuits différents sont dits dégénérés. « La dégénérescence est l'aptitude qu'ont les éléments structurellement différents d'un système à assurer la même fonction et à produire la même sortie. » (*Op. cit.*) La « dégénérescence » est donc la propriété de produire des informations de nombreuses façons, pas toutes identiques structurellement, alors que les circuits structurellement divers déclenchent des sorties similaires. Cette redondance des réseaux biologiques accroît leur robustesse ainsi que leur adaptabilité aux modifications environnementales.

N'oublions pas que, pour Edelman, le but de cette recherche est de découvrir les propriétés clés de la conscience, puis d'en rechercher les substrats neuronaux. Les propriétés fondamentales de la conscience

sont pour lui 1) l'intégration : chaque expérience consciente est unique et indivisible ; l'intégration est permise par la réentrée ; 2) la différenciation : le nombre d'états de conscience que l'on peut éprouver en quelques millisecondes est immense. Partant de cela, quels peuvent en être les substrats neuronaux ?

La conscience primaire est l'état qui permet de se rendre compte de la présence des choses dans le monde, d'avoir des images mentales. Concernant le système nerveux des mammifères, la catégorisation perceptive est assurée par des interactions entre les systèmes sensoriels et moteurs, ce qu'Edelman nomme des encartages globaux. L'aptitude à réaliser des catégorisations perceptives – à « donner sens » au monde – permet à un animal de découper les signaux provenant de son corps et de l'environnement en séquences qui donnent lieu à un comportement adaptatif. La catégorisation perceptive et la formation de concepts ne permettraient pas à un animal de s'adapter en l'absence de mémoire. Il faut noter que la mémoire n'est pas représentationnelle au sens où elle stockerait un code enregistré de façon statique pour un acte. Il vaut mieux y voir une propriété d'interactions non linéaires dégénérées dans un réseau multidimensionnel de groupes neuronaux. Ces interactions permettent de faire « revivre », mais non de façon identique, un ensemble d'actes et d'événements antérieurs, même si l'on a souvent l'illusion qu'on se souvient exactement de la façon dont un évé-

nement est arrivé. Les souvenirs sont nécessairement associatifs et jamais identiques. Pour autant, sous diverses contraintes, ils peuvent être suffisamment efficaces pour déclencher la même sortie.

Selon la théorie de la sélection neuronale, les catégorisations s'opèrent toujours par rapport à des critères de valeurs internes, qui ne déterminent pas des catégorisations spécifiques mais imposent des contraintes aux domaines dans lesquelles elles se produisent. Par exemple chez les animaux d'une espèce donnée, les bases de systèmes de valeurs ont déjà été établies par sélection au cours de l'évolution. Elles sont pré-encartées, si l'on peut dire, et elles se manifestent dans des régions du cerveau qui participent à la régulation des fonctions corporelles : rythme cardiaque, respiration, comportement sexuel et alimentaire, fonctions endocrines, fonctions végétatives.

Cette mémoire de « valeur-catégorie » est liée par des voies réentrantes (lignes renforcées) à la catégorisation perceptive actuelle des signaux du monde. C'est principalement ce processus qui engendre la conscience primaire. Quand elle apparaît à travers de nombreuses modalités (vue, toucher, etc.), la conscience primaire est celle d'une « scène » faite de réponses à des objets et à des événements, certains n'étant pas nécessairement connectés de façon causale les uns aux autres. Un animal doté d'une conscience primaire peut cependant effectuer des discriminations et relier ces objets et événements par la mé-

moire, son expérience antérieure comportant des valeurs. Cette aptitude favorise sa valeur de survie. Mais elle ne s'accompagne pas de conscience de soi, ni du présent et de l'avenir.

L'évolution ultérieure d'autres circuits réentrants a permis l'acquisition de la capacité sémantique et finalement du langage et a donné naissance à la conscience de niveau supérieur chez certains primates comme nos ancêtres hominidés (et peut-être aussi un grand nombre d'autres espèces de singes). La conscience de niveau supérieur procure la capacité d'imaginer le futur, de se rappeler explicitement le passé et d'être conscient d'être conscient. Contrairement à la conscience primaire, la conscience d'ordre supérieur ne se fonde pas sur l'expérience en cours mais sur la capacité de modéliser le passé et le futur. Pour expliquer ce phénomène, Edelman fait l'hypothèse de la présence d'un « noyau dynamique » rassemblant des groupes neuronaux qui, à un moment donné, interagissent plus entre eux qu'avec les autres groupes neuronaux. Le noyau donc est un faisceau fonctionnel qui se parle surtout à lui-même.

Le système thalamocortical, qui donne naissance au noyau dynamique, est représenté par un fin maillage d'aires corticales et thalamiques et par des connexions réentrantes. Les réponses déclenchées par le noyau dynamique réentrant peuvent stimuler des réponses non conscientes. Elles voyagent le long de voies parallèles, polysynaptiques et unidirectionnelles qui quit-

A p e r ç u d e l a t h é o r i e  
s é l e c t i o n n i s t e

d e G e r a l d E d e l m a n

tent le cortex, atteignent les divers composants des ganglions de la base et certains noyaux thalamiques, pour finalement retourner au cortex. De cette manière, les réponses au service de la conscience peuvent se connecter aux structures d'activité des aires non conscientes. Grâce à ces interactions, le noyau dynamique relie la mémoire valeur-catégorie à l'organisation perceptive. En outre, il sert à connecter les cartes mémorielles et conceptuelles les unes aux autres.

Un animal ou un nouveau-né fait l'expérience d'une scène en référence à un soi, mais il n'a pas de soi que l'on puisse nommer et qui puisse être différencié de l'intérieur. Un tel soi qu'on puisse nommer apparaît chez les humains lorsque la conscience d'ordre supérieur se développe pendant l'élaboration des aptitudes sémantiques et linguistiques, et des interactions sociales. Ce n'est qu'avec l'évolution de la conscience d'ordre supérieur, fondée sur les aptitudes sémantiques, qu'apparaissent les concepts explicites du soi, du passé et du futur. Cela implique que l'activité neurale fondamentale du noyau dynamique réentrant convertit les signaux issus du monde et du cerveau en « transformation phénoménale ». Être un animal conscient, c'est avoir des qualia. Le mot latin « quale » (pluriel « qualia ») signifie « quelle sorte » ou « quel type ». Ce sont donc à la fois des propriétés de l'expérience sensible et des phénomènes totalement privés (personne ne peut savoir la façon dont un autre

individu ressent la couleur rouge, ou entend un son, ou, selon la célèbre formule du philosophe Thomas Nagel, « quel effet cela fait d'être une chauve-souris ? »)

La transformation phénoménale (à savoir notre expérience des qualia) « reflète l'aptitude à effectuer des distinctions d'ordre supérieur qui seraient impossibles sans l'activité neurale du noyau. » (*Plus vaste...*) Edelman suppose que la transformation phénoménale est corrélée à cette activité neurale, mais qu'elle n'est pas causée par cette activité ; c'est plutôt une propriété simultanée de cette activité. Ce qui l'amène à la seconde question : la transformation phénoménale est-elle causale ? Edelman insiste sur le fait que *cette transformation phénoménale est un processus et non une substance*. Ce processus est déclenché par les activités du noyau. Étant donné la nature strictement causale du monde, c'est le noyau qui est causal, pas l'expérience phénoménale (qualia). Les animaux dotés d'un tel noyau au cours de l'évolution ont acquis la capacité d'effectuer des discriminations raffinées. Les qualia sont précisément ces discriminations, des formes multidimensionnelles de différenciation portées par un cerveau complexe, chacune suscitée par un état différent du noyau. Mais la question de savoir si les qualia existent dans le monde traduit une erreur semblable à l'assertion selon laquelle les catégories sensorielles comme la couleur et diverses autres per-

A p e r ç u d e l a t h é o r i e  
s é l e c t i o n n i s t e

d e G e r a l d E d e l m a n

ceptions existent indépendamment de l'esprit et du langage.

En résumé, les états conscients reflètent l'intégration des états neuraux dans le noyau. La conscience est une propriété des processus neuraux et ne peut elle-même agir de façon causale dans le monde. « Munis de ce tableau, nous pouvons clarifier un grand nombre d'erreurs logiques et d'incohérences sémantiques qui ont affligé les études sur la conscience. L'une de ces erreurs est l'incapacité à distinguer la causalité physique et l'entraînement logique. Les hypothèses selon lesquelles c'est l'action du noyau thalamocortical qui cause la conscience se heurtent à une difficulté. Puisque les causes précèdent les effets, ces hypothèses impliquent l'existence d'un écart temporel entre processus incommensurables. Au contraire, l'action neurale dans le noyau entraîne la conscience, tout comme le spectre de l'hémoglobine du sang est entraîné par la structure mécanique et quantitative de cette molécule. » (Edelman, *La Science du cerveau et la connaissance*)

Nous ne savons pas à quel moment de l'histoire de l'évolution la conscience primaire est apparue pour la première fois. Cependant, suppose Edelman, si l'on compare les structures neurales similaires requises pour son expression chez les humains et les autres vertébrés (par exemple, un système thalamocortical et des systèmes de valeur ascendants, ainsi que certaines structures comportementales), la conscience primaire semble être apparue

chez les vertébrés d'abord à la transition entre les reptiles et les oiseaux, puis lors de la transition entre reptiles et mammifères.

Lorsque les circuits réentrants indispensables ont évolué chez les primates supérieurs et finalement chez Homo sapiens, un concept du soi est apparu avec ceux du passé et du futur. L'aptitude à construire une scène liée à l'histoire des valeurs-catégories d'un individu marque l'apparition d'un soi. Un organisme doté d'un soi peut effectuer de riches discriminations fondées sur l'histoire de son apprentissage passé et il peut recourir à la conscience pour former des plans, du moins pendant la période représentée par le présent remémoré. L'intégration complexe du noyau dynamique, modulée par l'histoire comportementale et les souvenirs des événements d'apprentissage individuels, donne lieu au comportement adaptatif nécessairement idiosyncrasique à cet organisme individuel.

Les animaux ayant évolué de cette façon ont pu communiquer les états résultant des activités du noyau en termes phénoménologiques. Le fait que le monde soit causalement fermé et que seules les activités du noyau soient causales – *i.e.* causales dans le monde et non pas causes de la conscience – ne contredit pas le rôle que joue l'état phénoménal en tant que véhicule de communication. La transformation phénoménale est un moyen élégant d'exprimer les états intégrés du noyau à la première personne. Il n'existe pas d'autre

manière de faire directement l'expérience de ces événements neuraux. Même dans l'échange entre deux humains conscients, la transformation consciente fournit une indication des relations causales sans être causale elle-même (*Plus vaste...*).

L'état subjectif reflète les propriétés actuelles des états neuraux du noyau. Mais on ne sera pas surpris de lire que, même si le processus conscient implique la représentation, le substrat neural de la conscience n'est pas représentationnel. Les contenus représentationnels, images mentales et autres, apparaissent dans une scène de la conscience primaire en grande partie en vertu des mêmes processus neuraux qui font apparaître les images perceptives directes. D'autre part, les concepts ne doivent pas nécessairement reposer sur les images, mais plutôt sur les encartages globaux et certaines activités des systèmes moteurs qui n'engagent pas nécessairement le cortex moteur et ne donnent donc pas lieu à du mouvement.

Les états sous-jacents à la conscience ont une extraordinaire associativité et une remarquable ouverture à la sensation, à la perception, à la mémoire, à l'imaginaire et à diverses combinaisons de tout cela. Le grand encartage du noyau dynamique réentrant qui est réparti dans le cortex correspond à cette propriété. Dans l'imaginaire, par exemple, la réentrée engage plus ou moins les mêmes ensembles de voies qui sont mobilisées par la perception visuelle primaire, ainsi que d'autres voies associatives. La propriété d'associativité

vient de la réentrée et des interactions dégénérées entre les circuits thalamocorticaux qui constituent le noyau. La mémoire est non représentationnelle et a aussi des propriétés dégénérées qui permettent de riches associations avec tout un éventail de circuits en plus de ceux qui sont impliqués dans n'importe quel souvenir donné. Certaines données montrent que des primates comme les chimpanzés ont des aptitudes sémantiques, mais presque pas d'aptitude syntaxique, de sorte qu'ils n'ont pas de vrai langage. Pour autant, on sait qu'ils sont capables de reconnaître des images d'eux-mêmes et de raisonner sur les conséquences des actions d'autres chimpanzés ou d'humains. Dès lors et étant donné leurs aptitudes sémantiques, il est probable qu'ils sont dotés d'une forme de conscience d'ordre supérieur. Cela se produit lorsque les aires formant les concepts qui sont impliqués dans la conscience primaire sont liées par des circuits réentrants aux aires médiatisant la capacité sémantique. Présente chez les primates supérieurs, elle atteint son expression la plus développée chez les êtres humains, lesquels possèdent une véritable capacité linguistique. L'aptitude à relier des signes pour former un lexique au moyen d'une syntaxe accroît grandement l'éventail d'expression réentrante. Alors que la conscience d'ordre supérieur, lorsqu'elle apparaît, dépend encore de la conscience primaire, disposer de signes et de moyens de ce type permet à un individu de



A p e r ç u d e l a t h é o r i e  
s é l e c t i o n n i s t e

d e G e r a l d E d e l m a n

se libérer temporairement de ses liens avec le présent remémoré.

Fondamentalement, le cerveau fonctionne par reconnaissance de structures avant même l'apparition du langage. Il a ce qu'on pourrait appeler des capacités « prémétaphoriques ». La puissance de ces aptitudes analogiques repose sur l'associativité résultant de la dégénérescence des réseaux neuraux. On a vu que, pour Edelman, les états du noyau eux-mêmes ne « représentent » pas terme à terme une image, un concept ou une scène donnée. N'importe quelle représentation peut correspondre à de nombreux états neuraux sous-jacents et signaux dépendants du contexte. Dépendant des entrées, de l'environnement, de l'état du corps et d'autres contextes, les différents états du noyau peuvent sous-tendre une représentation donnée. Les résultats des aptitudes métaphoriques sont nécessairement ambigus, mais peuvent être très créatifs. Les cerveaux humains sont très performants pour établir des structures, mais ils sont en même temps constamment portés à l'erreur. On le voit dans les illusions de perception ainsi que dans les croyances de niveau supérieur. Mais comme le montre l'analyse de l'apprentissage, ils peuvent en général procéder à des corrections d'erreurs en réponse à des récompenses et à des punitions adaptées.

En tant que systèmes sélectionnistes, les cerveaux n'opèrent donc pas de prime abord par la logique et les mathématiques,

mais leur « mode précoce et fondamental de pensée, qui dépend beaucoup de la reconnaissance de structures, implique la métaphore. C'est un reflet de la portée et de l'associativité des réseaux extrêmement complexes et dégénérés du cerveau. Les productions de la pensée métaphorique peuvent se comprendre, mais pas se prouver, comme la comparaison ou les propositions logiques. Par exemple, si je dis : "Je suis au soir de ma vie", cet énoncé se comprend, mais il ne se prouve pas » (*La Science du cerveau et la connaissance*). Les métaphores ont un pouvoir allusif remarquablement riche. On peut recourir à la logique pour modérer les effets, mais elle ne peut elle-même être créative au même degré. L'expérience consciente fournit un échantillon de l'équilibre entre ces deux modes de pensée et des richesses infinies de leurs substrats neuraux sous-jacents. La conscience d'ordre supérieur peut être considérée comme un troc entre précision absolue et possibilités imaginatives riches. Ainsi, les discriminations multidimensionnelles ont une valeur adaptative. « Ce qu'elles perdent en précision absolue, elles le gagnent pour augmenter notre aptitude à généraliser, à imaginer et à communiquer dans un environnement riche. » (*Plus vaste...*)