

Que
sais-je ?

LA NEUROPHILOSOPHIE



Bernard Andrieu

puf

QUE SAIS-JE ?

La neurophilosophie

BERNARD ANDRIEU

Professeur d'épistémologie du corps et des pratiques corporelles

Faculté du sport /UHP/Nancy Université

ACCORPS /LHPS Archives Poincaré UMR 7117 CNRS

Deuxième édition mise à jour

6^e mille



Introduction

À la mémoire d'Édouard Zarifian (1941-2007)

« Il nous faut donc reconstruire le cerveau comme objet philosophique. »

D. Lecourt.

Le but de notre étude est d'établir les conditions d'un dialogue entre la philosophie et les neurosciences à partir de l'étude critique de la notion de « neurophilosophie ».

Le néologisme « neurophilosophie » apparaît comme le titre du premier ouvrage de la philosophe américaine Patricia Smith Churchland publié en 1986 *Neurophilosophy, toward a Unified Science of the Mind-Brain*. Dès 1980, William J. Davis faisait paraître un article à propos des « Neurophilosophical reflections on central nervous pattern generators » : il affirmait déjà sa volonté de dépasser les limites traditionnelles de la science et d'aller au-delà de la division de la philosophie, de l'épistémologie et de la logique. L'apparition de ce néologisme est limitée : nous en avons retrouvé des significations différentes. Elles ne renvoient pas systématiquement au sens défini par P.

S. Churchland : pour elle la neurophilosophie est la science unifiée de l'esprit-cerveau. Le néologisme a connu sa première traduction en langue française à l'Université de Liège en 1989 sous la plume de Michel Dupuis dans son opuscule Les neurones en personne. Six propositions pour une neurophilosophie. Claude Debru, à partir de son travail dans le laboratoire du neurophysiologiste du sommeil Michel Jouvet, publie en 1990 le premier titre français, Neurophilosophie du rêve, afin « qu'un champ spéculatif s'ouvre à la neurophilosophie, une philosophie qui ne serait pas seulement critique ». En 1991, le philosophe américain G. R. Gillet applique le néologisme au problème classique de la douleur dans son article « The neurophilosophy of pain ». En 1992, P. S. Churchland, associée à Y. Christen, investit le champ de la psychoneuropharmacologie en dirigeant l'ouvrage Neurophilosophy and Alzheimer's Disease.

En 1997, la neurophilosophie est le domaine des partisans de l'union du cerveau et de la pensée [\[1\]](#). Mais ce néologisme a été compris par nous plus comme une interprétation des résultats des neurosciences que comme une cause du développement des neurosciences dans le champ de la philosophie. Bien évidemment, la neurophilosophie a produit depuis 1986 un effet de regroupement des partisans des thèses éliminativistes et réductionnistes selon lesquelles les états mentaux sont réductibles à des états cérébraux. Nous faisons l'hypothèse que la production de ce néologisme est le résultat de l'histoire

des relations entre philosophie et neurosciences, si bien qu'une reconstitution précise des travaux, des concepts et des paradigmes a été rendue nécessaire.

En France, depuis la parution en 1983 de *L'homme neuronal* de J.-P. Changeux, l'entreprise de naturalisation des objets philosophiques, mais plus largement des objets traditionnellement étudiés par les sciences humaines, a assuré le succès de maisons d'édition et de nouvelles collections. Les neurosciences viendraient renouveler entièrement les modèles, les concepts et les théories usées produites entre 1850 et 1950 par les sciences humaines. Ainsi la conscience de soi, l'inconscient, le rêve, la perception, le langage, les émotions, les passions, le désir, l'agressivité, l'intelligence, les apprentissages, la mémoire, la sexualité, les pathologies mentales... ne pourraient plus avoir une définition socioculturelle. La crise du structuralisme aura été dû à son incapacité d'ouvrir aux sciences de la vie le développement de la psychanalyse, de la philosophie ou encore de l'anthropologie. Les sciences cognitives, très développées depuis une dizaine d'années, semblaient avoir résolu cette crise en privilégiant le sujet et le sens. S'il est vrai que les sciences cognitives renouvellent les théories de la complexité en situant les différents niveaux et en évitant la réduction des discours, l'arbre, si gros soit-il, cache la forêt.

Car les neurosciences n'ont jamais renoncé au statut de science unifiée. À la différence d'une explication

totalitaire du corps humain les neurosciences proposent une multiplication de travaux dont l'addition a produit la neurophilosophie de l'homme. Tous les scientifiques ne partagent pas la réduction de l'homme à son cerveau et l'élimination des sciences humaines : une histoire des neurosciences suffirait à le démontrer. Mais les neurosciences n'échappent pas à l'idéologie : la neurophilosophie défend une idéologie naturaliste face à d'autres idéologies ; son matérialisme réductionniste ne doit pas empêcher le philosophe de dégager un autre matérialisme : celui, qui depuis Diderot, comprend l'homme comme un corps dynamique et sensible au monde.

N'étant ni tout à fait neuroscientifique, ni tout à fait philosophe, le neurophilosophe voudrait être les deux en constituant la fusion des sciences exactes et des sciences humaines. Ce rêve de la science unifiée, s'il est une illusion, vise notre analyse de la matière vivante. Comment dès lors la matière produit-elle de l'homme ?

Notes

[1] William J. Davis, Neurophilosophical reflections on central nervous pattern generators, *The Behavioral and Brain Sciences*, 1980, 3, p. 543-544 ; P. S. Churchland, *Neurophilosophy, Toward a Unified Science of the Mind-Brain*, mit, 1986, 546 p. ; Michel Dupuis, *Les neurones*

en personne. Six propositions pour une neurophilosophie, Liège, Université libre, 1989, 29 p. ; Claude Debru, Neurophilosophie du rêve, Paris, Hermann, 1990, (rééd. 2006), 398 p. ; G. R. Gillet, The neurophilosophical of pain, Philosophy, 66, 1991, p. 191-205 ; P. S. Churchland et Y. Christen, Neurophilosophy and Alzheimer's Disease, Berlin, Springer-Verlag, 1992 ; C. Poirel, Le cerveau et la pensée. Critique des fondements de la neurophilosophie, Paris, L'Harmattan, 1997.

Chapitre I

Genèse de la neurophilosophie

L'école de San Diego

Après avoir obtenu, en 1962, le prix Nobel pour sa découverte, avec James Watson (1928), de la structure d'adn, le physicien et biochimiste Francis Crick (1916-2004) s'est orienté, depuis le milieu des années 1970, vers les neurosciences. Travaillant au Salk Institute de la Jolla, en Californie, il a pu réunir autour de lui les principaux partisans du matérialisme éliminativiste : Paul M. Churchland (1942), Patricia S. Churchland (1943), Terrence J. Sejnowski, Christof Koch... Pour P. S. Churchland et tout le courant du matérialisme éliminativiste, les états mentaux n'existent pas. Il n'y a que des états neurobiologiques. P. S. Churchland n'a jamais caché ses options matérialistes. Pour elle la compréhension de l'esprit passe nécessairement par les neurosciences. « Puisque j'étais matérialiste et par suite croyais que l'esprit est le cerveau, il me semblait évident qu'une bonne compréhension des neurosciences ne pouvait être utile que si je cherchais

à connaître comment nous voyons, comment nous pensons, raisonnons et décidons. » [1] Dès les années 1970 elle se tourne vers la neurobiologie. D'abord elle s'informe, rencontre les spécialistes, puis dès 1978 elle attaque très durement la psychologie cognitive du philosophe Jerry Fodor afin de se démarquer déjà du courant fonctionnaliste.

Trois étapes marquent la démonstration de sa thèse neurophilosophique :

- de 1980 à 1983 elle fait l'inventaire des problèmes de la neurobiologie en prenant systématiquement position contre J. Fodor, J. Eccles (1903-1997), K. Popper (1902-1995), les travaux du neurobiologiste Benjamin Libet. Durant cette période, l'affirmation du réductionnisme n'est pas entière ;
- de 1983 à 1985 elle rédige *Neurophilosophy* qui paraît en 1986 dans la collection « Computational models of cognition and perception » des éditions mit Press (traduit en français chez puf en 1999). Elle développe une réflexion critique sur le statut de la conscience dans son article « Consciousness, the transmutation of a concept ». Pourtant la perspective neurocomputationnelle est peu affirmée à ce moment-là, même si en 1983 elle publie, avec son mari Paul M. Churchland, *Stalking the Wild Epistemic Engine*. L'analogie

cerveau-ordinateur, présupposée par la notion de computation, sera d'ailleurs développée par les deux penseurs lors de la dernière étape ;

- depuis 1988 P. M. Churchland, P. S. Churchland et T. J. Sejnoswki publient de nombreux articles sur cerveau et computation. Le neurocomputationnisme, nouvelle forme de connexionnisme, permet à P. M. Churchland de publier en 1989 A Neurocomputational Perspective. The Nature of Mind and the Structure of Science, tandis que P. S. Churchland et Terrence J. Sejnoswki publient en 1992 The Computational Brain. P. M. Churchland publie quant à lui en 1995 The Engine of Reason, the Seat of the Soul. A Philosophical Journey into the Brain ;
- enfin depuis 1998 P. S. Churchland étudie le problème de la conscience, de la connaissance de soi de la volonté et de l'apprentissage à partir de Brain Wise. Studies in Neurophilosophy et d'une synthèse avec Paul Churchland sur les controverses neurophilosophiques On the Contrary : Critical Essays 1987-1997 (1998).

La neurophilosophie recherche les fondements naturels – c'est-à-dire les connaissances et les faits du monde vivant – qui peuvent être invoqués pour expliquer un comportement, une décision ou une prise

de position éthique qualifié de « méta-éthique biologique » ou « neuro-éthique » [\[2\]](#)

P. S. Churchland reconnaît son embarras à vouloir construire une théorie d'une science en cours d'élaboration et dont les résultats ne peuvent être définitivement établis. La première difficulté réside dans le passage d'une compréhension du comportement individuel du neurone (de ses propriétés membranaires et axoniques, des phénomènes synaptiques, de ses modes de connectivité, du transport des matériaux intracellulaires, de son métabolisme, de son épigénèse) à une théorie universelle des états neurobiologiques. Plus que le traditionnel passage du simple au complexe cette difficulté tient à l'objet même de la théorie neurophilosophique : son objet est lié à son activité, c'est-à-dire à sa capacité à décrire des réseaux au moment de la production de la tâche neurocognitive. La mobilité et la variabilité des situations cognitives ne permettent pas d'établir actuellement une théorie exhaustive des représentations où la forme serait reliée à un contenu. En cela le matérialisme éliminativiste se distingue d'un physicalisme radical et simpliste qui a cru qu'il suffirait d'établir les corrélations entre les états neurobiologiques et les états mentaux pour réduire les seconds aux premiers. Même si, par bien des côtés l'entreprise d'élimination des états mentaux s'apparente au physicalisme radical, elle s'en distingue par la prise en compte des réseaux neuronaux. Car la description par réseaux des

opérations neurocognitives se veut complexe.

La seconde difficulté tient au désir d'avoir une vue d'ensemble des neurones d'un point de vue, non pas neurobiologique mais neurophilosophique. L'entreprise de théorisation va à l'encontre même de l'évaluation qu'effectue dans les années 1980 certains neurobiologistes qui, à l'image de A. I. Selverston, déclarent : « Le temps pour les théories n'est pas encore arrivé, dans la mesure où nous ne connaissons pas assez le détail de la structure. » [3] Ce désir totalitaire d'embrasser l'ensemble des données pour parvenir à formuler une théorie scientifique est compréhensible : dans le pur respect de la méthode expérimentale, les neurosciences n'aboutiront à la théorisation qu'après avoir pu formuler la loi qui universalise les phénomènes observés. Le reproche est aussi celui de l'abstraction qui fournirait à cette théorie neuroscientifique un caractère incontrôlable expérimentalement, si elle était élaborée en deçà des résultats ultimes des recherches.

P. S. Churchland n'est pas insensible à l'argumentation d'autant qu'elle n'a eu de cesse d'accuser les positions dualiste et fonctionnaliste d'abstraction. En tenant compte des résultats neuroscientifiques, P. S. Churchland renverse l'argumentation critique de A. I. Selverston en estimant que les neurosciences seraient suffisamment avancées pour établir une théorie du fonctionnement cérébral qui correspondrait à l'actualité des recherches, sans que les scientifiques puissent

se sentir déjugés. En s'appuyant sur un travail critique de la thèse de Selverston, selon laquelle la théorisation du cerveau serait inactuelle et invérifiable, P. S. Churchland invalide la séparation entre théorie et pratique en proposant l'élaboration théorique de la complexité des niveaux d'organisation de la matière cérébrale. En effet les interactions entre les composants du cerveau ne sont pas linéaires ; elles supposent la description de la connectivité des neurones sous la double forme des réseaux et des modes de spécification attachés à la réalisation des tâches cognitives.

Pour cela P. S. Churchland se réfère, sans le développer, ici l'article fondateur de William J. Davis auquel elle emprunte l'expression de neurophilosophie en lui donnant une forme substantive : « Neurophilosophical reflections on central nervous pattern generators » (Behavioral and Brain Sciences, 1980, 3, p. 543-544). Cet article est un commentaire de l'article d'Allen I. Selverston (1980) publié dans la même revue et intitulé « Are central pattern generators understandable ? ». Contre le pessimisme de A. I. Selverston de parvenir à une compréhension totale du fonctionnement des neurones à partir des techniques présentes, la réaction s'organise puisqu'un certain nombre de neurobiologistes réclame une nouvelle génération de méthode expérimentale et théorique en neurobiologie, ou une prise en compte de la complexité des systèmes qui ne sont pas si simples qu'on pourrait le penser. L'article de William J. Davis se trouve

entre ces deux commentaires critiques.

Origine du néologisme

W. J. Davis appartient à ce moment-là aux Laboratoires Thimann, de l'Université de Californie, Santa Cruz. P. S. Churchland appartient aussi à l'Université de Californie, San Diego pendant la rédaction terminale du livre, si bien que l'on peut supposer qu'elle l'ait rencontré ou sinon lu. Pourtant aucune étude précise de l'article n'a eu lieu dans l'œuvre de P. S. Churchland, et sans cette référence bibliographique nul ne pourrait savoir l'origine du néologisme neurophilosophy.

Aussi pour rendre compte aussi précisément que possible des enjeux, nous proposons ici une traduction du court article de W. J. Davis :

« Au cours des années je suis arrivé pour l'essentiel à la même conclusion que A. I. Selverston, mais pour quelques différentes raisons. Le principal obstacle à la compréhension des modèles générateurs n'est pas, selon mon point de vue, "notre incapacité à identifier et à caractériser le pool total interneuronal qui fait un Modèle central générateur", comme l'écrit A. I. Selverston. Si cela venait de l'étendue du problème, on pourrait trouver simplement la solution en découvrant les cellules appropriées, ce qui est faisable dans les systèmes nerveux de

certaines invertébrés. Le principal problème est beaucoup plus profond, transcendant les limites de la science conventionnelle et gagnant la philosophie, l'épistémologie et la logique. Le fait est que, même quand nous aurons trouvé comment chaque neurone simple est relié à un réseau donné et cartographier toutes leurs interconnexions synaptiques, il sera encore impossible de comprendre, au strict sens du terme, comment le réseau remplit sa fonction. Le problème peut être abordé dans le contexte de deux questions plus générales : 1 / Pouvons-nous comprendre une propriété "émergente" en des termes élémentaires ? 2 / Quelle est la nature de la "preuve" que nous pouvons apporter pour nous convaincre et persuader nos collègues que nous avons véritablement atteint une compréhension non équivoque du fonctionnement d'un réseau ?

La première question a été abordée de manière instructive avec une métaphore introduite par Don Kennedy (1971) et reprise plus récemment par Ted Bullock (1980), nommément, le jeu de base-ball. Un stade plein avec des gens observant un jeu de base-ball est analogue à une collection de cellules cérébrales opérant de concert pour réaliser une fonction neuronale spécifique. Dans les deux le stade et le cerveau, les acteurs principaux du drame sont individuels, des éléments répétés (les gens et les neurones respectivement). Dans les deux cas il y a un "but" pour l'ensemble des

éléments, et ce but est en principe connu. Et dans les deux cas le but entraîne la coordination des éléments individuels dans des modèles d'activité spécifique selon un certain nombre de règles qui sont pareillement définies en principe.

Maintenant si un d'entre nous voulait tenter d'apprendre les règles du base-ball, en se rappelant des déclarations d'un simple fan, en observant les signaux manuels de l'attrapeur défensif, ou même en écoutant les bavardages de l'abri des entraîneurs de l'équipe adverse. Le jeu de base-ball ne peut pas, par nature, être compris seulement par les termes des unités simples de comportement dont il est composé. C'est l'action collective de centaines ou de milliers d'éléments individuels qui, ensemble, définit le jeu et lui donne tout le sens qu'il peut avoir.

Peut-on attendre moins du système nerveux ? Beaucoup des plus intéressantes et fondamentales fonctions du système nerveux ne peuvent être approchées à partir du niveau des éléments simples qu'il comprend. L'exemple de la personnalité humaine peut servir d'illustration sur ce point. Si vous êtes matérialiste, alors vous souscrivez à l'idée que nos personnalités émergent de nos cellules nerveuses dont les fonctions sont à leur tour spécifiées par le code génétique et formées par les histoires de nos apprentissages individuels. Et pourtant il doit être

évident que les neurones individuels ne possèdent pas de “personnalité” dans le même sens qu’un être humain en possède. La personnalité émerge à partir des interactions des billions de neurones, dans lesquelles chaque cas peut être analysé seulement au niveau à partir duquel il émerge. Les mécanismes de la personnalité humaine ne peuvent être situés au niveau cellulaire. Silverston, pour l’essentiel, affirme que la même chose peut être vraie même pour le problème simplement plus vaste du modèle central générateur, et avec des compétences majeures qui doivent être considérées, je crois.

La seconde question posée plus haut est même plus éthérée : Quelle est la nature d’une “preuve acceptable” dans les neurosciences ? Comme cela a été illustré par le récent BBS sur le traitement de la commande neuronale (Kupferman et Weiss, 1978), certains leaders dans le champ ne sont pas d’accord sur le critère de définition des éléments simples d’un réseau moteur. Dans le cas du modèle central générateur nous manquons également de terminologie, de concepts et de critères pour décider de ce qui est acceptable comme preuve pour un rôle ou une fonction neurale particulière. Un exemple saillant sont les expériences en phase de changement. La capacité de recomposer un rythme moteur par l’explosion interpolée d’une activité dans un neurone donné est considérée par certains scientifiques comme

aussi évident que le neurone stimulé est une partie du réseau du type générateur. Ce critère est rencontré par les membres d'une population de neurones cérébraux qui participe au comportement alimentaire du mollusque *Pleurobranchaea* (neurones paracérébraux), et pourtant le modèle alimentaire peut aussi bien être engendré quand la même cellule est silencieuse. Sont-ce les cellules paracérébrales qui appartiennent ou non au modèle générateur ? Je crois que le critère nous manque pour décider ; nous ne pouvons pas préciser ce qu'est "preuve acceptable" dans ce cas et dans bien d'autres aussi cruciaux. Peut-être que les problèmes comme ceux-ci reflètent-ils l'état embryonnaire de notre science ; ou bien, peut-être reflètent-ils les limites dans la capacité du cerveau humain à comprendre ses propres fonctions.

À partir de ces limitations, devrions-nous abandonner l'approche cellulaire pour comprendre le système nerveux ? Évidemment non. Elles sont des cas choisis dans lesquels le modèle de la génération devait être incompris dans les termes d'actions élémentaires. Les études d'avant-garde de Selverston sur le modèle de la génération dans le ganglion stomatogastrique peuvent fournir l'exemple le plus consistant. En outre, il y a une grande évidence que même les phénomènes comportementaux "complexes", incluant le choix et la plasticité, sont destinés au niveau cellulaire dans certaines préparations favorables. Les

opportunités pour obtenir des analyses cellulaires de comportement véritablement satisfaisantes sont limitées en nombre mais leurs existences encouragent à poursuivre l'exploration en utilisant une approche cellulaire.

Dans le long terme il est clair qu'il est nécessaire d'avoir une manière de regarder l'ensemble des neurones dans la réalisation de leurs rôles variés, comme nous pourrions observer un jeu de base-ball à partir d'un ballon suspendu en l'air au-dessus du stade. Les potentiels évoqués et ondes cérébrales représentent les versions primitives de telles technologies holistiques, et plus encore des possibilités fantastiques qui hantent les imaginations de nombreux neuroscientifiques à un moment donné de leurs carrières. Un jour ces fantaisies pourront être traduites dans la réalité technologique ; jusqu'à ce moment-là, l'approche cellulaire du fonctionnement du cerveau est le meilleur jeu de balle en ville » (op. cit., notre trad.).

Ce qui surprend dans ce texte, c'est tout d'abord le contraste entre le titre du commentaire de l'article de Selverston « Réflexions neurophilosophiques sur le modèle des générateurs du système nerveux central », et le contenu de ce commentaire. Jamais, William J. Davis ne justifie l'usage du néologisme, qui, à la différence de la reprise nominale de P. S. Churchland, n'est ici qu'à titre d'adjectif aux réflexions. Ce néologisme « neurophilosophie » n'est jamais

expliqué dans sa formation ni par celui qui l'a formé, ni par celle qui l'utilise. Il paraît correspondre à un désir d'unir la science exacte et la science humaine, mais sans que l'on puisse en être certain. C'est à propos de la recherche d'un modèle théorique que l'expression de neurophilosophie apparaît dans le travail de la réflexion. Comme si la recherche d'un paradigme relevait de la philosophie, mais comme il s'agit de trouver le paradigme pour les neurosciences, le terme composé de neurophilosophie serait le reflet de cette intersection. La philosophie sert de contre-champ théorique pour le développement du champ des neurosciences. Si bien que l'on est en droit de faire au moins deux hypothèses sur l'apparition de ce néologisme :

- soit il s'agit d'une plaisanterie de scientifique qui s'essaye à la philosophie en introduisant un terme nouveau afin de situer le niveau de son commentaire ; ce qui dans le jeu des publications ne peut être totalement exclu ;
- soit les réflexions neurophilosophiques trouvent leur justification à l'intérieur même de la démarche épistémologique du neuroscientifique qui évalue la pertinence du modèle proposé par A. I. Selverston.

Nous allons développer ce second aspect car la thèse de A. I. Selverston est essentielle pour P. S. Churchland car elle pose le problème de l'organisation cohérente

et dynamique d'un réseau de neurone. Les réticences de Neurophilosophy disparaîtront en 1992 dans The Computational Brain puisque les travaux de A. I. Selverston sur le ganglion stomatogastrique du homard épineux lui ont permis de mettre à jour la totalité du réseau de 28 neurones ainsi que le rythme fonctionnel par lequel l'action musculaire s'accomplissait. La perspective neurocomputationnelle était déjà présente dans le relevé théorique du modèle central de génération qui constituait, pour A. I. Selverston, le point de départ vers le réseau. Donc entre le pessimisme de 1980 et l'article de A. I. Selverston de 1988, « A consideration of invertebrate central pattern generators as computational data bases », la théorie neurocomputationnelle se trouve projetée comme apportant les solutions au souhait d'intégralité descriptive émis par A. I. Selverston en 1980.

Cette recherche des modes de connectivité entre les neurones est reconnue comme nécessaire par W. J. Davis : il faut relier la description des unités simples qui composent le réseau avec la perspective holistique sans laquelle le fonctionnement perdrait toute cohérence dynamique. Le recours à l'analogie du jeu base-ball prouve combien dans le début des années 1980 le modèle neurocomputationnel n'est pas encore clairement désigné par les neuroscientifiques comme le moyen de théoriser le fonctionnement du cerveau. Cette relation entre analogie et modèle, courante dans l'histoire des sciences, trouve ici toute sa force à cause de la quantité d'informations à insérer dans le modèle.

Ou plutôt c'est l'analogie qui sert provisoirement de modèle théorique pour décrire la nature des relations entre les neurones et la structure du réseau. L'analogie n'est pas hissée au statut de modèle, mais les réflexions neurophilosophiques sont articulées autour de deux questions épistémologiques : « 1 / Pouvons-nous comprendre une propriété "émergente" en des termes élémentaires ? 2 / Quelle est la nature de la "preuve" que nous pouvons apporter pour nous convaincre et persuader nos collègues que nous avons véritablement atteint une compréhension non équivoque du fonctionnement d'un réseau ? »

La première question est traditionnelle en philosophie mais a trouvé depuis sa solution au sein du courant de la neurophilosophie : en effet la théorie de l'émergence, selon Francisco Varela, est apparue à partir de 1958 lorsqu'on a pu considérer « que les cerveaux opèrent de manière distribuée sur la base d'interconnexions massives, de sorte que les connexions effectives entre les ensembles de neurones se modifient en fonction du déroulement de l'expérience » [\[4\]](#). Frank Rosenblatt, en 1958 par la construction du « perceptron », a posé les principes de la reconnaissance sur la seule base de changements de connectivité entre des composants analogues à des neurones. Comme F. Varela l'indiquait déjà en 1989 la stratégie « est de construire un système cognitif à partir, non pas de symboles et de règles, mais de constituants simples qui peuvent dynamiquement être reliés les uns aux autres de manière très dense. Ici,

chaque constituant fonctionne seulement dans son environnement local de sorte que le système ne peut être actionné par un agent extérieur qui en tournerait en quelque sorte la manivelle. Mais grâce à la nature configurationnelle du système, une coopération globale en émerge spontanément lorsque les états de chaque “neurone” en cause atteignent un stade satisfaisant. Un tel système ne requiert donc pas d’unité centrale de traitement pour contrôler son fonctionnement. Ce transfert de règles locales à la cohérence globale est le cœur de ce qu’il était convenu d’appeler l’auto-organisation pendant les années de la cybernétique. Aujourd’hui, on préfère parler de propriétés émergentes ou globales, de réseaux dynamiques, ou non linéaires, de systèmes complexes, ou encore même de synergetique » [\[5\]](#).

Ainsi au moment où paraît Neurophilosophy en 1986, ce dont témoigne son chapitre 10, les travaux de T. Sejnowsky (le collaborateur de P. S. Churchland à partir de 1988 et le coauteur de The Computational Brain en 1992) de C. Rosenblatt, D. Rumelhart, G. Hinton et R. Williams développent ce modèle de propriétés émergentes et des réseaux neuronaux [\[6\]](#). Modèle que P. S. Churchland commence à inventorier dans ce chapitre 10 de Neurophilosophy sans toutefois pouvoir le concrétiser faute de connaissances pratiques des réseaux neuronaux. L’émergence reste donc une théorie du fonctionnement du cerveau au moment de la rédaction et de la publication de Neurophilosophy. W. J. Davis ne remet pas en cause les travaux de A. I.

Silverston, même si cette idée de modèle central générateur fausse sans doute la perspective de l'émergence dans la mesure où il faut prendre en compte à la fois la motricité des éléments dans le système et la complémentarité fonctionnelle de ces éléments à l'intérieur du système. A. I. Silverston lui-même, par rapport à ces affirmations de 1980, a dû constater l'échec de la recherche d'un modèle centripète en se dégageant de la référence métaphorique à « central » dans l'objet de ses travaux sur le système nerveux central. Les neurocomputations lui permettent ainsi de présenter la causalité motrice du système nerveux central mais à l'intérieur d'un réseau de détermination aussi variée que l'inhibition fonctionnelle, l'excitation fonctionnelle, la modularité et les effets électriques.

En effet, l'obstacle à la théorisation est l'impossibilité de comprendre le tout à partir de l'observation d'une seule de ses parties. C'est l'échec de l'induction dans la méthode expérimentale. L'objet neuroscientifique ne peut extraire du particulier les lois universelles, non seulement à cause de la spécificité des opérations mais aussi à cause de l'étendue des connexions pour l'accomplissement de ces opérations. L'avantage du modèle central générateur est de proposer un principe moteur à partir duquel s'organise la structure et se développe le fonctionnement. La comparaison du cerveau avec le jeu de base-ball sert de point analogique à partir duquel W. J. Davis compare la structure formelle commune. L'analogie n'est pas

l'homologie et W. J. Davis a bien conscience de la différence entre :

« Un stade plein avec des gens observant un jeu de base-ball est analogue à une collection de cellules cérébrales opérant de concert pour réaliser une fonction neurale spécifique. Dans les deux, le stade et le cerveau, les acteurs principaux du drame sont individuels, des éléments répétés (les gens et les neurones respectivement). Dans les deux cas, il y a un "but" pour l'ensemble des éléments, et ce but est en principe connu. Et dans les deux cas le but entraîne la coordination des éléments individuels dans des modèles d'activité spécifique selon un certain nombre de règles qui sont pareillement définies en principe. »

Les gens et les neurones composent respectivement le stade et le cerveau, mais le tout est supérieur à la somme des parties ; tel est le principe holistique. Seulement si l'on se contente d'une approche trop abstraite et générale, ou d'un détail appartenant à un morceau du tout, la compréhension ne pourra s'effectuer de manière pertinente. Comment peut-on faire correspondre le modèle à la nature et à l'activité des réseaux de neurones ?

On peut dès lors se demander si la neurophilosophie n'est pas seulement une étape pour parvenir, à partir de 1986, à une théorie des propriétés émergentes dans une perspective neurocomputationnelle. Cette

étape est-elle nécessaire au point de constituer maintenant la représentation d'un courant de science unifiée. La neurophilosophie n'est-elle dans la recherche de P. S. Churchland que l'élaboration théorique au moment des années 1970-1980 des données neuroscientifiques ? Il faut répondre positivement aux deux questions : d'une part, la neurophilosophie appartient à un mouvement idéologique du positivisme logique dans son intersection avec les neurosciences ; et d'autre part, le terme de neurophilosophie est un moyen en 1984 de s'éloigner définitivement de la psychologie ordinaire et du cognitivisme en choisissant nettement une méthodologie connexionniste. Comme si le mot lui-même devait traduire cette connexion entre le cerveau et l'esprit.

La seconde question de W. J. Davis, celle de la « preuve » incontestable par l'ensemble des neuroscientifiques d'une explication non équivoque d'un réseau de neurones, se pose dans la mesure où le développement des connaissances au début des années 1980 ne permettrait pas d'établir une description définitive. Ce défaut de connaissance n'interdit pas à W. J. Davis d'espérer, comme tout neurobiologiste, le temps où l'on pourra expliquer la totalité des réseaux neuronaux. La cartographie des réseaux neuronaux serait enfin la preuve que l'accumulation des constats pathologiques et cliniques par les techniques d'imagerie sont en réalité les effets d'une causalité complexe de réseaux matériels. À ce

moment-là, et selon les conclusions de l'article de W. J. Davis, il n'y aurait plus de différence entre l'analogie et le modèle puisque la réalité matérielle aurait entièrement confirmé la prédiction. La conclusion de W. J. Davis est une preuve supplémentaire du remplacement, au cours des années 1980, de l'approche cellulaire par celle du réseau. La solution du réseau permet d'apporter une réponse à la question holistique sans pour autant nier la spécificité de chaque élément. Là où W. J. Davis regrettait l'écart entre le tout et les parties, le réseau neurocomputationnel serait le modèle idéal pour tenir à la fois la spécificité fonctionnelle des éléments et la cohérence finale de l'ensemble.

Les trois arguments de la neurophilosophie américaine

Trois arguments définissent la neurophilosophie américaine : élimination-réduction-unification, représentant trois relations tendant à nier la différence entre la philosophie et les neurosciences : élimination de l'existence des états mentaux, réduction des objets traditionnels de la philosophie, l'unification de toutes les sciences par les neurosciences.

– L'élimination provient de la tentation, dans la

rencontre de la philosophie et des neurosciences, de naturaliser les concepts traditionnels de la philosophie (mais aussi ceux de la psychologie ordinaire) comme la conscience, l'intentionnalité, le désir... Les progrès méthodologiques et techniques des neurosciences légitimeraient cette élimination dans la mesure où la philosophie des facultés ne serait plus adaptée pour rendre compte de la description fonctionnelle du cerveau. Ce remplacement des facultés par les fonctions exigerait la formation d'une neurophilosophie susceptible de réinterpréter, à la lumière des neurosciences, les phénomènes traditionnellement attribués à l'esprit. L'élimination des états mentaux serait la conséquence nécessaire de l'analyse neuronale des productions du cerveau.

Pourtant en refaisant le chemin qui légitimerait cette élimination, l'histoire de la physiologie des nerfs depuis la phrénologie nous prouve combien le rapport structure-fonction n'autorise pas ce processus. Au contraire la localisation fonctionnelle n'aura été que le moyen de transformer, et non d'éliminer, la relation de la philosophie à la science des nerfs. Les physiologues du xviii^e et du xix^e siècle n'ont eu de cesse d'interroger la philosophie en s'appropriant sa méthode afin de réfléchir aux conséquences des découvertes de la science des nerfs pour la définition de l'homme. L'intention fut moins d'éliminer que de hisser la pratique des localisations au rang d'une théorie philosophique de l'esprit humain. Mais il nous

semble que la description des lésions fonctionnelles par ces physiologues ne conduit pas à l'affirmation positive de la nécessité d'éliminer la philosophie. La fragmentation des facultés en fonctions nerveuses sert d'argument aux éliminativistes pour faire du cerveau le lieu et la cause de la mémoire, du langage et de la sensibilité. Bien sûr le combat contre l'idéalisme est clairement annoncé, mais, sous le couvert du projet matérialiste, la philosophie est réduite à une analyse de la Nature. Même le courant psychophysiologique à partir de 1850 ne peut être interprété comme l'avant-garde de l'éliminativisme : la conscience, même mesurée en intensité électrique ou en procédés chimiques, continue de constituer un objet de réflexion à partir duquel la philosophie demeurerait en position de synthèse ultime.

L'élimination se définit avant tout par une volonté arbitraire de refuser toute référence à une philosophie mentaliste. Car admettre l'existence des états mentaux serait reconnaître à la philosophie un champ spécifique. Il faut comprendre l'élimination comme une interprétation de l'histoire des neurosciences plutôt que comme son résultat : la présenter comme l'effet du développement des neurosciences serait admettre que les objets philosophiques sont entièrement devenus des objets neuroscientifiques. L'élimination voudrait déduire la philosophie des neurosciences. L'état mental ne serait qu'une illusion subjective dont l'examen des conditions de production rétablirait l'objectivité neurophysiologique. Ce retour constant à

l'objectivité rend le sujet exsangue de toute dimension mentale. La lutte de l'éliminativisme contre le dualisme mentaliste et le matérialisme fonctionnaliste s'inscrit dans ce projet. En refusant de reconnaître à l'état mental le moindre degré de réalité subjective, l'élimination prive l'esprit de toute causalité spécifique.

Il s'agit de situer avec précision les conditions du débat entre le cognitivisme et le fonctionnalisme d'une part, et les thèses éliminativistes de la neurophilosophie d'autre part. La reconnaissance de l'état mental nous apparaît nécessaire mais pas comme le cognitivisme et le fonctionnalisme le conçoivent à partir de l'activité exclusive des représentations. Mais parce que l'état mental existe comme réalité produite par le cerveau : notre thèse est de lier l'activité mentale à sa condition neurobiologique sans la réduire. Il s'agit ni d'éliminer les états mentaux, ni de les réduire à un déterminisme neurobiologique. L'état mental est un état matériel produit par le cerveau.

– La réduction traditionnelle est rejetée par la neurophilosophie qui y perçoit une attitude caricaturale et facilement criticable. S'inscrivant dans une perspective interthéorique, la réduction vise à éliminer les thèses de la psychologie ordinaire en les remplaçant par des arguments constitués à partir des résultats des neurosciences. La réduction interthéorique repose dans la neurophilosophie sur la sélection des travaux des neurosciences qui permettrait de définir les conditions neurobiologiques

de la conscience, de la perception, de l'intentionnalité. Ce mode d'interprétation utilise une sélection arbitraire de ces travaux dans l'histoire des neurosciences depuis le xviii^e siècle. En reprenant l'histoire des neurosciences autour de ses trois paradigmes (l'électrophysiologie, la neurochimie, la neurobiologie du développement), les différences peuvent être établies entre les réductions méthodologiques nécessaires au travail des neurosciences et les réductions idéologiques : celles-ci sont une extension qui cristallise les résultats des travaux neuroscientifiques en proposant par là une explication universelle. Ces extensions ne sont pas seulement dangereuses par le défaut de cohérence avec les réductions méthodologiques. Le danger de ces réductions idéologiques réside dans les discours et les représentations produits par ces mouvements d'extension : ainsi la dérive neuropsychopharmacologique nous paraît transformer réellement les comportements ; même si des progrès thérapeutiques apportent des solutions particulières, c'est au prix d'une fragmentation des pathologies et d'une réduction chimique des états mentaux que se développe cette idéologie sociale.

Contre les effets sociaux des réductions idéologiques, il convient de situer les réductions méthodologiques dans leurs domaines respectifs et dans leurs implications. Ainsi l'électrophysiologie rejoint la chimie du cerveau dès lors que les derniers travaux

s'inscrivent dans l'interaction de leurs champs. Ainsi la neurobiologie du développement limite les effets de réduction idéologique : le développement intègre le temps dans l'ontogenèse, tandis que la plasticité a modifié l'opposition traditionnelle de l'inné et de l'acquis par la découverte du rôle de l'environnement. Il existe donc bien une alternative philosophique à la neurophilosophie éliminativiste : celle-ci consiste, à l'intérieur du champ de la neurobiologie du développement, en l'élaboration d'une théorie matérielle du cerveau vivant dans sa phylogenèse et son ontogenèse. Il nous semble qu'un certain nombre de travaux converge vers le partage d'une même référence : la capacité de la matière cérébrale à s'auto-organiser. Pour notre part, nous avançons l'hypothèse de l'autonomie relative afin de tenir compte à la fois des déterminations neurobiologiques et les états mentaux vécus subjectivement : le cerveau est la condition objective de la réalité spirituelle du sujet humain car il la produit.

Il existe aussi une alternative épistémologique à la réduction interthéorique qui avance l'élimination de la psychologie ordinaire comme une nécessité devant les progrès des neurosciences. Le lent travail de reconstitution de l'histoire des neurosciences tend à prouver la fécondité des réductions méthodologiques dans leurs paradigmes respectifs. La sélection partielle de certains travaux ne nous semble pas conforme à l'histoire des neurosciences.

– L'unification est un projet qui a pour but de constituer une science unique dans laquelle la philosophie et les neurosciences n'auraient plus aucune différence de champ et d'objet. Le projet de la Naturphilosophie sert de référence à ce rêve d'une science unifiée : la confusion s'effectue en liant la quête philosophique de l'unité de la Nature aux développements de la chimie, de la physiologie et de l'électrophysiologie. Comme sous-titre à Neurophilosophy, P. S. Churchland précise *Toward a Unified Science of Mind-Brain*. Le rêve d'une science unifiée, dont Neurophilosophy serait le nom, trouverait sa raison par la dissolution de toutes les sciences humaines dans les neurosciences. La neurophilosophie servirait de théorie unificatrice pour toutes les sciences nouvelles issues des neurosciences : neuropsychologie, neuropsychiatrie, neuropsychopharmacologie...

Les trois thèses de la neurophilosophie, élimination-réduction-unification, nous obligent à un travail de genèse et de critique. L'aspect cumulatif des thèses de la neurophilosophie 1, 2, 3 est essentiel : thèse 1, discours fondateur car constitutif de la réalité neurobiologique de l'être humain ; thèse 2, discours d'élimination et classification des théories scientifiques selon leur intérêt neuroscientifique ; thèse 3, discours de domination de la neurophilosophie comme la science unifiant toutes les sciences. La science unifiée résorberait ainsi les différences méthodologiques et de contenu entre les sciences humaines (comme la philosophie et la psychologie) et les sciences de la vie.

Résumons cela par un tableau :

Thèse 1 : Analyse des états neurobiologiques
argument scientifique

Thèse 2 : Réduction de la psychologie ordinaire à la neuropsychologie **argument épistémologique**

Thèse 3 : Science unifiée de l'esprit-cerveau **argument idéologique**

Comment certains philosophes ont-ils pu consentir à une telle position ? Qu'est-ce qui, dans les relations de la neurologie et de la philosophie, a pu autoriser la création d'une telle expression ?

J.-P. Changeux et la neurophilosophie française

Les recherches de J.-P. Changeux (1936) s'inscrivent dans la réduction méthodologique opérée par la chimie du cerveau, même si, comme nous l'avons montré, il utilise la méthode dégagée par l'électrophysiologie des années 1970, pour observer la communication dans le canal ionique.

J.-P. Changeux aura donc réalisé le projet d'une

neurophilosophie française en fournissant, par ses travaux sur les neurotransmetteurs, une causalité explicative immédiatement traductible en une pharmacologie thérapeutique. Plutôt qu'une philosophie moléculaire [7], si nous parlons de neurophilosophie française, c'est en raison de ce déplacement du champ de la science vers celui de la philosophie au nom d'un matérialisme réductionniste qui cache encore son nom. Aujourd'hui professeur au Collège de France, directeur de recherche à l'Institut Pasteur, président du Comité consultatif national d'éthique [8], auteur de L'homme neuronal (1983)..., J.-P. Changeux a voulu l'extension et la reconnaissance institutionnelle de la neurobiologie. Mais, derrière cette réussite, se cachent bien des questions sur la limite entre la science et la philosophie, sur la pertinence de la psychologie ordinaire à rendre compte des comportements par-delà la neuropsychologie chimique et sur la place de l'inconscient face à l'installation du nouveau paradigme du cerveau.

En 1983, J.-P. Changeux propose sa synthèse la plus connue, L'homme neuronal, en divisant en trois moments sa thèse inaugurale : « Les objets mentaux », « Le pouvoir des gènes », « Épigenèse » [9].

Tout d'abord, les objets mentaux : « L'hypothèse adoptée ici est que percept, image de mémoire et concept constituent des formes ou des états divers d'unités matérielles de représentation mentale, que nous regrouperons sous le terme général d'"objets

mentaux". » [\[10\]](#) Ainsi les objets mentaux ne sont donc pas localisables dans le cerveau, car le cerveau de l'homme construit des représentations, en l'absence de l'objet réel. Mais cela ne veut pas dire que les images mentales n'existent pas : au contraire, elles possèdent une matérialité, puisqu'elles sont reproductibles d'un sujet à l'autre ; mais cette matérialité n'est pas entendue. S'inspirant de la différence de degré empiriste entre l'impression et l'idée, J.-P. Changeux fait, de l'image et du concept, des objets de mémoire ne possédant qu'une faible composante sensorielle. L'objet mental « est identifié à l'état physique créé par l'entrée en activité (électrique et chimique), corrélée et transitoire d'une large population ou "assemblée" de neurones distribués au niveau de plusieurs aires corticales définies » [\[11\]](#).

Ainsi peut-on faire une carte d'identité de la représentation, selon son degré d'activité. Mais, dès que l'on quitte le percept primaire, qui est le résultat de la prise directe avec l'objet extérieur, par la projection d'organes des sens au niveau des cartes ou homoncles des cortex primaire et secondaire, les processus d'autonomisation se développent : les propriétés associatives des objets mentaux produisent des enchaînements et des liaisons, qui développeront d'autant plus les apprentissages, qu'il y aura isomorphisme du percept à l'objet extérieur, et comparaison positive ou négative du percept avec le concept ou une image. Les objets mentaux sont essentiels aux opérations de la conscience, bien qu'on

ne puisse la définir autrement que comme « un système de régulations en fonctionnement » [\[12\]](#) des opérations sur les objets mentaux.

Puis, le pouvoir de gènes est nécessaire pour définir le cadre et les limites des opérations sur les objets mentaux. Leur caractère mental ne délivre pas, pour autant, ces objets des logiques déterministes. En effet, la compréhension du déterminisme génétique de l'organisation cérébrale « passe par le déchiffrement de l'adn et de son expression en protéines » [\[13\]](#). Le pouvoir des gènes surdétermine l'organisation cérébrale. Ce principe d'universalité donne, sous l'autorité des gènes, une « unité du cerveau humain » au sein de l'espèce, par-delà les différences ethniques et raciales. Le cerveau serait, du point de vue de sa structure, le point d'égalité entre tous les hommes. Car, du point de vue du fonctionnement, il faut distinguer les déterminants géniques, qui se développent au cours de la maturation postnatale du cerveau, des déterminations acquises par l'interaction avec un milieu plus ou moins riche, susceptibles d'actualiser seulement ce qui est en puissance. « ... Les gènes "se nouent", s'imbriquent, s'enchaînent en s'exprimant de manière séquentielle et différentielle au cours du développement pour créer l'organisation de l'encéphale propre à l'espèce humaine. » [\[14\]](#)

Enfin l'épigenèse : J.-P. Changeux utilise l'expression « d'enveloppe génétique », pour souligner le rapport forme-matière, pour lier la nécessité génétique avec

l'utile développement des capacités induites par elle. L'ouverture à la variabilité phénotypique individuelle concerne la croissance des câbles nerveux dans les réseaux de neurones, et non la réplication de l'adn. Ainsi y a-t-il un renouvellement permanent de l'architecture moléculaire de la synapse adulte, sous influence du milieu extérieur. L'hypothèse de la stabilisation sélective de 1975 est confirmée comme l'ultime stade de redondance maximale de contacts synaptiques. Aussi « apprendre, c'est stabiliser des combinaisons synaptiques préétablies. C'est aussi éliminer les autres » [\[15\]](#). Le cerveau permet d'apprendre par sélection des informations prises dans le contenu de cette sélection. La révolution neuronale, pour autant qu'elle évite tout réductionnisme localisationniste [\[16\]](#), définit une prédestination formelle : la méthode structuraliste s'applique donc à la science neuronale.

En cela, J.-P. Changeux rejoint les thèses des neurophilosophes américains (P. S. Churchland, T. Sejnowski) : en effet, les neurocomputations développées comme modèle explicatif, dès 1986, par le courant américain du matérialisme éliminativiste, sont une solution pour tenir compte de la complexité de la raison. Le glissement épistémologique est assuré par la notion de réseaux, qui sert de soubassement matériel à la production fonctionnelle des facultés supérieures du lobe frontal. P. S. Churchland et J.-P. Changeux partagent les mêmes sources bibliographiques, puisqu'ils se fondent tous deux sur la

théorie générale de l'organisation fonctionnelle du cerveau humain en trois unités de A. Luria (1973) : « 1 / L'ensemble réticulo-frontal règle l'éveil et le tonus mental, et rend possible les formes les plus complexes d'activité consciente ; 2 / La partie postérieure du cerveau (aires sensorielles primaires et d'association) traite l'information et prépare les programmes ; 3 / Le lobe frontal dirige, planifie et règle par le truchement du langage, crée les programmes et les évalue. » [\[17\]](#) Désormais J.-P. Changeux va spécialiser son travail sur la détermination des contenus des représentations et sur ce qu'il va appeler, en juin 1992, « les neurones de la raison ».

En 1989, avec Alain Connes, J.-P. Changeux, dans *Matière à pensée*, poursuit cette réflexion autour de l'épigenèse. À l'inverse des psychologues fonctionnalistes, comme J. Fodor ou P. N. Johnson-Laird, pour qui il suffit de décrire les processus de pensée sous des formes algorithmiques, fondant ainsi une homologie mathématique entre le cerveau et l'intelligence artificielle, J.-P. Changeux refuse de séparer structure et fonction. En effet, il distingue deux sortes de darwinisme appliqué au cerveau : d'une part, un darwinisme neural de l'évolution de la connectivité de la vie embryonnaire à la petite enfance ; c'est l'évolution par épigenèse. D'autre part, un darwinisme mental qui, « à l'échelle du temps psychologique, produit des changements d'efficacité synaptique, plutôt qu'une évolution du nombre de connexions » [\[18\]](#). Le cerveau, alors adulte, donne espoir aux

neurophysiologues de modéliser, à terme, « certaines étapes du déroulement de la pensée » [\[19\]](#).

Ainsi, le modèle de l'homme neuronal ne se confondrait pas tout à fait avec celui de « l'homme machinal » [\[20\]](#). La machine resterait une métaphore ou une analogie. Et sa référence est comparative, afin d'élaborer, en contrepartie, le cerveau d'un corps humain. Bien sûr, la tentation est grande de confondre la forme avec la matière de la construction, pour autant que le déterminisme génétique structure l'expérience mondaine du cerveau à travers son corps. Mais la différence entre le neuronal et le mental serait garantie par ce darwinisme généralisé qui distingue des différences de niveaux.

Dans un article paru initialement en anglais, « Neuronal models of cognitive functions » [\[21\]](#), J.-P. Changeux et S. Dehaene situent leurs travaux à l'opposé de l'entreprise du matérialisme éliminativiste de P. S. Churchland : la référence à Neurophilosophy est explicite en note 1 de la page 50. Une telle opposition peut paraître surprenante, comparée au commentaire qui accompagne, en quatrième de couverture la parution du livre *The Computational Brain* de P. S. Churchland et T. Sejnowski, en 1992 : « Il n'y a pas d'équivalent à *The Computational Brain*, une synthèse unique du champ étendu des modèles neuronaux. Une expérience qui mérite d'être lue. » En 1989, les raisons avancées sont les suivantes : « Parallèlement, les sciences du système nerveux ont fait

des progrès considérables. Des études, portant sur des cellules nerveuses isolées et sur leurs composants moléculaires, ont enraciné sans aucune ambiguïté, les processus neuronaux dans la chimie physique, en introduisant ainsi un large éventail de contraintes, en particulier sur la dynamique de ces processus. De plus, des recherches d'anatomie comparée chez les vertébrés supérieurs et chez l'homme, portant sur la connectivité cérébrale, la biochimie et la physiologie, ont fourni des vues nouvelles sur la complexité de l'organisation du cerveau adulte et sur sa morphogenèse. Sur de telles bases, on a avancé un programme radical exactement opposé, selon lequel, en dernière instance, la science de la vie mentale serait réduite aux sciences neurales, et les doctrines de la psychologie de sens commun seraient éliminées. » [\[22\]](#)

En se référant à la complexité de l'organisation du cerveau, ils refusent le réductionnisme du mental au neural, ce qui est une nette évolution, semble-t-il, par rapport aux premières positions de 1981. Il s'agit d'articuler la frontière entre le cognitivisme et les neurosciences, sans la détruire, et de se présenter comme lien entre les disciplines. Pourtant le passage en revue des concepts psychologiques, comme l'innéité, l'apprentissage, les représentations internes, l'intentionnalité, dans la grille des contributions neuroscientifiques, introduit un déplacement de la philosophie (« le but de notre texte n'est pas, une fois de plus, de discuter ce débat philosophique », *ibid.*) à

la neuropsychologie. Tout en excluant la référence à la méthode philosophique, il s'agit, quand même de statuer sur les objets philosophiques ; même si le but est « d'offrir des modèles pertinents qui, pour la plupart, restent toujours spéculatifs » (ibid.). La reconnaissance de l'imparfaite modélisation en termes neuronaux autorise bien cette démarche spéculative, et constitue une position éthique, par rapport au réductionnisme. Mais est-elle une garantie ? La base neurologique est complexe, ce qui devrait nous garantir de la tentation réductionniste de P. S. Churchland. Mais, en ajoutant que « les données sont trop rares », ne risque-t-on pas d'infirmer la dimension complexe pour attendre des données neuroscientifiques susceptibles de réduire l'écart entre le mental et le neural ?

Pour lutter contre cette confusion fondamentale dans les niveaux d'organisation de l'implémentation neurale des modèles de fonction cognitive, J.-P. Changeux et S. Dehaene proposent une réflexion sur l'organisation. En effet, trois niveaux peuvent être distingués, là où, en 1981, J.-P. Changeux avait surtout mis l'accent sur le seul premier :

3 / Le méta-circuit

2 / Le niveau-circuit

1 / Le niveau cellulaire

Le niveau cellulaire peut être représenté par l'image traditionnelle de l'arborescence cellulaire et de ses connexions synaptiques. Cette topologie rend déjà compte de la complexité opérationnelle à l'intérieur d'un système, le système nerveux. Aucune inférence directe entre l'architecture physique et les fonctions cognitives ne peut être établie positivement ; même si un certain nombre de contraintes limite le codage des représentations mentales, les computations, les modalités de stockage et de récupération. Ainsi, la mémoire dépend du mode de stockage des traces mnésiques. Ainsi, les processus mentaux internes, comme l'attention, sont soumis aux conditions du système pour être produits. Il y a là une sorte de déterminisme systémique, où l'état et la structure des mécanismes, fussent-ils physicochimiques, imposent des voies de communication. L'architecture des fonctions cognitives, selon la métaphore, dépend d'un architecte dont l'auto-organisation est cellulaire. Sans être réductionniste, il y a une nécessité cellulaire qui s'impose dans l'architecture même des fonctions cognitives, ce qui interdit toute perspective dualiste.

Le niveau-circuit est la connection de plusieurs cellules dans la réalisation d'un comportement simple. Le connexionnisme tente, d'ailleurs, d'imiter ces circuits, sans parvenir à en reproduire la complexité. La difficulté provient, à la fois, du caractère phylogénétique,

et de la singularité ontogénétique. Mais l'idée se dégage d'une coopération de plusieurs domaines spécialisés (anatomiques, physiologiques, neuropsychologiques...), pour comprendre la perception sensorielle, le contrôle moteur, la production du langage, la mémoire... Si bien que, là aussi, la complexité des relations réciproques des circuits neuronaux, tout en définissant un autre niveau d'organisation, interdit le réductionnisme simpliste. L'interconnection, si elle ne conduit pas à l'homogénéité des processus, exige la description de chaque champ.

Le métacircuit est celui des fonctions cognitives. Mais la position de J.-P. Changeux, que l'on retrouve dans *Matière à pensée*, est la plus difficile pour nous à soutenir. En effet, reprenant la distinction kantienne de l'entendement et de la raison, il veut la justifier par « une mise en rapport avec des concepts récents issus de la psychologie expérimentale. De cette façon, les processus de l'entendement seraient, surtout, modulaires, automatiques et inconscients, bien que l'attention puisse réguler leurs inputs et leur accès aux réserves mnésiques, tandis que, pour la raison pure, le traitement serait essentiellement (quoique non exclusivement) conscient, non modulaire, et demanderait de l'attention » [23]. Ainsi à cause de leur caractère dynamique, réfléchi et anticipatoire, il affirme que « les aires corticales frontales, par exemple, contribuent aux architectures neurales de la raison » [24]. Les expériences sur les lésions cérébrales

préciseraient l'existence de deux niveaux dissociés, très proches de l'entendement et de la raison.

Alors qu'en 1981, J.-P. Changeux affirmait que le cerveau était « une machine chimique », qu'en 1983, il parlait de « machine cérébrale » [\[25\]](#), en 1989 il s'oppose à la machine sémantique de D. Dennett pour proposer l'expression de « machine intentionnelle » [\[26\]](#). D'une part, le cerveau est un système évolutif, dont la finalité n'est pas prédéterminée, mais dont le rapport structure-fonction admet une réciprocité déterminante au cours de l'épigenèse grâce à la stabilisation sélective. Et, d'autre part, il reconnaît qu'une étape élémentaire vers l'implémentation physique du sens a été réalisée ; mais, en se référant aux thèses de Intentionality : An Essay in the Philosophy of Mind (1983) de J. Searle, il reconnaît que toute tentative d'éclaircir a priori les bases neurales de l'intentionnalité est irréaliste.

En s'interdisant de présenter « une théorie neurale de l'intentionnalité » [\[27\]](#), il laisse au philosophe dualiste le soin de s'opposer au matérialiste, au nom d'une théorie a priori. Mais le projet reste virtuel, il suffirait d'utiliser les bases neurologiques présentées : « Nous nous limitons plutôt à quelques remarques et à quelques arguments qui illustrent comment une telle théorie pourrait être construite en termes de buts et de plans » (ibid.).

Comment, dès lors, ne pas situer la responsabilité des

travaux de J.-P. Changeux dans la production du néologisme hybride de neurophilosophie ? Ainsi, l'intentionnalité dépend des fonctions prospective et rétrospective du cortex préfrontal. Les intentions sont au sommet de la hiérarchie des réseaux cérébraux, elles représentent donc le niveau terminal du métacircuit. On ne fait pas de philosophie à cause de sa réflexion théorique a priori, tout en lui fournissant les véritables bases, les fondements physiologiques, anatomiques et neurologiques, sans lesquels aucune réflexion sur les sciences de l'esprit ne serait possible ! Ainsi la complexité lui permet d'éviter l'accusation de réductionnisme pur et simple ; mais cette hiérarchie de niveaux n'articule pas l'écart entre le corps et l'esprit autrement que sous le mode matérialiste, au point de confondre les facultés et les fonctions. Le paradigme reste, donc, celui d'un matérialisme mécaniste, même si la complexité systémique permet d'adoucir la reprise du modèle. Dans *Matière à pensée*, J.-P. Changeux distingue trois approches des machines à penser : l'intelligence artificielle, la modélisation du cerveau humain et de ses fonctions, et celle des machines neuro-mimétiques ; celle qui est « de loin la plus importante » [\[28\]](#), est la seconde : « Il s'agit d'un travail de recherche plus profond, qui suppose la contribution multidisciplinaire de compétences issues des mathématiques, de la physique, de la neurobiologie et de la psychologie. Cette modélisation exploite des données d'ordre anatomique et physiologique, les résultats de la biologie moléculaire et, bien entendu,

l'observation du comportement qui relève de la psychologie et de l'éthologie » (ibid.). Point de philosophie ici, et le rôle de la psychologie reste encore essentiel : J.-P. Changeux ne voit aucun obstacle théorique à la recherche des bases neurales des fonctions supérieures du cerveau ; en effet, « la méthode dite de réduction, ou mieux encore de reconstruction, que nous employons tous dans les sciences expérimentales, consiste à rechercher l'explication à un niveau inférieur à celui que l'on veut expliquer. On se fonde alors sur l'organisation, les règles d'interaction et les propriétés des éléments qui composent le niveau inférieur pour expliquer les propriétés du niveau supérieur » [29]. En expliquant le supérieur par l'inférieur, la reconstruction est un autre mot pour le réductionnisme.

Notes

[1] P. S. Churchland, *Neurophilosophy*, op. cit., p. IX, traduit en français chez puf en 1999.

[2] J. Illes, *Neuroethics in a new era of neuroimaging*, *American Journal of Neuroradiology*, 24 : 1739-1741, octobre 2003 ; D. Rees et S. Rose, *New Brain Sciences Perils and Prospects*, Cambridge University Press, 2004 ; N. Levy, *Neuroethics : Challenges for the 21st Century*, Cambridge University Press, 2007.

[3] Cité par P. S. Churchland, op. cit., p. 404, Allen I. Selverston, *Are central pattern generators*

understandable (1980).

[4] F. Varela, E. Thompson et E. Rosch, L'inscription corporelle de l'esprit. Sciences cognitives et expérience humaine, chap. 5 : « Les propriétés émergentes et le connexionnisme », Paris, Le Seuil, 1993, p. 131, F. Rosenblatt, The perceptron : A probabilistic model for information storage and organization in the brain, Psychological Review, 1959, 65, p. 368-408.

[5] Op. cit., p. 135-136 ; et F. Varela, Connaître les sciences cognitives, tendances et perspectives, chap. 4 : « Troisième étape. L'émergence : une alternance à l'orientation symbolique », Paris, Le Seuil, 1989, p. 61.

[6] W. Bechtel et A. Abrahamsen (1991), Le connexionnisme et l'esprit. Introduction au traitement parallèle par réseaux, trad. J. Proust, Paris, La Découverte, 1993.

[7] Nous utiliserons le livre de Claude Debru, Philosophie moléculaire. Monod, Wyman, Changeux, Paris, Vrin, 1987. Il montre l'articulation entre l'électrophysiologie et la neurochimie autour de la définition et de la fonction de l'acétylcholine

[8] D. Memmi, Les gardiens du corps. Dix ans de magistère bioéthique, Paris, ehess, 1996, p. 88-96.

[9] J.-P. Changeux, L'homme neuronal, Paris, Fayard, 1983, coll. « Pluriel », 1989.

[10] Ibid., p. 168.

[11] Ibid., p. 174.

[12] Ibid., p. 211.

[13] Ibid., p. 225.

[14] Ibid., p. 252.

[15] Ibid., p. 304.

[16] « Nous savons qu'il n'est pas possible d'assigner une fonction cérébrale intégrée à un "centre" unique ou à un seul neurotransmetteur, mais à un système d' "étapes de transit" ou se "nouent" des états d'activité électrique et chimique » (ibid., p. 252).

[17] Changeux, Molécule et mémoire, Gourdon, Éd. Bedou, p. 126.

[18] Changeux et Connes, Matière à pensée, Paris, O. Jacob, 1990, p. 153.

[19] Ibid., p. 154.

[20] Selon l'expression de A. Green, L'homme machinal, Le temps de la réflexion, Paris, Gallimard, 1983, p. 345-369.

[21] Changeux et Dehaene, Neural models of cognitive functions, publié dans Cognition, 33, 1989, p. 63-109, et traduit en français par Michel Dupuis, sous le titre Modèles neuronaux des fonctions cognitives, Philosophie de l'esprit et sciences du cerveau, Paris, Vrin, 1991, p. 49-69.

[22] Ibid., p. 50.

[23] Ibid., p. 55.

[24] Ibid., p. 58.

[25] Ibid., p. 212.

[26] J.-P. Changeux et A. Connes, La matière à pensée, op. cit., p. 69.

[27] Ibid., p. 62.

[28] Ibid., p. 204.

[29] Ibid., p. 206.

Chapitre II

Élimination des états mentaux

Le potentiel évoqué, définition du seuil de conscience

Après la découverte de Richard Caton (1875), l'invention de eeg à travers l'enregistrement, en 1928, des rythmes alpha par H. Berger (1873-1941), la démonstration, en 1933, par Edgar Adrian (1889-1977) et Brian Matthews (1933) a permis de mettre au point un oscilloscope : au lieu que le signal venant du nerf déplace un très fin ressort, c'est la position d'un faisceau d'électrons de masse quasi nulle qui, dans l'oscilloscope, sert d'indicateur de tension. Chaque cellule possède un potentiel d'énergie constant. Ce potentiel, appelé potentiel de membrane, est de l'ordre du dixième de volt (90 millivolts). L'intérieur de la cellule est négatif par rapport à l'extérieur. Si la perméabilité de la membrane change, le potentiel de membrane

change rapidement. Quand la cellule reçoit un signal électrique qui l'excite, les ions sodium se précipitent à l'intérieur et les ions potassium à l'extérieur de la cellule. Ce rapide changement de l'équilibre ionique est la source du potentiel d'action. Comme nous le verrons dans la seconde partie, la technique d'enregistrement intracellulaire (1953) permet à des scientifiques comme Alan Hodgkin, Andrew Fiedling Huxley (1917-) et Bernard Katz (1911-) de démontrer que le potentiel d'action a pour cause le flux rapide d'ions élémentaires à travers la membrane semi-perméable de la cellule.

Les potentiels évoqués sont obtenus par la mise en série de données d'électro-encéphalogramme (eeg). Les potentiels évoqués sont la sommation des potentiels d'action par stimulation sensorielle de la vue, de l'audition, du toucher. Georges D. Dawson, dans les années 1940, enregistra par la pose d'électrodes sur le scalp les effets d'amplitude évoqués par la stimulation électrique du cortex. Cette technique ne constitue pour Churchland qu'une variante de l'électro-encéphalographie. Pourtant, ces résultats sont le moyen de critiquer définitivement Libet qui avaient été dans les années précédant Neurophilosophy, l'adversaire désigné. En effet, comme nous l'avons montré, Churchland s'appuyait sur un temps de réponse de 358 ms contre la position de Libet qui soutenait la base de 500 ms comme un minimum pour le temps de réaction de la conscience. Churchland voulait établir une différence entre 500 ms

(temps d'intégration sensori-motrice de la stimulation du cortex à la sensation consciente) à 358,33 ms (temps de stimulation de la main à la sensation consciente). Elle préférait cette seconde méthode pour abaisser le seuil de la conscience. S'appuyant sur de récents travaux de E. Donchin « Event-related potentials in the study of consciousness » (Donchin, McCarthy, Kutas et Ritter, 1983), qui depuis 1975 cherchent à dégager les modes de reconnaissance du cerveau par l'électrophysiologie des potentiels, elle démontre la flexibilité très grande de l'évaluation neurocognitive : celle-ci dépend de la puissance de la stimulation électrique, du temps de réaction et de la localisation de la stimulation.

Les méthodes électrophysiologiques ont été et restent les méthodes de choix pour l'analyse des signaux nerveux à des échelles allant de la milliseconde à la seconde. À partir de 1950, la généralisation de l'utilisation des micro-électrodes intracellulaires a été décisive dans la compréhension des potentiels d'action, des potentiels synaptiques et des potentiels de récepteurs. Churchland s'en tient ici à l'examen des potentiels d'action pour souligner le rôle sensori-moteur des fonctions cérébrales : sous l'influence d'un stimulus électrique galvanique appliqué d'un muscle à un nerf ou sur le point moteur d'un muscle en repos, on obtient une réponse musculaire dont on peut apprécier la forme, l'amplitude et la durée. Le but de cet examen est de rendre compte de la réactivité du complexe neuromusculaire. Churchland y voit le moyen de relier

les neurosciences et la psychologie : « En admettant que la justification méthodologique de la recherche des potentiels évoqués est de franchir l'intervalle entre le niveau microscopique des neurosciences et le niveau macroscopique de la psychologie, on doit aussi la destiner pour les données de la matière du substrat neuronal. » Churchland n'ignore pas les autres potentiels dans la mesure où ils fournissent les termes de chaque étape distincte dans le processus d'information de l'ensemble neuronal. Le but est l'approfondissement de la méthode analytique afin de décomposer au maximum et de situer dans chaque niveau la hiérarchie des causes.

Méthode de stimulation cérébrale, cette technique permet de faire surgir des convergences neurocognitives par l'observation des résultats. L'objectif est de morceler l'observation des états neurobiologiques afin de donner de manière détaillée le degré, la nature et la fréquence de l'activité étudiée. Le potentiel évoqué repose sur la détection de la réponse évoquée au niveau du cortex cérébral chez l'homme par la stimulation d'un récepteur sensoriel ou d'un nerf périphérique à l'aide d'électrodes placées à la surface du scalp. Churchland relève les travaux de S. A. Hillyard sur l'attention dirigée pour préciser le seuil entre l'insensibilité et la sensibilité, ceux de C. C. Duncan-Johnson et E. Donchin sur l'attention sélective, et ceux de Sutton sur la perception spatiale. Dans tous les cas il s'agit de faire la démonstration que la conscience est liée à l'activité électrique du cerveau,

puisque les potentiels peuvent évoquer l'attention et la perception qui sont les deux caractéristiques de la définition neurobiologique de la conscience. Par ces descriptions microfonctionnelles, la thèse 1 se trouverait renforcée puisque aucun état mental n'a de consistance suffisante pour résister à la méthode de décomposition analytique mis en œuvre par l'électrophysiologie.

Les états neurobiologiques sont mis en valeur par les réponses aux stimulations : l'idée d'évocation de potentiel renforce la critique de la théorie selon laquelle la conscience serait pleinement actuelle par le travail de la pensée. Contre ce dualisme, la thèse 1 de la neurophilosophie élimine le concept de conscience car il devient inutile à partir du moment où les données électrophysiologiques apportent la preuve d'une activité potentielle : il y a différents niveaux dans l'organisation cérébrale si bien que ce que la psychologie ordinaire désigne sous le vocable de conscience n'en est qu'un niveau. Le critère de définition de la conscience devient l'attention à une stimulation visuelle ou tactile : plutôt qu'une réflexion qui développe une conscience de soi par la référence abstraite à un cogito, il conviendrait, selon Churchland, d'apercevoir les limites de la connaissance de la sensation à travers des seuils quantitatifs de stimulations électriques. Plutôt qu'une référence au sensualisme, l'analyse de réponses aux stimulations fournit une définition neurologique de la sensibilité qui ne met pas en jeu le travail corporel de réceptivité. En allant en deçà de nos sens, les

potentiels évoqués ont pour but de manifester des capacités neurocognitives endogènes. La conscience n'est plus une unité transcendante aux états neurobiologiques qui les maîtriserait par la volonté. Au contraire, la déliquescence de la conscience trouverait sa cause interne dans la neurobiologie des nerfs sans lesquels la possibilité d'un retour réflexif sur soi-même serait interdit. Il n'y aurait donc pas de caractère spécifique à la conscience qui déterminerait sa qualité ; en mesurant le seuil de la conscience par la sensibilité aux stimulations nerveuses, Churchland voudrait détruire tout recours à une expérience subjective indépendante. Le cerveau et ses activités produisent des états neurobiologiques dont la composition et la complexité d'organisation fourniraient les informations suffisantes à la sensibilité nerveuse. Churchland opère une réduction de la conscience à une réaction de la sensibilité à une stimulation. Le potentiel suppose une actualisation par stimulation et indique une permanence et une constante disponibilité neurocognitive.

L'imagerie rapide du débit sanguin cérébral par tomographie d'émission de positrons (tep) permet de mettre en évidence les réseaux de neurones impliqués dans l'exécution de certaines fonctions cognitives chez l'homme. Churchland rend compte de la comparaison entre un cerveau normal et un cerveau pathologique. En effet, l'utilisation de l'imagerie cérébrale s'est développée pour pallier au défaut d'une observation clinique externe du cerveau qui ne pouvait plus décider

de la nature et de l'étendue de la pathologie. La comparaison entre les cerveaux « normaux » et les cerveaux de psychotiques chroniques schizophréniques est déjà à l'œuvre depuis 1982 à travers les travaux de D. H. Ingvar « Mental illness and regional brain metabolism » (1982) : la distribution du sang, différente dans le cerveau des patients schizophrènes, est moins importante dans les régions frontales et plus élevée dans les régions postcentrales. La neuropsychiatrie trouve là une matérialisation du déficit mental dans la circulation et la distribution sanguine qui expliquerait, sinon la cause, du moins l'effet de la différence de performance cognitive. Une telle perspective scientifique, si elle nous paraît pertinente du point de vue du constat neuropathologique, relève tout de même d'une idéologie de sélection sociale par laquelle on veut parvenir à délimiter les effets sociaux et comportementaux de la dégradation neurologique.

Le temps de réaction de la conscience n'est pas une preuve de son existence

À partir de 1981, P. S. Churchland s'attaque aux travaux de Benjamin Libet, neurophysiologue, qui étudie les états de conscience par rapport au temps. P. S. Churchland continue son travail critique contre J.

Eccles à travers B. Libet, car ils défendent la thèse du manque d'identité entre les états cérébraux et les états mentaux. L'examen critique des données neurophysiologiques de B. Libet doit révéler l'insuffisance de la conclusion dualiste.

Les travaux de B. Libet se fondent sur des études électrobiologiques afin de mesurer la prise en compte consciente d'un stimulus [\[1\]](#). Le moyen de cette mesure est le potentiel évoqué (pe) qui est une réponse bioélectrique à phases multiples, les unes précoces et les autres plus tardives à l'application d'un stimulus adéquat, cutané, visuel ou auditif. L'enjeu du débat avec P. S. Churchland est la quantité de la réaction bioélectrique qui indiquerait le degré de conscience. L'hypothèse d'une intégration sensori-motrice infraconsciente ou préconsciente vient supposer des opérations sensori-motrices antérieures à la conscience qu'en a le sujet. Ce qui interdit la thèse de l'identité des états cérébraux et des états mentaux. Cet écart entre la stimulation bioélectrique et la sensation consciente subjective est le principe d'une non-réduction de la conscience et du cerveau. B. Libet sollicite ses patients à partir de deux stimulations différentes : la sensation de picotement dans la main stimulée de l'extérieur et la sensation stimulée de l'intérieur du cerveau lui-même. Or, le temps des expériences conscientes ne serait pas le même que celui des stimulations cérébrales.

Comme le précise D. Dennett, « Libet a comparé le

temps mis par des pincements induits sur le cortex à celui de sensations produites de façon usuelle, en appliquant une brève impulsion électrique à la main elle-même... À partir des données recueillies, il a soutenu qu'alors que dans chaque cas cela prenait un temps considérable (environ 500 ms) du départ de la stimulation jusqu'à l' "adéquation neuronale" (le point où selon lui les processus corticaux culminent pour produire une expérience consciente d'un pincement), quand la main elle-même était stimulée, l'expérience était "automatiquement" » renvoyée à rebours dans le temps « et était ressentie comme arrivant avant le pincement produit par la stimulation cérébrale elle-même » [2]. P. S. Churchland ne peut que s'attaquer à l'hypothèse d'une intégration antérieure à la conscience, car elle refuse à la conscience une activité spécifique qui ne dépend pas directement d'une causalité neurophysiologique.

Selon P. S. Churchland, la thèse de B. Libet se déroule selon quatre affirmations :

1. « Il y a un déplacement temporel d'une sensation tel qu'il apparaît plus tôt qu'il ne devrait dans l'ordre des événements expérimentés.
2. « La sensation épidermique est ressentie plus tôt que les états cérébraux nécessaires pour sa production.

3. « Le temps subjectif est différent du temps physique.
4. « Les états mentaux ne peuvent être identiques aux états physiques. » [\[3\]](#)

La première critique faite par P. S. Churchland consiste à attirer l'attention sur la méthodologie de B. Libet : « Le temps de l'expérimentation des sensations n'est pas établi sur la base d'une réponse immédiate du sujet telle que la pression d'un bouton ou le fait de dire go aussitôt après les avoir ressentis. » Lorsque le sujet répond au stimulus périphérique sur le mode d'un temps de réaction, il va plus vite selon B. Libet que lorsqu'il doit attendre l'intégration sensori-motrice après un ébranlement prolongé du système nerveux. P. S. Churchland voit dans cette comparaison un manque de rigueur méthodologique dans la mesure où ce délai favorise l'écart entre le tri terminal des sensations dans l'intégration et le report mécanique de la stimulation dans le temps de réaction. P. S. Churchland « a essayé de discréditer la première thèse de Libet, la durée longue requise pour atteindre l'« adéquation neuronale » pour la conscience, en demandant aux sujets d'une expérience de dire go dès qu'ils seraient conscients de la présence d'un stimulus sur leur peau comparable à ceux utilisés par Libet. Elle a relevé un temps moyen de 358 ms pour neuf sujets, ce qui, selon elle, montrait que les sujets doivent avoir atteint l'adéquation neuronale vers 200 ms au plus tard » [\[4\]](#).

L'objectif de P. S. Churchland est de détruire l'argument dualiste que B. Libet tire du constat d'un temps de latence entre la stimulation du cortex et la sensation épidermique. Selon B. Libet le temps de l'intégration sensori-motrice est de 200 à 500 ms. Pendant ce temps l'activité neurophysiologique s'exerce indépendamment de la conscience. La contestation de cette donnée quantitative est essentielle dans la stratégie critique de P. S. Churchland ; se référant à d'autres travaux qui tiennent compte de la réponse réflexe des sujets, elle trouve un seuil de réceptivité sensible beaucoup plus bas : la rétroaction ou réaction serait de l'ordre de 50-100 ms.

Cette différence entre 500 ms (temps d'intégration sensori-motrice de la stimulation du cortex à la sensation consciente) à 358,33 ms (temps de la stimulation de la main à la sensation consciente) pose le problème de la définition de la conscience. Car la sensation épidermique précède la sensation consciente par stimulation du cortex. S'il en était ainsi, il faudrait que la sensation épidermique soit ressentie plus tôt que les états cérébraux nécessaires à sa production. Ce qui est paradoxal pour P. S. Churchland, qui ne distingue pas la phénoménologie électrobiologique de la phénoménologie de la conscience, contrairement à B. Libet. Pour elle la sensation consciente ne peut apparaître a posteriori après une latence de l'activité cérébrale. Car la stimulation électrique soit du cortex, soit de la main modifie seulement le temps d'évaluation de la

sensation consciente, mais ne transforme pas la nature de la conscience. Tout dépend du lieu de stimulation dans l'activité nerveuse qui produit la conscience.

Selon la neurophilosophie le travail sur les potentiels évoqués ne devrait pas faire prendre la notion de potentialité pour une notion temporelle. Le potentiel évoqué précède la conscience comme sa cause, même si son effet est ressenti après coup dans la sensation. Mais la conscience est déjà là dans le potentiel évoqué. L'utilisation du potentiel évoqué vise à éliminer toute différence temporelle pour souligner la simultanéité de l'enchaînement logique. Seuls les dualistes comme B. Libet ou J. Eccles, qui se réfèrent aux processus d'un esprit non physique, ont intérêt à séparer l'ordre des sensations (critère de la conscience) de l'ordre des stimulations. Qu'ils s'agissent de la perception, de l'anticipation ou de la durée, les capacités temporelles du cerveau joueraient un rôle déterminant dans la constitution de ces états.

Le dualisme invoque des mécanismes non physiques à cause du défaut d'explication des neurosciences : le dualisme n'aurait de pertinence que durant ce temps d'échec des neurosciences : « Si dans notre ignorance et notre frustration nous invoquons des mécanismes non physiques pour expliquer ce que nous ne pouvons pas encore expliquer physiquement, nous pourrions satisfaire notre curiosité. » [\[5\]](#) Le refus du dualisme n'est pas dogmatique. La neurophilosophie reconnaît

que le dualisme représente une des meilleures théories pour lier les neurosciences et la psychologie, en l'absence d'explication exhaustive. La thèse de B. Libet est donc un partenaire fiable, mais qu'il faut invalider.

La réponse de B. Libet s'organise autour de la thèse selon laquelle « la distinction entre les phénomènes mentaux et physiques est bien sûr tout à fait compatible avec la théorie d'une correspondance entre les événements mentaux et physiques du cerveau » [6]. La différence temporelle qu'il mesure lui semble être un argument suffisant pour créer des difficultés à la théorie de l'identité.

Quatre arguments sont proposés :

1. Concernant l'erreur de méthode selon laquelle B. Libet n'aurait pas retenu la réaction verbale du sujet à l'impulsion, il répond que le sujet peut très bien être préconscient ou inconscient de l'effet avant qu'il ne se produise réellement dans la sensation exprimée. Si l'on s'en tenait à la seule mesure verbale, on ne pourrait différencier le moment de l'impulsion du stimulus dans le cortex avec son effet ressenti dans l'expression.
2. Le délai entre le processus cérébral et le stimulus tactile est de 500 ms, ce qui est remis en cause par P. S. Churchland. Or, précise B.

Libet, qu'il y a bien une modification de ce délai selon la localisation du stimulus, mais cela ne le fait pas disparaître pour autant car il reste entre 300 et 600 ms. Le délai de la rétroaction du stimulus tactile sur le processus cérébral ne peut non plus être imputé à une interférence entre les deux.

3. Il faut bien reconnaître, malgré les critiques de P. S. Churchland sur le caractère un peu rigide de la valeur du délai de 500 ms, qu'il y a un délai cérébral après une durée minimum de stimulus cortical. Quelle que soit la localisation du stimulus, ce délai est établi par les expériences.
4. Les solutions alternatives proposées par P. S. Churchland font appel à un mécanisme mystérieux qui relierait de manière plus rapide les deux pôles par le moyen des processus neurophysiologiques. Le désir d'abandonner le délai de la substance cérébrale provient de la volonté de confondre le mental et le cérébral, et de n'accorder aux processus de conscience sensori-motrice que la dimension secondaire par rapport au processus cérébral.

Après avoir rappelé ses premiers travaux, B. Libet en vient à démontrer que le délai neuronal existe pour l'intention consciente : en effet, s'il est vrai que le système encéphalique est nécessaire pour produire un

état conscient, il n'en est pas la condition suffisante. Les changements de durée du délai neuronal pour traiter une activité qualifient les modes d'apparition de la conscience.

Ces changements déterminent l'activité inconsciente ou consciente du sujet : « 1 / Une durée minimum de l'activité neuronale appropriée, dont les conditions peuvent aller jusqu'à 500 ms, est requise pour faire apparaître une expérience consciente ou de prise de conscience d'un événement... 2 / Quand l'activité neuronale a une durée plus brève que ce qui est requis pour la prise de conscience, il peut encore s'interposer une fonction mentale inconsciente, sans aucune prise de conscience subjective de celui-ci. » [\[7\]](#) L'existence d'une activité inconsciente à travers l'activité sensorimotrice et les perceptions tend à prouver que la conscience n'est pas le seul critère de mesure. B. Libet a prouvé qu'il faut un certain facteur temps de la substance neuronale pour que la conscience se produise de l'ordre de 500 ms. Mais ce n'est pas pour réduire toute l'activité de l'esprit/cerveau à ces critères : au contraire ses recherches l'amènent à étudier ce qui se passe pendant ce délai de production, qui contient les fonctions mentales inconscientes.

Conscience et cerveau : une synthèse

interactionniste

John Eccles (1903-1997), prix Nobel de médecine 1964, a développé une conception originale dont K. Popper est largement redevable. Auteur en 1970 de *Facing Reality : Philosophical Adventures by a Brain Scientist*, il présente en 1989 une synthèse de ces thèses dans son ouvrage *Évolution du cerveau et création de la conscience*. À la recherche de la vraie nature de l'homme [8]. Il a une position très critique envers la neurophilosophie ; il estimait, à propos de la conscience, dans son dernier livre en 1994 : « La philosophie et la neurophilosophie contemporaines ne sont plus, à ce stade, qu'une cacophonie matérialiste composée de structuralisme et de fonctionnalisme et qui culmine dans le robotisme, avec L'Apprenti sorcier en musique de fond. » [9] Après avoir rendu compte de l'évolution biologique, de l'évolution humaine, de l'évolution du cerveau humain, des théories de l'apprentissage... on pourrait s'attendre à une défense d'un matérialisme évolutif à la manière de G. M. Edelman. Or, s'il reconnaît que les « événements mentaux sont nécessairement liés au continuum des processus neurobiologiques du cerveau » [10], il refuse de donner au statut ontologique de la conscience une définition purement moniste. La théorie darwinienne de l'évolution n'aurait pas répondu selon J. Eccles à la formation et au développement des expériences mentales non matérielles. Il déclare : « Je pense que

l'apparition de la conscience est un peu "cadavre dans le placard" de l'orthodoxie évolutionniste... aussi longtemps qu'on la considère comme un processus exclusivement naturel au sein d'un monde exclusivement matériel. » [11] Reprenant la tripartition de K. Popper, il critique le matérialisme absolu comme celui qui ne se réfère qu'au Monde 1, c'est-à-dire à l'ensemble du monde matériel, y compris les cerveaux. Dans le matérialisme absolu, le Monde 1_m, qui est la petite partie du monde matériel qui est associée à des états mentaux, est éliminé ; il en est de même pour le Monde 2 qui est l'ensemble des expériences subjectives ou mentales. Seuls des « matérialistes intransigeants » refusent de reconnaître l'existence de la conscience dans les événements mentaux. De fait pour J. Eccles : « Le matérialisme absolu s'élimine de lui-même ! » Les théories matérialistes cherchent à préserver le caractère fermé du Monde 1, même si le panpsychisme, l'épiphénoménisme, la théorie de l'identité (défendue par H. Feigl) tentent de proposer des liens avec la vie mentale de la conscience.

Face à cela J. Eccles défend une thèse du dualisme interactionniste. Il le définit ainsi :

Monde 1 = Monde 1_p + Monde 1_{ci} : Monde 1_m = Monde 2

Le Monde 1 : ensemble du monde matériel, y compris les cerveaux.

Le Monde 1_p : est ce qui, dans le monde matériel, est dépourvu d'états mentaux.

Le Monde 1_{cl} : est l'interaction dans le Cerveau de Liaison (cl) avec le Monde 2.

Le Monde 1_m : est la petite partie du monde matériel qui est associée à des états mentaux grâce au Monde 1_{cl} .

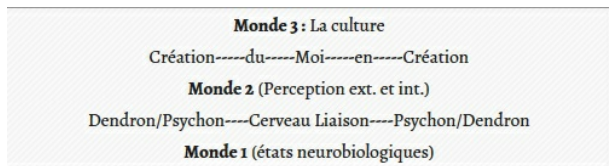
Le Monde 2 : ensemble des expériences subjectives ou mentales.

Par l'hypothèse du cerveau de liaison, J. Eccles établit le lien entre le Monde 1 et le Monde 2. Pour cela, dès 1986, il fait l'hypothèse des microsites : selon celle-ci, « le réseau vésiculaire présynaptique donne l'occasion – offre à l'intention mentale une chance – de sélectionner, de choisir qu'une des vésicules d'un bouton libère une exocytose. Le même phénomène se produirait pour tout l'ensemble des synapses dendritiques qui sont activées en cet instant et dont le nombre atteint probablement plusieurs milliers... » [\[12\]](#). Même si J. Eccles prend la précaution de l'italique, la finalité interne à l'organisation de la matière ressort comme le moyen de l'interaction entre le mental et le neuronal. L'interaction est justement située dans les microsites présynaptiques.

L'unité concernée par cette interaction représente trois dendrons (unités neurales) et chacun est pénétré par une unité mentale ou « psychon ». Comme l'indique J. Eccles lui-même : « L'hypothèse possède à la fois une grande simplicité et une grande puissance d'explication. Elle diffère de la théorie de l'identité en ce qu'elle accorde une entière autonomie au Monde 2. Une intention mentale agissant par l'intermédiaire d'un psychon, par exemple celui des carrés pleins, dispose sur son dendron de dizaine de milliers de réseaux vésiculaires présynaptiques déjà activés, dont les vésicules n'attendent plus d'être choisies. » [\[13\]](#) Sans aboutir à une localisation neurophysiologique, les microsites assurent aux événements mentaux un taux de production et d'efficacité comparable aux champs de probabilité de la mécanique quantique. J. Eccles instaure une analogie entre la mécanique quantique et la neurophysiologie afin de donner aux états mentaux une probabilité d'existence sans la localisation réductionniste. Comme si les neurosciences devaient accomplir le même saut qualitatif que la physique moderne.

Pourtant si cette interaction problématique dans le champ d'un dendron permettait au psychon d'accomplir matériellement son intention mentale, Eccles maintient le concept de création du moi et réserve seulement l'évolution au cerveau. S'il est vrai qu'au plan de l'ontogenèse J. Eccles reconnaît la pertinence de la thèse de l'émergence de la conscience de soi dans la construction du cerveau et

de ses rapports avec le monde extérieur, l'unité vécue du moi ne provient pas de la complémentarité hémisphérique du cerveau. Aussi le néocortex est associé à une large variété de fonctions cognitives : la conscience et la conscience de soi, la réflexion, la mémoire, les sentiments, l'imagination et les facultés créatrices. Mais, dans le Monde 2 (composé d'éléments provenant de la perception extérieure (lumière, couleurs, sons, odeurs, goût, douleur, toucher), et d'éléments provenant de la perception intérieure (pensée, sentiments, souvenirs, rêves, imagination, intentions) et de l'entité appelée le Moi, le Soi ou l'Âme) les relations avec le cerveau de liaison (cl) ont besoin de cette unité du moi pour rassembler l'identité du Soi. Le terme de création de la conscience ne se trouve pas justifier autrement : il vient indiquer que le Monde 2 produit dans son autonomie une zone unitaire et identitaire qui est le Moi. Celui-ci produira le Monde 3 de manière entièrement indépendante du Monde 2.



Nous proposons ce schéma comme synthèse de la position de J. Eccles, qui reste un dualisme malgré son interactionnisme. P. S. Churchland critique cette

position dans laquelle elle reconnaît la conservation d'un vocabulaire métaphysique. La théorie de Eccles reste pour elle « métaphorique » [\[14\]](#) ; et son hypothèse du « psychon » n'est pas sans nous rappeler, indique-t-elle, la recherche de R. Descartes de la glande pinéale.

Le cerveau divise la conscience

Dans ce contexte, P. S. Churchland veut préciser la spécificité des fonctions neurocognitives en ayant recours aux travaux sur le split-brain (littéralement : « cerveau coupé en deux »), c'est-à-dire sur la latéralisation hémisphérique des fonctions. En 1865, comme nous l'avons montré, P. Broca donna la première preuve d'une asymétrie des fonctions entre les deux hémisphères cérébraux en localisant la perte de la parole dans une lésion de l'hémisphère gauche. Si bien que pendant très longtemps on a cru que l'hémisphère gauche était « dominant » et que l'hémisphère droit n'était que « mineur ». P. S. Churchland voit là un risque de hiérarchiser les hémisphères en privilégiant celui de gauche afin d'en faire le centre des activités cognitives. Ce qui serait pour elle un retour à une position tout aussi dualiste que celle de R. Descartes : en effet ce serait accorder à un seul hémisphère une prépondérance causale, alors

que les études sur les différenciations fonctionnelles tendraient à prouver que l'unité de la conscience n'existe pas. Car s'appuyant sur ces travaux, P. S. Churchland veut réfuter l'idée que l'addition des deux hémisphères constituerait un tout dont l'activité serait la conscience (tout qui serait même supérieur à la somme des parties) ; et plus encore elle ne veut pas attribuer à un seul hémisphère les fonctions spécifiques du langage, de la perception, de la lecture visuelle et de la mémoire... car le centre cognitif serait reconstitué en un seul hémisphère.

Au cours des années 1940, William van Wagenen et R. Yorke Herren démontrèrent cette technique chirurgicale : la communication entre les hémisphères est assurée par le corps calleux. « Ce moment est provenu d'observations effectuées sur les effets d'interventions chirurgicales consistant à supprimer le corps calleux – c'est-à-dire une masse de fibres nerveuses reliant les deux hémisphères – chez les patients atteints d'épilepsie. » [\[15\]](#) En effet l'épilepsie se prête particulièrement à cette dissection dans la mesure où une crise cérébrale provoque une perturbation des fonctions. Ainsi le point de départ des études sur le split-brain trouve dans l'épilepsie une approche clinique des dysfonctionnements cérébraux. Il s'agit moins de décider en faveur de la seule localisation intégrale et réductive de la crise d'épilepsie dans l'hémisphère gauche, que de comparer les performances relevées à l'occasion de cette séparation des hémisphères. Dans les années 1950 Ronald

Myers et Roger Sperry (1913-1994) démontrèrent leur propre méthode : ils ont enseigné à un hémisphère d'un cerveau de chat certaines tâches pendant que l'autre hémisphère l'ignore ; l'expérience permit de « montrer que le sectionnement chez le chat de la ligne médiane du chiasma optique, du corps calleux et de la commissure antérieure rend l'animal incapable de fournir, à l'aide de son œil non entraîné, une discrimination visuelle acquise au moyen de l'œil opposé. Ainsi un chat ayant un cache sur un œil est entraîné et surentraîné à discriminer entre deux figures ; il se révèle ensuite incapable d'accomplir la même discrimination à l'aide de l'œil obturé en cours d'apprentissage. En effet, la seconde moitié du cerveau doit apprendre la discrimination depuis le début. La comparaison entre les courbes d'apprentissage des deux hémisphères montre qu'elles sont à peu près identiques » [\[16\]](#). La déconnexion d'une structure cérébrale, réalisée chirurgicalement, remet en cause la théorie de la communication des propriétés « mentales » entre les deux hémisphères : en coupant le chiasme optique, par lequel toutes les informations de l'œil droit sont dirigées vers l'hémisphère gauche, les chercheurs sollicitent la réponse d'un seul hémisphère à une stimulation visuelle.

R. W. Sperry développe la thèse du mentalisme. La thèse de l'interaction esprit-cerveau lui semble favoriser une causalité neuronale des phénomènes mentaux qui posséderaient une trop grande réalité cognitive par rapport à la dépendance observée dans

les cas de split-brain. Le mentalisme est un monisme matérialiste par lequel R. W. Sperry refuse la thèse d'une séparation des fonctions mentales des matériaux neurophysiologiques. La neurophilosophie américaine trouve dans ces tests de perception simultanée de visages par les deux hémisphères cérébraux les résultats d'un abandon définitif de l'existence des états mentaux séparés.

L'expression de « mentalisme », propre à R. W. Sperry, est pourtant en contradiction avec le projet éliminativiste des états mentaux. R. W. Sperry n'attribue pas à l'hémisphère droit un fonctionnement automatique qui le priverait de toute conscience. Le comportement émotionnel du sujet, lorsqu'on propose à l'hémisphère droit des photographies chargées d'affects, est évalué de manière identique avec les résultats obtenus avec l'hémisphère gauche.

Si bien que R. W. Sperry avance sa thèse que deux consciences coexistent chez le même individu. Cette double conscience est le moyen d'éviter un dualisme des hémisphères qui accorderait à l'hémisphère gauche une supériorité dans l'organisation cognitive à cause de ses activités dominantes comme celle du langage. Selon Dominique Laplane, « la conscience n'est pas un épiphénomène acausal, elle n'est pas la face intérieure (internal aspect) de l'activité cérébrale, elle ne s'identifie pas aux événements neuronaux, elle n'est pas non plus un pseudo-problème injecté dans nos esprits par des gymnastiques sémantiques et

destinées à disparaître grâce à une approche linguistique appropriée. La conscience n'est plus interprétée comme un corrélat simplement passif et parallèle des événements neuronaux, leur face passive ou leur sous-produits. Elle est un déterminant actif, causal, essentiel du contrôle cérébral normal » [\[17\]](#). P. S. Churchland ne peut être en accord avec ces conceptions philosophiques de R. W. Sperry qui défend l'idée que chaque hémisphère du cerveau possède un esprit conscient qui organise la direction et le contrôle du comportement. Comme déterminant actif et causal du contrôle cérébral normal, la conscience serait une propriété émergente de chaque hémisphère. Refusant d'une part le matérialisme éliminativiste et d'autre part l'idéalisme du dualisme, R. W. Sperry défend la thèse que la conscience est le principe organisateur de la pensée cognitive. P. S. Churchland, encore une fois, cite les travaux de R. W. Sperry sur l'asymétrie hémisphérique fonctionnelle sans adhérer pour autant à sa position philosophique : ainsi elle ne peut admettre que la conscience soit un principe organisateur des états neurobiologiques de chaque hémisphère. La neurophilosophie américaine cherche au contraire à éliminer le concept de conscience afin de ne décrire les états neurobiologiques qu'à partir de la complémentarité hémisphérique.

L'unité de la conscience

Le dualisme ne serait donc pas la bonne réponse car il attribue le centre de la conscience au seul esprit. Au contraire les patients du split-brain « ont deux quelque chose là où nous en avons seulement une seule. Mais deux quoi ? Deux esprits peut-être, ou deux âmes, ou deux soi, ou deux personnes, deux centres de la conscience, deux centres de la cognition, deux centres de contrôle, deux volontés... » [18]. Limité par son adhésion aux thèses cartésiennes, le dualiste défend l'idée que l'esprit est une substance non physique. Il refuse, selon, P. S. Churchland, d'admettre que la déconnection anatomique des hémisphères remet en cause l'unité de soi.

On peut ainsi opposer les diverses fonctions spécifiques de l'hémisphère dominant et de l'hémisphère mineur :

« Hémisphère dominant	Hémisphère mineur
– Liaison vers la cse ind.	– Liaison vers la cse d'exister
– Verbal	– Presque non verbal
– Désignation linguistique	– Musical
– Idéation	– Sens de l'image et de la forme
– Ressemblances conceptuelles	– Ressemblances visuelles
– Analyse à long terme	– Synthèse à long terme
– Analyse du détail	– Holiste ; images
– Opérations arithmétiques ;	– Opérations spatiales et géométriques.
	computs. »

J. Eccles refuse d'adhérer à la conclusion selon laquelle il y aurait quasi-égalité des hémisphères. Il met l'accent sur la « différence considérable entre compréhension et expression dans les performances de l'hémisphère droit » : en effet l'hémisphère droit a accès au vocabulaire auditif, il est capable de relier des mots présentés par l'ouïe ou par la vue à des représentations graphiques, il répond aussi bien au nom d'action qu'au verbe correspondant ; mais il ne possède pas de capacité à exprimer verbalement ou par l'écriture la compréhension du langage qu'il a. J. Eccles renforce la thèse de l'asymétrie plutôt que de la symétrie fonctionnelle là où la neurophilosophie y développe, au contraire, la destruction de l'unité de la conscience. Pour J. Eccles : « Il existe une certaine conscience de soi dans l'hémisphère droit, mais celle-ci est étroitement limitée et ne permet pas de lui attribuer une personnalité d'après les critères que nous venons d'évoquer. Ainsi, la coupure du corps calleux a détaché un fragment de l'esprit conscient, mais la personne elle-même et son unité mentale sont pour l'essentiel restées intactes, dans leur association spécifique avec l'hémisphère gauche... Nous reconnaissons néanmoins que les émotions induites par l'hémisphère droit peuvent impliquer le gauche à travers le système limbique qui, lui, n'a pas été sectionné ; sur le plan émotionnel la personne reste donc bien attachée à l'hémisphère droit. » [\[19\]](#)

Non seulement J. Eccles sépare compréhension et expression en distinguant respectivement hémisphère

droit et hémisphère gauche, mais il ne reconnaît pas à l'hémisphère droit d'autre rôle dans l'unité de la personne que son influence affective. Il ne faut pas confondre, précise J. Eccles dans le chapitre X consacré à la personne humaine, le fait que le cerveau soit le siège de la personnalité consciente avec les données fournies par les dommages causés aux parties du cerveau qui peuvent perturber sérieusement la conscience de la personne humaine. L'ensemble des hémisphères cérébraux donnent son unité à la conscience. Faire de la conscience l'unité des deux hémisphères (thèse de la complémentarité de J. Eccles et R. W. Sperry) est bien une affirmation dualiste : la conscience est maintenue au-dessus de l'organisation complémentaire du système nerveux. La neurophilosophie, quant à elle, par la théorie du cerveau séparé attribue à chaque hémisphère des fonctions équipotentielles. Aussi cette thèse de l'équipotentialité tend à nier les différences de construction et de fonctionnement de chaque hémisphère cérébral.

À l'inverse de J. Eccles, la neurophilosophie refuse de distinguer compréhension (hémisphère droit) et compréhension-expression (hémisphère gauche). Ces cas peuvent être considérés comme de parfaits contre-exemples de l'hypothèse dualiste situant dans l'hémisphère gauche l'activité consciente. Le centre de la conscience n'aurait aucune unité empirique : en effet les études sur l'hémisphère droit prouvent que la connexion par le corps calleux n'est pas nécessaire

pour permettre au sujet de ressentir l'unité de soi-même. Et le critère de la production linguistique comme résultat d'une activité consciente privilégie une lecture hiérarchique des deux hémisphères.

Par cette revalorisation de l'hémisphère droit, la neurophilosophie espère démontrer l'équipotentialité des fonctions cérébrales. La vision, l'audition et le toucher constituent un lieu privilégié d'observation et d'étude dans la mesure où on peut distinguer nettement l'origine des stimulations et des réponses dans chaque hémichamp ; ces travaux ont commencé dans les années 1960 : « Pour prendre l'exemple de la vision, toute information apparaissant dans le champ visuel gauche stimule la moitié droite de la rétine de chaque œil et est ensuite transmise au cortex visuel de l'hémisphère droit, et l'inverse se produit pour l'information apparaissant dans le champ visuel droit. » [\[20\]](#) La référence aux travaux sur la distinction verbale - non verbale souligne l'asymétrie fonctionnelle du cerveau. Les recherches sur les patients souffrant de lésions à l'hémisphère droit montrent certains troubles de la compréhension verbale, ce qui remet en cause l'attribution traditionnelle au seul hémisphère gauche de la fonction de compréhension linguistique.

S'appuyant sur les travaux de M. Gazzaniga, J. E. Bogen et R. W. Sperry, P. S. Churchland distingue différentes capacités dans l'hémisphère droit : capacité d'expression, capacité de compréhension du langage et de manipulation des représentations linguistiques et

la capacité de penser dans le langage. Ces distinctions établissent des différences de degré entre compréhension et expression linguistique, cela afin de se préserver de la tentation de réduire l'hémisphère droit aux capacités connues et dominantes de l'hémisphère gauche. Pour M. Gazzaniga : « L'homme dont le corps calleux a été sectionné en deux sphères de conscience séparées, partageant un tronc cérébral, une tête et un corps commun. » [\[21\]](#) Cette division de la conscience en deux sphères correspondantes aux deux hémisphères cérébraux vise à prouver l'indépendance de chaque moitié lors de la section des commissures cérébrales.

Normalement celles-ci sont responsables du transfert interhémisphérique de l'apprentissage et de la mémoire. L'unité de la perception serait due à la complémentarité des hémisphères lors du transfert. M. Gazzaniga refuse « une conception unifiée du fonctionnement du cerveau » [\[22\]](#) : cependant il n'affirme pas la supériorité de l'hémisphère droit ; car les faits observés présentent un déficit de base de la mémoire à court terme lors des tests visuels.

Le double hémisphère

L'hémisphère droit servirait seulement à préciser et à nuancer les jugements faits par l'hémisphère gauche, mais en aucun cas à se substituer à lui. Cette fonction

de modération permet d'inscrire la complémentarité des hémisphères dans la thèse de leur relativité fonctionnelle. M. Gazzaniga reconnaît lui-même : « La section médiane sagittale des commissures cérébrales provoque un état de dédoublement mental ; deux sphères conscientes séparées coexisteraient dans un même crâne. Sans se donner la peine de définir ce qu'est ou n'est pas la conscience, on a simplement essayé de préciser ce que chaque hémisphère d'un cerveau dédoublé peut faire ou ne pas faire, séparément et indépendamment. » [\[23\]](#) Même si, comme nous l'avons vu, la théorie de l'existence d'une double conscience a été contestée par J. Eccles, la neurophilosophie affirme celle-ci sans la fonder autrement que par des résultats cliniques. Il s'agit d'une naturalisation de la conscience, puisque chaque hémisphère est crédité d'une activité consciente ou plutôt qu'au dédoublement du cerveau correspondrait un dédoublement de la conscience.

En divisant la conscience unifiée, la neurophilosophie confond ici la cause et l'effet : les deux hémisphères sont nécessaires pour produire la conscience unifiée, même s'il est vrai que chaque hémisphère est capable de performances singulières. Mais l'évaluation morcelée de celles-ci ne peut servir d'argument à leur assimilation comme une partie de la conscience ; en désignant trait pour trait la relation causale entre telle activité neurologique et telle possibilité de la conscience, on risque de réduire et la nature et la qualité de la conscience à l'état cérébral de l'individu

qui la produit. En abondant dans cette réduction n'affirmerait-on pas qu'une même conscience est à l'œuvre tant chez l'individu normal que dans un cas de split-brain ? La conscience n'est-elle que le résultat de la qualité et de l'intensité des transferts interhémisphériques ? Car chaque hémisphère possède un degré de conscience lié à ses possibilités neurocognitives. Mais R. W. Sperry et M. Gazzaniga vont plus loin dans l'hypothèse de l'existence de la double conscience : il ne s'agit pas d'une conscience qui exercerait la même activité dans deux hémisphères différents ; mais, plutôt, d'une attribution de conscience à chaque hémisphère dans la mesure où chacun est capable de percevoir, de comprendre et d'agir. Cette promotion d'une conception opératoire de la conscience la réduit à la somme des opérations cognitives évaluées dans chaque hémisphère. Or, la conscience ne se réduit pas à partir de l'état neurobiologique du cerveau car le cerveau est la condition d'une organisation cognitive autonome que nous nommons subjectivement la conscience.

Du point de vue des résultats le problème fut de déterminer la nature et l'étendue de la compréhension de l'hémisphère droit. L'hémisphère droit est incapable d'associer les images aux noms, de distinguer le passé à partir du présent, de déterminer à partir de la syntaxe le sujet et l'objet. Mais alors que J. Levy et R. W. Sperry en 1968 « remarquèrent que des formes géométriques et des visages étaient reconnues en fonction de leurs propriétés globales lorsqu'ils étaient

traités par l'hémisphère droit et en fonction de leurs détails ou composantes lorsqu'ils étaient traités par l'hémisphère gauche » [\[24\]](#). E. Zaidel en 1975 présenta une nouvelle technique d'évaluation de l'hémisphère droit dans son article « A technique for presenting lateralized visual input with prolonged exposure » : il proposa de présenter les mots durant une longue période afin de prouver les capacités de compréhension de l'hémisphère droit ; au moyen d'une lentille spéciale, il parvint à restreindre la diffusion de l'information à l'hémichamp visuel droit sollicitant une réponse exclusivement de la main gauche. Il démontra ainsi que l'hémisphère droit est capable d'assembler les lettres d'un mot, de reconnaître une rangée de mots ainsi que les verbes. La théorie selon laquelle le contrôle moteur de la parole et la compréhension étaient latéralisés uniquement dans l'hémisphère gauche devait prendre en compte ces résultats.

Michel Denis précise que : « L'ensemble de ces données nous invite donc à concevoir, à côté d'un système spécialisé dans le traitement des données verbales, l'existence d'un système différencié, sans doute relié à l'autre, mais fondamentalement distinct, dont le fonctionnement intéresse plus probablement certaines structures du système nerveux que d'autres, et dont la spécificité qualitative peut être caractérisée par le fait qu'il actualise des événements psychologiques descriptibles comme des analogues figuratifs des contenus perceptifs. » [\[25\]](#) S'il reconnaît l'« existence d'un système différencié », Michel Denis en

reste à une théorie analogique de l'hémisphère droit alors que P. S. Churchland y aperçoit un mode de cognition réel avec des capacités actives.

En s'appuyant sur les travaux de D. Kimura (« Functional asymmetry of the brain in dichotic listening », 1967), P. S. Churchland démontre les compétences linguistiques de l'hémisphère droit : ces études « démontrent que l'organisation cérébrale, en particulier pour le langage, n'est pas identique chez les droitiers et chez les gauchers. Grâce aux données neuropsychologiques et expérimentales telles que le test de Wada et l'écoute dichotique, on admet généralement une plus grande ambilatéralité dans la structuration du langage chez les gauchers » (D. Kimura, 1961 ; T. Rasmussen et B. Milner, 1977). Ainsi les techniques de stimulation auditive permettent à D. Kimura de sélectionner les performances de chaque oreille si bien que l'écoute dichotique prouve les compétences de l'hémisphère droit. Les travaux de Harold W. Gordon (« Right hemisphere comprehension of verbs in patients with complete forebrain commissurotomy : Use of the dichotic method and manual performance » [1980]) confirment l'intérêt de la méthode de D. Kimura : même si le langage est une opération relevant strictement de l'hémisphère gauche, les patients au cerveau divisé sont capables de quelques rudiments de compréhension linguistique.

P. S. Churchland conclut en ces termes : « L'hémisphère gauche est prééminent dans le contrôle

moteur de la parole et généralement devance de loin l'hémisphère droit dans la compréhension ; l'hémisphère droit est généralement en dessus des performances de l'hémisphère gauche en ce qui concerne les tâches de manipulations spatiales. Est-ce à dire que l'hémisphère gauche est latéralisé pour le langage ? Cela dépend de ce que l'on entend par "latéralisé" ? Si l'on entend par là que l'hémisphère gauche dans de nombreux cerveaux a des performances meilleures sur les tâches verbales que l'hémisphère droit, alors vraiment le langage est latéralisé dans l'hémisphère gauche. Si l'on veut dire que l'hémisphère droit n'a pas de capacités linguistiques significatives, ou qu'il ne contribue en rien aux opérations linguistiques normales, ou que le tissu neural dans l'hémisphère gauche est spécialisé pour les capacités linguistiques mais pas celui de l'hémisphère droit, alors la chose reste à prouver. Il y a une abondance de données qui donnent à penser et un ensemble de trucs ingénieux pour étudier le cerveau, mais justement tous les sens de ces expressions de capacités définissables restent profondément inexpliqués. » [\[26\]](#)

La stratégie de défense de la thèse réductionniste s'illustre parfaitement ici : P. S. Churchland reconnaît elle-même que les travaux sur l'hémisphère droit ne sont pas probants. Mais ils sont suffisamment corrosifs pour les définitions classiques de la répartition fonctionnelle des hémisphères pour qu'un doute autorise sinon l'élimination, du moins le

suspens de la théorie de l'asymétrie. Encore une fois P. S. Churchland ne conclue pas positivement en faveur de l'hémisphère droit : elle accumule des travaux qui portent sur peu de patients mais qui ont été répétés suffisamment de fois pour accréditer l'idée d'une remise en cause de l'unité de la conscience. La preuve de la supériorité cognitive de l'hémisphère gauche ne se trouve pas renforcée par les études des cas de cerveau divisé puisqu'elles montrent des patients qui manifestent des capacités de compréhension et d'expression avec l'hémisphère droit.

L'étude des cerveaux lésés vient remettre en cause les certitudes établies sur les cerveaux fonctionnant normalement : ainsi, comme pour la localisation des fonctions par le moyen d'un raisonnement inductif sur les lésions, comme nous l'avons montré pour Broca et son époque, P. S. Churchland utilise les cas de lésions pour introduire un doute dans la répartition des fonctions. « Comment la question est-elle posée ? », voilà l'interrogation que P. S. Churchland retourne contre les évidences de la théorie de la conscience unifiée. La latéralisation traditionnelle des fonctions ne serait pour elle qu'une conséquence de la conception unitaire de la conscience qui attribue une supériorité à l'hémisphère gauche et au mieux une relation de complémentarité fonctionnelle des deux hémisphères. La somme des parties forme un tout supérieur à celle-ci. Au contraire P. S. Churchland estime inutile la conservation du concept de conscience qui n'apporte rien de plus par rapport aux évaluations

neurocognitives effectuées par les recherches cliniques. Plus que la double conscience de R. W. Sperry et M. Gazzaniga elle préfère l'élimination du concept de conscience.

Notes

[1] Pierre Buser, a présenté ses travaux dans son article Neurobiologie et conscience, Revue de métaphysique et de morale, n° 2, avril-juin 1992, p. 165-184.

[2] D. Dennett, La conscience expliquée, (1991), trad. Pascal Engel, Paris, O. Jacob, 1993, p. 196-197. Pascal Engel présente un résumé du débat B. Libet - D. Dennett dans son livre Introduction à une philosophie de l'esprit, Paris, La Découverte, 1994, p. 196-201.

[3] P. S. Churchland, On the alleged backward referral of experiences and its relevance to the mind-body problem, Philosophy of Science, 48, 1981, p. 167.

[4] D. Dennett, La conscience, op. cit., p. 203.

[5] Ibid., p. 179.

[6] B. Libet, The experimental evidence for subjective referral of a sensory experience backwards in time : A reply to P. S. Churchland, Philosophy of Science, 48, 1981, p. 182-197, ici p. 196.

[7] B. Libet, The neural time-factor in perception, volition and free will, Revue métaphysique et morale, n° 2, 1992, p. 255-272, ici p. 265.

- [8] J. C. Eccles, Évolution du cerveau et création de la conscience, (1989), Paris, Fayard, 1992.
- [9] J. C. Eccles, Comment la conscience contrôle le cerveau, (1994), Paris, Fayard, 1997, p. 204.
- [10] Ibid., p. 232.
- [11] Ibid., p. 236.
- [12] Ibid., p. 254-255.
- [13] Ibid., p. 256.
- [14] P. S. Churchland, Neurophilosophy, op. cit., p. 319-320.
- [15] Ibid., p. 119.
- [16] G. Gazzaniga, Le cerveau dédoublé, (1970), Bruxelles, Mardaga, 1976, p. 16.
- [17] D. Laplane, La neuropsychologie s'intéresse-t-elle à la pensée, dans Le cerveau et la pensée, sous la dir. de H. Barreau, Paris, cnrs, 1992, p. 3-14, ici p. 11.
- [18] P. S. Churchland, Neurophilosophy, op. cit., p. 179.
- [19] J. Eccles, Évolution du cerveau et création de la conscience, (1989), Paris, Fayard, 1992, chap. IX, p. 280.
- [20] J. Sergent, Dilemmes de la gauche et de la droite, Psychologie et cerveau, sous la dir. de X. Seron, Paris, puf, 1990, p. 131.
- [21] M. Gazzaniga, Le cerveau dédoublé, (1970), Bruxelles, Mardaga, 1976, p. 15.
- [22] Ibid., p. 217.
- [23] Ibid., p. 220.
- [24] J. Sergent, Dilemmes de la gauche et de la droite, Psychologie et cerveau, op. cit., p. 138.
- [25] M. Denis, Les images mentales, Paris, puf, 1979,

p. 142-143.

[26] X. Seron et C. Laterre, Rééduquer le cerveau. Logopédie, psychologie, neurologie, Bruxelles, Mardaga, 1986.

Chapitre III

Naturalisation de l'inconscient

L'élimination de la psychanalyse

Deux courants psychologiques s'affrontent : d'une part, avec Sigmund Freud (1856-1939) la thèse de la fonction psychique du rêve (1901), et d'autre part, avec Édouard Claparède (1832-1871), sous l'influence de conceptions évolutionnistes, la thèse de la fonction biologique du rêve. Cette dichotomie arrange : il y aurait d'une part, la psychanalyse qui étudierait le retour du refoulé représenté dans et par la vie psychique de l'individu, et d'autre part, la psychophysiologie du rêve qui prouverait les conditions neurologiques du rêve. Il est vrai que si l'on s'en tient à cette coupure, le neurophysiologue français Michel Jouvet (1925) peut écrire : « Freud, inventeur de la métapsychologie, écarte tout ce qui est sommeil parce que c'est de la physiologie. S. Freud fait du rêve l'expression d'un désir et le gardien du sommeil. Il construit un véritable

appareil psychique en dehors du cerveau. Cette topique permet de considérer les espaces correspondants au Ça, Moi et Surmoi. Concepts qui firent fortune et qui attendent toujours une impossible réfutation expérimentale. » [\[1\]](#)

Pourtant en lisant le Freud de l'Esquisse d'une psychologie scientifique (dans sa correspondance Freud fait mention du titre « Psychologie à l'usage des neurologues ») le chapitre 19 de la première partie « Processus primaires. Sommeil et rêves » s'attache à décrire une psychophysiologie du sommeil. De l'aveu même de S. Freud : « Nous avons cherché à faire entrer la psychologie dans le cadre des Sciences naturelles, c'est-à-dire à représenter les processus psychiques comme des états quantitativement déterminés de particules matérielles distinguables, ceci afin de les rendre évidents et incontestables. » [\[2\]](#)

Ainsi S. Freud décrit le sommeil en suivant le modèle énergétique : le sommeil est « une diminution de la charge endogène dans le noyau psy... pendant qu'il dort, le sujet se trouve dans un état idéal d'inertie alors que son stock de quantité (Qn)... » (op. cit., p. 353). Les neurones psy se distinguent des neurones phi par leurs capacités à être imperméables, c'est-à-dire résistants et rétenteurs de quantité (Qn). D'eux dépend les mécanismes de la mémoire. Ainsi la mémoire est « représentée par les frayages se trouvant entre les neurones psy » (op. cit., p. 320). Cette opposition entre les neurones phi de la perception et les neurones psy oblige S. Freud à une conception passive du sommeil.

Tel pourrait être le sens de l'oubli de S. Freud des neurophysiologues contemporains du sommeil dans la mesure où la découverte du sommeil à mouvements oculaires rapides ne sera faite par Eugène Aserinsky et Nathaniel Kleitman qu'en 1953. Dans le contexte de l'époque l'idée d'inertie est cohérente avec les travaux, souligné par Michel Jouvet, du D^r Jean Baptiste Édouard Gélinau (1828-1906) qui en 1880 individualise la narcolepsie ; « ... l'une des caractéristiques primordiales du rêve, c'est-à-dire l'absence totale de tonus musculaire. » [3] S. Freud n'aurait développé qu'une conception seulement psychique du rêve ; mais, même dans L'interprétation des rêves (1899), S. Freud reprend la thèse de la diminution d'activité consciente selon laquelle « la condition essentielle de la formation des rêves est le sommeil de l'esprit » [4]. S. Freud y reprend sa théorie du système psy sans toutefois délaisser la thèse des traces mnésiques de nos impressions par le moyen des neurones : « Si maintenant nous trouvions confirmation de ce fait que la mémoire et la qualité qui caractérisent la conscience s'excluent l'une l'autre dans les systèmes psy, nous aurions des aperçus gros de promesse sur les conditions de l'excitation des neurones » (op. cit., p. 458). Il est vrai aussi que S. Freud refuse de défendre un matérialisme réductionniste lorsqu'il écrit, en 1901, dans Le rêve et son interprétation à propos des rêves d'enfant y voir « un motif de plus de ne pas vouloir ramener le rêve à

une activité partielle du cerveau dans le sommeil, car pourquoi cette réduction des fonctions psychiques serait-elle propre au sommeil de l'adulte et non pas à celui de l'enfant ? » [5]. S. Freud n'est pas contre la production cérébrale du rêve, mais il distingue la production psychique du contenu représentatif du rêve avec sa cause physiologique :

Contenu de représentation du rêve

Effet psychique dans le système Psy

Cause neurologique : neurone Psy

S. Freud ne renie pas la détermination neurologique, mais il accorde au contenu représentatif du rêve une autonomie à l'intérieur du système psy. Bien sûr on n'a jamais observé de neurone psy, mais dans le contexte de l'époque aucune théorie ne se fonde sur un enregistrement de l'activité du cerveau dans le sommeil.

Au chapitre VII de son livre, *Le sommeil et le rêve*, Michel Jouvet se livre à une critique précise de la thèse freudienne de 1894 : S. Freud concevait les neurones comme des réservoirs passifs d'énergie susceptibles de recevoir des quantités d'énergie extérieures. À l'inverse, la conception moderne attribue aux neurones leur propre source d'énergie métabolique pour maintenir un gradient de potentiel électrique entre l'extérieur et l'intérieur de la membrane cellulaire.

Ignorant cette fonction homéostasique du neurone, S. Freud construit un système d'échange quantitatif entre l'extérieur et l'intérieur qui ne prend pas en compte les principes de la théorie électrophysiologique moderne :

- Chaque cellule génère son propre « potentiel de repos ».
- Influence mutuelle des neurones par rapport à la polarisation préexistante.
- Faiblesse de l'énergie échangée et de la fréquence des décharges.

Les neurones restent chez S. Freud des conduits pour la transmission de l'énergie de l'extérieur du cerveau à l'inverse des récepteurs sensoriels qui, aujourd'hui, sont pensés comme transformateurs d'énergie.

Sur des bases aussi fausses, la théorie du rêve de S. Freud paraît incongrue au regard de la théorie neurophysiologique actuelle : la notion de sommeil paradoxal ne pouvait exister dans le système freudien car « il ne postule, à aucun endroit, des systèmes de neurones automatiquement régulés, ou leur contrepartie. La conclusion qui s'impose est que la théorie de S. Freud doit être abandonnée à cause de l'absence d'activité autonome et de l'absence de régulation et d'énergie endogène du cerveau » [\[6\]](#). En défendant l'idée que la force agissante au cours du sommeil paradoxal est une activation biologique et non

pas l'effet d'un désir refoulé, M. Juvet reconnaît que le rêve est devenu un phénomène physiologique à partir du moment où l'eeg (électro-encéphalogramme) a permis une évaluation quantitative des rem (mouvement oculaire rapide) au cours du sommeil. La psychanalyse, s'il peut se comprendre dans le cadre épistémologique de l'époque, ne peut se poursuivre, à notre avis, que dans la dimension de l'interprétation symbolique des rêves et non de l'explication psychophysiologique du cerveau. Le divorce entre les neurosciences et la psychanalyse alimente un vain conflit : chacun possède son champ de validité à la seule condition qu'il s'en tienne au seul domaine de compétence de son activité. La théorie freudienne paraît désuète par l'effet du progrès de l'observation de l'activité électrophysiologique du cerveau durant le sommeil, mais pas son travail d'interprétation.

Électrophysiologie du rêve

En 1905, Édouard Claparède, psychologie genevois, attribue une fonction biologique instinctuelle au rêve : dimension utilitaire du rêve, cette conception n'en vient pas encore à une description neurofonctionnelle dans la mesure où elle demeure dans une théorie des facultés. Au cours de son œuvre de Henri Piéron (1881-1964), anticipa les résultats obtenus par l'électro-

encéphalographie à partir des années 1950. En 1913, il « réussit en effet à transférer par voie sanguine ou ventriculaire les « hypnotoxines » d'un chien privé de sommeil à un receveur et à induire chez celui-ci un comportement de sommeil profond » [7]. Son livre, paru la même année, *Le problème physiologique du sommeil* (Paris, Masson, 1913) pose les principes d'une méthode expérimentale du sommeil.

Entre 1880 et 1953, les travaux ne parviennent cependant pas à proposer une synthèse active des étapes du sommeil. Ainsi l'utilisation de l'invention de l'électro-encéphalographie en 1929 de Hans Berger (1873-1941) n'a fourni que des travaux partiels : comme « en 1937, l'Allemand Klaue découvre chez le chat des périodes de *Tiefen Schlaf* (sommeil profond) accompagnées d'une activité électrique corticale rapide très différente de l'activité corticale lente du sommeil... En 1944, l'Allemand Ohlmeyer décrit chez l'homme – dans un journal de physiologie – un cycle d'érection périodique au cours du sommeil. Ce cycle débute quatre-vingt-dix minutes après l'endormissement et les phases d'érection d'une durée moyenne de vingt-cinq minutes ont une périodicité moyenne de quatre-vingt-cinq minutes. Ce sont les caractéristiques exactes des périodes de rêve mais l'érection ne fut pas alors reliée au rêve » [8].

Alfred Loomis publia en 1937 (avec E. Newton Harvey et G. A. Hobart) un article « *Cerebral states during sleep, as studied by human brain potentials* » (*Journal*

of Experimental psychology, 1937, 21) ; il y établit l'existence d'états de sommeil distincts et en proposa une classification. Mais il associa les différents types d'ondes à différents niveaux de conscience, ce qui le condamnait à localiser le rêve à l'intérieur d'un état de sommeil à faible voltage. Cette interprétation n'était pas toujours confirmée au regard du témoignage subjectif des sujets.

Nathaniel Kleitman, physiologiste, professeur à Chicago, eût dès 1939 l'idée d'utiliser la méthode électro-encéphalographique en la combinant avec une interrogation des sujets sur leurs rêves. Comme le montre Claude Debru la localisation ne pouvait plus être statique depuis que John Hughlings Jackson (1835-1911) « avait proposé une classification des niveaux de l'organisation cérébrale fondée sur le critère de l'automatisme » [\[9\]](#). La différence entre les centres inférieurs à organisation simple et automatique et les centres supérieurs, complexes, modifiables et faiblement organisés, pour permettre les apprentissages et les associations, convient tout à fait au projet de P. S. Churchland. Elle y voit, dans cette partie de Neurophilosophy sur les fonctions supérieures du cerveau, le moyen de prouver l'activité cérébrale indépendamment du contrôle de l'attention consciente. Dans son ouvrage de 1939, *Sleep and Wakefulness* (University of Chicago Press), N. Kleitman, selon Claude Debru, attribue aux centres supérieurs corticaux la fonction « d'analyser et d'interpréter les signes afférents, de former des

associations nouvelles, et également d'exercer une action inhibitrice sur la pensée et sur l'action, par suite de l'expérience et de l'entraînement acquis... le rêve est la réponse non critique des niveaux inférieurs de l'organisation corticale aux signaux afférents » [\[10\]](#).

Même si P. S. Churchland ignore explicitement le travail de 1939, elle en partage les options philosophiques dans la reprise de 1953 : en effet au début des années 1950, Eugène Aserinsky, étudiant en cours de thèse au laboratoire de physiologie de N. Kleitman, dégage une certaine organisation périodique de l'activité oculaire en observant les mouvements des paupières au cours du sommeil des nouveau-nés. Une technique d'enregistrement électrique, l'électro-oculographie, permet de définir des mouvements oculaires non plus lents et pendulaires mais rapides, saccadés et conjugués. Ce que l'on désigna du nom de « mouvements oculaires rapides », soit rem. P. S. Churchland n'en resta pas à ce premier résultat qui cantonne le rêve seulement à cette période rem, mais y voit avec D. Foulkes (1966, *The Psychology of the Sleep*) un état cognitif qu'elle désigne du nom de « mentation ». P. S. Churchland met l'accent sur l'existence d'une sorte de mentation-cognition qui prouve que les états neurobiologiques produisent une activité en dehors du traditionnel contrôle de la conscience. Ou plutôt, dans la mesure où pendant le sommeil le cerveau produit des « mentations » alors la séparation entre l'activité consciente et inconsciente n'a plus de sens si l'on continue à la conjuguer sur

l'opposition négatif-positif.

La neurophilosophie américaine se réfère aux travaux de S. Clemes et W. Dement (The effect of rem sleep deprivation on psychological functioning, Journal of Nervous and Mental Disorders, 1967, 144) : William Dement est un élève de N. Kleitman qui poursuit l'étude de l'association de la motilité oculaire rapide et de l'électro-encéphalogramme de bas voltage ; l'étude des bases neurobiologiques de la conscience est un des enjeux principaux de cette recherche car le rêve est défini comme un niveau de conscience qui correspond à un certain modèle électro-encéphalographique. Comme le souligne Claude Debru c'est la thèse du parallélisme psychophysique qui est défendue à travers ces travaux : en effet, en mettant l'accent sur des critères d'observation électrographique le constat est accentué d'une correspondance entre niveau d'activité de la conscience et niveau d'eeg. P. S. Churchland utilise les résultats de S. Clemes et W. Dement à propos de la privation de sommeil : elle rapporte que la corrélation entre la privation de sommeil rem et les données comportementales comme l'irritabilité et la sexualité a été établie, mais qu'aucune démonstration solide n'a été suffisante pour prouver le lien entre la privation de sommeil rem et quelque chose comme la désintégration cognitive. Claude Debru confirme que : « Dement ne put établir de différence significative entre psychotiques et sujets normaux dans l'organisation périodique du sommeil à mouvements oculaires rapides. » [\[11\]](#)

Au plan épistémologique on peut se demander si P. S. Churchland ne reproduit pas le défaut dénoncé par Claude Debru en se référant toujours à l'école de Chicago pour valider la différence de niveaux dans l'organisation des fonctions supérieures du cerveau : « La théorie cyclique et continuiste des niveaux de conscience, le cadre conceptuel jacksonien adopté par Kleitman faisaient écran pour la perception de la véritable nature de ce syndrome. En faisant du sommeil à mouvements oculaires rapides un stade de sommeil léger, associé au rêve, les chercheurs de Chicago se sont spontanément conformés à la tradition philosophique du rêve comme sommeil léger, à l'intérieur de laquelle la conception freudienne du rêve comme satisfaction d'un désir refoulé et gardien du sommeil pouvait trouver sa place. La description minutieuse des phénomènes, leur interprétation dans le cadre du parallélisme psychophysique sont des traits particuliers de cette science américaine. » [\[12\]](#)

P. S. Churchland trouve utile de se référer à cette science américaine qui poursuit la méthodologie de J. Hughlings Jackson au point d'établir une continuité de degré dans les états neurobiologiques. Ce qui satisfait son exigence d'homogénéité dans tous les phénomènes cérébraux.

Le paradoxe du sommeil

Le sommeil paradoxal, découvert par Michel Jouvet en 1959, vint s'opposer à la conception de l'école de Chicago qui assimilait le rêve à un demi-sommeil intermédiaire entre le sommeil profond et l'éveil. Ayant commencé ses premiers travaux par l'étude de la plasticité cérébrale, des phénomènes de conditionnement et d'attention, M. Jouvet séjourne un an, en 1955, dans le laboratoire du P^r Horace Magoun (1907-1991) à Los Angeles qui étudiait le déterminisme des divers états électro-encéphalographiques de vigilance et d'attention. En 1956, il fait paraître *Étude électro-encéphalographique de l'établissement des liaisons temporaires dans le système nerveux central* (Lons-le-Saunier, imprimerie Maurice Declume) : Claude Debru voit dans le début de ce travail « le rappel d'une histoire philosophique, dont le neurophysiologiste ne veut pas se couper. Michel Jouvet a ainsi été l'un des premiers lecteurs de l'ouvrage de Georges Canguilhem sur la formation du concept de réflexe » [\[13\]](#). En 1962, M. Jouvet remarqua que ses chats, lorsqu'ils s'ennuyaient et tombaient dans le sommeil, perdaient leur tonus musculaire. Cette atonie musculaire, causée par l'inhibition spontanée du pont du tronc cérébral que, en 1962, M. Jouvet reconnût comme étant le centre cérébral du sommeil cérébral.

L'étude de l'activité électrique de diverses structures cérébrales au cours du sommeil se réalisa par l'ablation du cortex sur le chat. Le neurologue François

Michel, collaborateur de M. Jouvét, lui suggéra d'enregistrer l'électromyogramme. « L'addition de ce deuxième canal de mesure, l'emg, eut pour Jouvét et Michel autant d'importance que l'addition de l'électro-oculogramme (eog) en avait eu pour Aserinsky et Kleitman. » [\[14\]](#) Il caractérisa avec son collègue François Michel une « phase paradoxale » marquée « par l'apparition de bouffées d'activité rythmique de fuseaux de grande amplitude, se répétant de façon régulière au niveau de la formation réticulée pontique. En même temps que s'installe cette activité, deux ordres de phénomènes apparaissent : d'une part, une activation rhinencéphalique continue à rester rapide, bien que parfois des "fuseaux" y apparaissent ; d'autre part, une disparition totale de toute activité tonique musculaire » [\[15\]](#).

Le caractère paradoxal provient donc dans cette phase de la coexistence d'une activité cérébrale rapide et de la disparition totale du tonus musculaire. Selon M. Jouvét l'augmentation du sommeil paradoxal est en relation avec la maturation du cerveau. Ainsi plus un mammifère nouveau-né est immature, plus le temps occupé par le sommeil paradoxal est important : 50 % à 60 % de la durée du sommeil pour un nouveau-né humain. Les périodes de sommeil paradoxal ont une durée de six minutes et surviennent toutes les vingt-cinq minutes au cours du sommeil. J. A. Hobson précise la modification de paradigme effectuée par Michel Jouvét : « En lisant l'article de Dement, écrit en 1958, sur le sommeil paradoxal des chats, Jouvét

saisit immédiatement la signification de ce que Michel et lui avaient observé. Dans cet exemple classique de changement de paradigme (changement dans la manière dont un chercheur envisage ses résultats), Jouvét fut capable de passer du point de vue de l'apprentissage (transformations suscitées par des stimuli expérimentaux) à celui du changement d'état endogène et cyclique qui décide de la facilité de l'animal à lui répondre. Autrement dit, il est passé du paradigme du réflexe (stimulus-réponse) au paradigme de l'état. » [\[16\]](#)

Ce changement de paradigme indique combien l'électrophysiologie du rêve constitue une rupture par rapport à la notion mécanique de réflexe. Les différences de niveaux décrivent les variations d'intensités des états successifs. Rêver devient une activité physiquement diagnosticable par le moyen de l'enregistrement de la correspondance entre onde et état.

Neurophilosophie du rêve

L'accueil des travaux de Michel Jouvét par les hypnologues américains se traduira par une mention trois ans plus tard, en 1962, de la découverte de la phase paradoxale. La référence de P. S. Churchland s'effectue à partir de l'article de 1967 « Neurophysiology of the states of sleep », paru dans

Physiological Review. Cet article s'inscrit dans le volume dirigé par Michel Jouvet, Neurophysiologie des états du sommeil, qui fut publié aux Colloques internationaux du cnrs en 1965. P. S. Churchland situe les travaux de Michel Jouvet autour de la question de l'activité du cerveau dans la période rem du sommeil et dans la période non-rem. Elle se réfère à trois thèses :

- Le sommeil paradoxal joue un rôle dans les processus d'oubli ou de désapprentissage (F. Crick et G. Mitchison, 1983).
- Sommeil paradoxal et ontogenèse. Hypothèse sur les fonctions du sommeil paradoxal dans le développement ontogénétique (F. Roffwarg, J. N. Muzio et W. C. Dement, 1966).
- Le rêve constitue le moment privilégié de l'interaction entre les événements épigénétiques et des schèmes génétiquement programmés de comportements instinctuels. Programmation itérative génétique (M. Jouvet, 1967).

P. S. Churchland connaît les travaux de Francis Crick comme elle l'écrit dans la préface de Neurophilosophy : « Je suis reconnaissante à Francis Crick, dont la compréhension générale du champ entier des problèmes théoriques en neurobiologie de la vision et des sens a pu être une solution directe dans ma compréhension des questions fonctionnelles. » [\[17\]](#) La

thèse défendue par F. Crick et G. Mitchison (On the function of dream sleep, Nature, 1983, 304), affirme que le sommeil paradoxal joue un rôle dans les processus d'oubli ou de désapprentissage : M. Jouvet analyse cette thèse ainsi : « Le rêve serait alors un processus de nettoyage où le cerveau, fonctionnant en circuit fermé, pourrait se libérer de toutes les modalités parasites en raison de la création de nouveaux circuits d'informations... Ainsi, la fonction du rêve serait de réaliser un « apprentissage en sens inverse » (a reverse learning mechanism) qui modifierait le cortex, par exemple, en altérant la résistance des synapses... Crick et Mitchison suggèrent ainsi qu'au cours du sommeil paradoxal, nous désapprenons nos rêves inconscients. "Nous rêvons dans le but d'oublier". » [\[18\]](#)

Michel Jouvet nous permet de comprendre l'attachement de P. S. Churchland pour cette thèse dans la mesure où le paradigme qui la rend possible est celui du réseau neuronal et d'informatique. Comme nous la montrons dans cette partie, la perspective de la neurocomputation est la finalité de la neurophilosophie car le réseau neuronal peut reproduire in vitro les mouvements et activités in vivo du cerveau. En s'associant à un informaticien, Mitchison, Francis Crick pense que « le cortex cérébral, dans son immense complexité, constitue un réseau neuronal ».

« Au sein de ce réseau, l'information va être distribuée à un grand nombre de synapses. Un tel réseau peut devenir surchargé s'il doit traiter

simultanément des modèles trop différents ou trop vastes. Cette surcharge du réseau devrait donc entraîner des associations “bizarres” ou reproduire toujours le même état d’association (des “obsessions”) ou, enfin, surtout si le réseau possède une “action en retour” fermée sur lui-même, il répondra par des “hallucinations” à des signaux qui devraient normalement n’entraîner aucune réponse. »

Le sommeil paradoxal aurait donc pour objectif d’éliminer les mémoires sans importance. Le rêve serait le moment et le moyen de déconnexion des entrées et des sorties du réseau neuronal. Pour être aussi efficace qu’un superordinateur, le cerveau devrait posséder « l’irruption périodique de signaux stochastiques afin d’effacer les mémoires saturées par des signaux dépourvus de signification » [\[19\]](#). Comme le souligne M. Jouvet cette thèse est la reprise de celle de W. Robert, *Der Traum als Notwendigkeit erklärt* (Hambourg, 1886) cité par S. Freud dans *L’interprétation des rêves* : en effet, S. Freud écrit : « Robert se représente donc le rêve “comme un processus somatique d’élimination dont nous prenons connaissance par notre réaction à lui”. Le rêve est élimination de pensées étouffées dans l’œuf... Le rêve joue, pour le cerveau surchargé, le rôle de soupape de sûreté. Les rêves ont un pouvoir de soulagement, de guérison. » [\[20\]](#)

Ainsi c’est la surcharge du cerveau qui demande à se

libérer afin de réguler la dynamique entre les entrées et les sorties. Cette conception de F. Crick reprend la perspective quantitative de W. Robert en actualisant le propos par une modélisation informatique. Cette corrélation entre le volume cérébral et la quantité de sommeil paradoxal présente un obstacle lorsque l'on observe les conséquences cliniques du traitement des états dépressifs par des inhibiteurs des monoamines oxydases ou des antidépresseurs tricycliques ; M. Jouvet y voit là une preuve suffisante de l'absence de corrélation systématique entre rêve et oubli : en effet, ces drogues suppriment tout sommeil paradoxal sans que l'on ait pu constater un dysfonctionnement de la mémoire ou une augmentation des troubles cognitifs.

Aussi en appliquant le modèle du circuit informatique au réseau neuronal, F. Crick développe l'idée d'une circularité économique où la quantité de mémoire est toujours égale même si la qualité en est constamment renouvelée par le processus d'élimination utilitaire pratiqué durant le sommeil paradoxal. P. S. Churchland estime qu'il est encore trop tôt pour invalider la thèse de F. Crick, mais elle pose la validité des deux autres thèses. L'avantage de la thèse de F. Crick est de présenter une version économique du cerveau qui anticipe les travaux sur le cerveau computationnel de P. S. Churchland. La responsabilité de F. Crick dans la bascule théorique de P. S. Churchland de la neurophilosophie américaine à la neurocomputation est importante puisque en 1992 dans *The Computational Brain*, les travaux de M. Jouvet sont

définitivement écartés au profit de ceux de F. Crick.

Enfin P. S. Churchland trouve dans la thèse de F. Crick une version éliminativiste de la mémoire : au sein de sa thèse de l'élimination des états mentaux, ce fonctionnement neurobiologique du rêve assure une description homogène des états de non-conscience. Ce serait pour l'équilibre du cerveau qu'une régulation onirique rétablirait l'homéostasie neurofonctionnelle.

La thèse de H. Roffwarg, Muzio et W. Dement repose elle sur des travaux effectués entre 1960 et 1966 sur l'ontogenèse du cycle éveil-sommeil. Selon cette thèse « le sommeil paradoxal fournirait une stimulation intense, d'origine endogène, au moment où l'organisme est dépourvu d'excitations exogènes. Les décharges issues du pace-maker ponto-bulbaire pourraient ainsi contribuer au développement et à la myélinisation des centres supérieurs thalamo-corticaux. Cette stimulation pourrait "anticiper" ou "préparer" le cerveau à réagir de façon adaptée aux futures stimulations sensorielles » [\[21\]](#). Cette thèse de la stimulation fonctionnelle ne peut être entièrement retenue par P. S. Churchland. En effet, elle présente le défaut de suivre un modèle exclusivement physiologique pour expliquer l'activité du cerveau. Par stimulation, la thèse avance l'idée d'une continuité causale à l'œuvre tout au long de l'existence qui maintiendrait l'activité du cerveau à un niveau suffisant. Le sommeil paradoxal jouerait un rôle essentiel dans la maturation [\[22\]](#) du système nerveux central au cours

de la vie fœtale, néo-natale et tout au long de la vie adulte. Un tel développement du cerveau présuppose un modèle évolutif dans lequel la constance et l'intensité de l'activité du cerveau sont soumises à une régulation croissante.

Or, P. S. Churchland accorde peu d'importance au modèle darwinien de la construction du cerveau dans la mesure où cela l'obligerait à délaisser la perspective neurocomputationnelle qui suppose une analogie fonctionnelle entre le réseau neuronal formel et le réseau neuronal vivant. La limite du connexionnisme de P. S. Churchland s'exprime ici dans sa difficulté à reconnaître le lien entre l'épigenèse et l'ontogenèse : en effet, la stimulation fonctionnelle lors du sommeil paradoxal implique la reconnaissance d'un mouvement interne pour la maturation du cerveau. Et si P. S. Churchland accepte une maturation cognitive, elle refuse d'accorder au seul sommeil paradoxal la cause de la maturation cérébrale :

- La thèse de M. Jouvet intéresse P. S. Churchland dans le strict débat de l'inné et de l'acquis. En effet, l'hypothèse de la programmation itérative génétique au cours du sommeil paradoxal repose sur un certain nombre de principes.
- L'impossible conservation de l'individuation génétique par neurogenèse permanente chez les mammifères après la période postnatale fait

supposer qu'un nouveau mode de programmation génétique est apparu avec l'homéothermie.

- Selon une hypothèse sélective, l'information de la programmation héréditaire dépendrait alors seulement de la sélection des neurones qui sont, ou ne sont pas, excités par une programmation endogène dont le codage peut être stochastique.
- Il devrait être possible, dans certaines conditions, de distinguer une composante génétique dans les patterns d'activité des motoneurones. C'est au cours du sommeil que la programmation génétique du cerveau peut s'effectuer sans représenter un danger pour celui-ci : à ce moment-là l'inhibition des afférences intero- et exteroceptives et le blocage de l'activité musculaire permettent la réalisation de cette programmation génétique du cerveau.

Ce modèle spéculatif a été proposé par M. Jouvet en 1977, soit dix ans après la référence de P. S. Churchland. Pourtant elle en connaît les enjeux, Neurophilosophy étant écrit en 1982-1983, car elle précise : « Par contraste avec la fonction du rêve obtenu par Crick-Mitchison, l'hypothèse de Jouvet envisage que le pattern du rêve est génétiquement programmé et que la consolidation et la rationalisation du contrôle moteur en sont l'objet. » [\[23\]](#) Même si P. S. Churchland

relativise la portée de la thèse de Michel Jouvét en indiquant qu'elle ne porte que sur le cerveau « d'animaux inférieurs » comme les chats, elle reconnaît l'importance de l'enjeu face à la thèse de Crick-Mitchison. Claude Debru explique que « la programmation génétique désigne une action, elle-même génétiquement programmée, qui porte sur l'expression des propriétés génétiques dans le plan de câblage du système nerveux. Il s'agit d'un système supplémentaire de contrôle, de modulation, de renforcement de la structure du système nerveux telle qu'elle est génétiquement déterminée et établie » [\[24\]](#) Ainsi la réorganisation des programmes instinctuels complexes a lieu pendant le sommeil paradoxal par le moyen de la transformation de message biochimique en un message électrique ; par là le code génétique biochimique deviendrait le code nerveux du cerveau. M. Jouvét avance que non seulement le sommeil paradoxal préserve les caractéristiques génétiques de base de l'espèce mais aussi les variantes phénotypiques des comportements issus des apprentissages.

P. S. Churchland ne peut, à l'intérieur d'une perspective neurocomputationnelle du connexionnisme, admettre les conclusions philosophiques du « couple » Michel Jouvét - Claude Debru : en effet, ils défendent la thèse d'un parallélisme psychophysique renouvelée car le sommeil paradoxal n'exclut nullement l'existence de fonctions psychiques du rêve : « On invoquera ici l'aspect polyfonctionnel des structures et fonctions

vitales. Fonctions biologiques et fonctions psychiques doivent coexister. Mais il se pose ici un redoutable problème qui, à nouveau, met en jeu le parallélisme psychophysique et la validité que l'on peut lui reconnaître. Si elles peuvent coexister, fonctions biologiques et fonctions psychiques peuvent-elles pour autant être sans aucun rapport les unes avec les autres ?... En raison de son caractère périodique, itératif, le sommeil paradoxal doit avoir une fonction d'homéostasie liée) la plasticité cérébrale, qu'il s'agisse d'une fonction d'incorporation, de mise en mémoire des événements de l'éveil (peut-être liée au stress) ou, au contraire, d'une fonction de désinscription, au profit de la reprogrammation de propriétés idiotypiques. Les fonctions psychiques du rêve doivent dériver de fonctions de ce genre. À tout le moins, elles ne doivent pas être incompatibles avec elles... » [25] Au sein de la thèse 1, pour laquelle il n'existe que des états neurobiologiques, ce genre de dérivation est incompatible avec la méthode éliminativiste de ce matérialisme : le modèle de la dérivation suppose une causalité inductive et continue par laquelle les états neurobiologiques seraient la cause des états psychiques. Ce qui suppose une différence qualitative à partir d'une production exclusivement neurobiologique.

L'utilisation du néologisme de « neurophilosophie » « dans l'expression de « neurophilosophie du rêve » conduit à une conclusion différente de celle du matérialisme éliminativiste de P. S. Churchland.

Notes

- [1] P. S. Churchland, Neurophilosophy, op. cit., p. 193.
- [2] É. Claparède, Esquisse d'une théorie biologique du sommeil, Archives de psychologie, 1905.
- [3] M. Jouvet, Le sommeil et le rêve, Paris, O. Jacob, 1992, p. 40.
- [4] S. Freud, La naissance de la psychanalyse, Paris, puf, 3^e éd., 1973, p. 315.
- [5] M. Jouvet, Le sommeil et le rêve, op. cit., p. 41.
- [6] S. Freud, L'interprétation des rêves, Paris, puf, 1967, p. 447.
- [7] S. Freud, Le rêve et son interprétation, (1901), Paris, Gallimard, coll. « Folio/Essais », 1985, p. 30.
- [8] M. Jouvet, Le sommeil et le rêve, op. cit., p. 41.
- [9] Ibid., p. 42.
- [10] C. Debru, Neurophilosophie du rêve, op. cit., p. 45.
- [11] Ibid., p. 46
- [12] Ibid., p. 55. Voir aussi William C. Dement, Dormir, rêver, Paris, Le Seuil, 1981.
- [13] C. Debru, Neurophilosophie du rêve, op. cit., p. 58.
- [14] Ibid., p. 63.
- [15] J. A. Hobson, Le cerveau rêvant, op. cit., p. 55.
- [16] Cité par C. Debru, M. Jouvet et F. Michel, Corrélations électromyographiques du sommeil chez le chat décortiqué et mésencéphalique chronique, Comptes rendus des séances de la société de biologie, 1959, 153

[17] J. A. Hobson, Le cerveau rêvant, op. cit., p. 185.

[18] P. S. Churchland, Neurophilosophy, op. cit., p. X.
C'est tout un réseau qui s'organise à La Jolla, Californie autour de Francis Crick qui a été recruté comme professeur associé en psychologie. Arrivés en 1984 au département de philosophie Paul et P. S. Churchland se trouvent en relation avec la psychologie psychophysique de la vision de Bob Boyton, qui sera poursuivie par Don MacLeod et V. S. Ramachandran (avec qui P. S. Churchland publiera des articles en 1992). Ces spécialistes de la neurovision forment avec le physicien du Campus d'Irvine de l'Université de Californie, Gordon Shaw, le club Helmholtz. Du côté du traitement cerveau-ordinateur, Francis Crick collabore avec David Rumelhart et Jay Mac Clelland qui formulent le traitement parallèle par réseaux (pdp) en 1986. Terry Sejnowsky (avec qui P. S. Churchland publiera The Computational Brain en 1992) travaille avec Charles Rosenberg au réseau NetTalk. Enfin de 1980 à 1986 à l'Institut de neurosciences Max Cowan collabore avec Francis Crick. Au mit R. Andersen et Simon Le Vay (auteur en 1993 de Sexual Brain) complètent le réseau de la neurophilosophie

[19] M. Jouvet, Le sommeil et le rêve, op. cit., p. 163-164.

[20] P. S. Churchland, Neurophilosophy, op. cit., p. 208.

[21] C. Debru, Neurophilosophie du rêve, op. cit., p. 155.

[22] Ibid., p. 186-187.

[23] P. S. Churchland, Neurophilosophy, op. cit., p. 254.

[24] Ibid., p. 265

[25] O. Neurath, Énoncés protocolaires, dans Manifeste du Cercle de Vienne, Éd. A. Soulez, Paris, puf, 1985, p. 214-231, ici p. 223.

Chapitre IV

Connexionisme et neurocomputationnisme

La neurophilosophie trouve dans le connexionisme, plutôt que dans le cognitivisme, la possibilité matérielle de montrer comment les réseaux neuronaux sont représentationnels tout en refusant les représentations symboliques traditionnelles.

Les réseaux neuronaux

L'ordinateur est-il une représentation seulement symbolique des réseaux neuronaux, ou, au contraire, permet-il de représenter les mouvements réels de ceux-ci. P. S. Churchland se réfère ici au débat qui opposait en 1984, à la Société de science cognitive, David Rumelhart et Geoffrey Hinton, défenseurs de la modélisation par réseaux, à Zénon Pylyshyn et Kurt Van Lehn, qui soutenaient que les réseaux étaient des instruments inadéquats pour les activités cognitives. L'intérêt des premiers est d'avoir développé avec (J. A. Anderson, Bower, Norman...) des réseaux extrêmement localistes dans lesquels les unités

(appelés des nœuds) codaient des concepts. Mais P. S. Churchland ne veut pas s'en tenir à ces premiers réseaux sémantiques, car il leur manque la liaison avec l'architecture fonctionnelle du cerveau. Elle reconnaît l'intérêt de ces processus de diffusion de l'activation du réseau qui lui donne un caractère distribué.

Mais la spécificité des neurones individuels est telle que l'on ne peut s'en tenir à des connexions symboliques. L'introduction des modèles connexionnistes distribués par D. Rumelhart, P. Smolensky, J. L. Mac Clelland et G. Hinton dans les années 1980 repose sur l'idée que le comportement du système cognitif n'est pas régi par des règles, mais il est seulement (approximativement) décrit par des règles. Cette approche approximationniste présente l'avantage de répondre aux conditions du système sans tomber dans une description mécanique et universelle. On voit bien quelle est la stratégie générale des approximationnistes : ils cherchent à simuler la performance de systèmes symboliques en utilisant des configurations de connectivité plutôt que des règles et des représentations symboliques. P. S. Churchland ne peut utiliser de tels modèles qui ne sont qu'un intermédiaire entre ceux du cerveau computationnel et ceux des représentations symboliques. Les approximationnistes ne vont pas assez loin dans la description du cerveau.

Le projet de la neurophilosophie américaine est de

construire un connexionnisme neuronal plutôt qu'un connexionnisme paralléliste. Pour elle l'étude des lésions des fonctions mnésiques du cerveau repose sur des principes qui sont différents de ceux des ordinateurs. La conservation et la diffusion en réseau de l'information dans le système nerveux exigent une modélisation spécifique de l'action neuronale. Les réseaux distribués n'admettent pas l'individualisation spécifique des neurones à cause de la nécessité paradigmatique de proposer un système cohérent et unitaire dans le traitement parallèle par réseaux informatiques. Il faut distinguer ici deux approches : dans les réseaux localistes, on assigne à chaque concept une unité ; dans les réseaux distribués, la représentation de chaque concept est distribuée sur plusieurs unités. J. A. Anderson a bien aperçu le risque des réseaux localistes de reposer uniquement sur un processus externe d'interprétation sémantique par lequel la correspondance entre l'unité et le concept est établie par le langage de l'investigateur. La modélisation apparaîtrait subjective et éphémère. En supposant que le cerveau constitue un réseau de neurones, le but est de modéliser les phénomènes cognitifs dans des systèmes présentant certaines propriétés fondamentales de ce réseau.

Les réseaux formels de neurones

Ils furent introduits par l'article de Warren Mac Culloch (1892-1969) et Walter Pitts publiés en 1943 : ils

formaient un modèle simple d'unité computationnelle analogue à des neurones. L'équivalence est supposée entre la réponse d'un neurone et une proposition qui énonce ses stimuli appropriés. L'augmentation de la mémoire du réseau permettait de réaliser la machine de Turing universelle (1937). Par cette convergence de la logique moderne, de la neurophysiologie et de la simulation automatique du calcul, W. S. Mac Culloch estime que « le cerveau est une machine logique » [1]. Cette méthode computationnelle, engagée théoriquement par Alan M. Turing (1912-1954), se réalise à partir de l'étude du cerveau. En 1950, comme le précise Alain Tête, deux voies s'offrent à la recherche : « D'une part, celle de l'intelligence artificielle qui vise à compliquer indéfiniment la compétence de ses automates ; de l'autre, celle de la neuropsychologie cognitive dont les neurones formels devront simuler les neurones "humides", donc incarner les processus cognitifs dans un corps humain virtuel. » [2] Seulement la théorie formelle des réseaux de neurones ne propose qu'une modélisation des fonctions de neurones minimales.

Les réseaux de neurones réels

Ils sont établis à travers les travaux de J. von Neumann (1903-1957). À partir de la voie ouverte par W. Pitts et W. Mcac Culloch il reste à étudier la nature concrète des réseaux à l'intérieur même du fonctionnement cérébral [3]. La technologie du silicium et des circuits

intégrés à haute densité a produit une remise en cause qualitative des architectures d'ordinateurs, au point de permettre aux ordinateurs de ressembler de plus en plus aux tissus du cerveau.

J. von Neumann décrit ces réalisations dans *L'ordinateur et le cerveau* (en 1957 juste avant sa mort) : l'analogie entre l'ordinateur et le cerveau se trouve pensée ici dans le caractère digital du fonctionnement du système nerveux. Le composant du système est la cellule nerveuse, le neurone dont la fonction est de produire et de diffuser un influx nerveux. « Il est clair qu'on peut considérer les impulsions nerveuses comme des marqueurs (à deux valeurs)... l'absence d'une impulsion représente alors une valeur (disons le chiffre binaire 0) et la présence d'une impulsion représente l'autre valeur (disons le chiffre binaire 1)... » [\[4\]](#) Il distingue un hardware cérébral qui représente la structure tandis que le software est constitué par les apprentissages sélectifs, à l'image de la différence entre la possibilité du langage et l'apprentissage d'une langue. Cependant, J. von Neumann ne confond pas l'analogie avec une homologie : s'il est vrai que « le système nerveux envisagé comme un ordinateur » a une même structure arithmétique et logique, il faut reconnaître que le langage du cerveau n'est pas le langage des mathématiques. P. S. Churchland fait remarquer que l'analogie introduite par J. von Neumann repose sur l'hypothèse selon laquelle l'ordinateur et l'esprit-cerveau ont trois niveaux d'organisation :

le niveau du contenu	LA SÉMANTIQUE
le niveau de l'algorithme	LA SYNTAXIQUE
le niveau de l'implémentation structurelle	LA MÉCANIQUE

La critique de P. S. Churchland repose ici sur le caractère simplificateur de l'analogie de l'ordinateur. En effet, la complexité des processus cérébraux suppose « beaucoup de niveaux d'organisation entre le niveau le plus élevé et le niveau des dynamiques intracellulaires » [5]. La réduction au niveau 1, celui de la mécanique, de toutes les opérations structurelles tend à accréditer l'idée d'une opposition entre la matière et la forme, entre la structure et le contenu. Or, une telle conception tendrait à séparer la représentation de ses conditions neurobiologiques.

P. S. Churchland préfère distinguer les niveaux d'organisation suivant :

le comportement
le circuit
l'assemblée cellulaire
la synapse
la cellule
la membrane

La complexité des réseaux interdit la description exhaustive, pour le moment, du passage continu de la membrane au comportement.

La théorie des réseaux réels de neurones analyse deux catégories principales d'ordinateur, les ordinateurs à procédure analogique, et les ordinateurs à procédure digitale. Comme dans une machine analogique « chaque nombre est représenté par une quantité physique appropriée, dont la valeur, mesurée selon une unité définie au préalable, est égale au nombre en question » [6], le fonctionnement du système nerveux est digital. C'est la complexité de l'influx nerveux, qui diffuse l'information des cellules nerveuses ou neurones, qui rend impossible le modèle analogique car on ne peut associer directement à chaque neurone une valeur fonctionnelle directe. « Dans une machine digitale décimale, chaque nombre est représenté... comme une suite de chiffres décimaux. Chaque chiffre décimal est, quant à lui, représenté par un système de marqueurs. » [7] Ce sont seulement les impulsions nerveuses qui sont considérées comme des marqueurs que J. von Neumann va distinguer en deux valeurs : « L'absence d'une impulsion représente alors une valeur (disons le chiffre binaire 0) et la présence d'une impulsion représente l'autre valeur (disons le chiffre binaire 1). » [8]

Autour de la mémoire, J. von Neumann illustre sa description du cerveau comme machine digitale. La présence d'une mémoire est « une hypothèse que justifie et confirme toute notre expérience avec les ordinateurs et les automates » [9]. Le mathématicien s'appuie sur l'établissement expérimental de

l'existence dans le cerveau de plusieurs types de mémoire de temps d'accès et de durée de conservation différents.

On peut établir le tableau suivant à propos de la mémoire :

<i>Machines digitales composées d'organes actifs</i>	<i>La mémoire du système nerveux</i>
Les capacités de mémoire sont de l'ordre de 10^5 à 10^6 bits.	Capacité supérieure : 1 010 neurones $\times$ 14 impressions digitales par seconde pendant 60 ans = $2,8 \times 10^{20}$ bits.
« Flips-flops » : paires de tubes à vide qui se déclenchent et se contrôlent mutuellement.	Les systèmes de cellules nerveuses se stimulent les uns les autres de façon cyclique.
Formes de mémoire techniquement différentes des organes actifs de base de la machine.	Même si pas de connaissance des types d'entités physiques, mémoire de très grande capacité associée aux organes de base.

J. von Neumann, même si la différence quantitative existe, établit une analogie de structure dans les principes de la mémoire. En décrivant la structure logique du système nerveux, il le conçoit comme le fonctionnement d'un des automates complexes. Il convient du caractère différent du langage utilisé par le système nerveux central et le langage des mathématiques qui ne parvient qu'à fournir une série de données statistiques des états neurobiologiques. Si bien que « quand nous parlons des mathématiques,

nous sommes peut-être en train de parler d'un langage secondaire, bâti sur le langage premier réellement utilisé par le système nerveux central. Ainsi les formes extérieures (visibles) de nos mathématiques ne sont-elles pas absolument pertinentes pour évaluer quels sont les véritables langages mathématiques et logiques utilisés par le système nerveux central » [\[10\]](#) (op. cit., p. 80-81). Sans renoncer au modèle mathématique de la structure logique du système nerveux central, ce texte exprime la limite de l'arithmétique pour modéliser le cerveau à partir de ces unités, processus et fonctionnement. Il faudra attendre la reprise des travaux de W. Pitts et W. Mac Culloch (1947) sur les manières dont un réseau pouvait effectuer des tâches de reconnaissances des formes.

Les réseaux à couches et multicouches

Les travaux de Frank Rosenblatt (1958) sont cités en référence mais ne constituent pas, pour la neurophilosophie américaine, un modèle suffisant. Le perceptron élémentaire, nom du système, comportait les entrées fournies dans les quatre unités sensorielles horizontales, tandis que les sorties sont produites par les deux unités motrices situées en bas du système. L'intersection représente les synapses. « Il compléta les réseaux de W. Mac Culloch et W. Pitts en rendant les forces de connexions (généralement appelées les poids) entre unités continues au lieu de

binaires, et en introduisant des procédures permettant de changer ces poids, ce qui permet d'apprendre aux systèmes à changer leurs réponses. » [11] F. Rosenblatt expliquait en 1958 qu'« en tant qu'analogue au cerveau biologique, le perceptron, ou plus précisément la théorie de la séparabilité statistique, paraît mieux à même de remplir les exigences posées par l'explication fonctionnelle du système nerveux qu'aucun système jamais proposé » [12]. Pourtant, s'il est vrai que F. Rosenblatt augmente la proportion d'unités activées par entrée par le jeu des interactions, le modèle reste encore bien inférieur à la plasticité fonctionnelle du cerveau. Les couches ne sont pas déterminées par rapport à la réalisation de fonctions. Elles se modifient au fur et à mesure de l'activité neurobiologique.

– La microstructure de la cognition dans le traitement parallèle de distribution (Parallel Distributed Processing) est formulée par J. L. Mac Clelland et D. E. Rumelhart dès 1981. Ils proposent un modèle d'activation interactive qui illustre la manière dont un réseau multicouches peut reconnaître des formes visuelles. L'avantage de ce modèle est de concilier l'architecture de réseau au plan global et la dynamique fonctionnelle à l'intérieur du réseau qui fait appel à des multicouches. Les dimensions synchrone et diachronique existent lors de la réalisation d'une activité cognitive comme la reconnaissance visuelle de formes. Le réseau peut à la fois reconnaître la couche supérieure et les couches intermédiaires par le biais

des points dynamiques d'intersection. Comme réseau interactif, la propagation de l'activation continuera sur un grand nombre de cycles de traitement. P. S. Churchland repère l'avantage de ce type de réseau qui pourra être en connexion directe avec les activités variables du cerveau. La neurophilosophe ne parle pas de couches, ou de niveaux, mais de sous-réseaux (subnetwork) établissant par là une règle de composition interne au système nerveux. Il y a des réseaux principaux dans lesquels d'autres réseaux assurent la continuité fonctionnelle. La différence de point de vue ou de niveau provient de cette composition structurelle entre réseaux et sous-réseaux.

– Analogie thermodynamique et stratégie probabiliste de réseaux, telles sont les données du modèle présentées par G. Hinton et T. J. Sejnowski. Les couches ne sont plus nécessaires dans le modèle élaboré par G. Hinton et T. J. Sejnowski de 1983. Au moment de la rédaction de *Neurophilosophy*, P. S. Churchland ne connaît pas encore les travaux de T. J. Sejnowski et F. Rosenberg publiés en 1986 « NetTalk : A parallel network that learns to read aloud », travaux qui vont servir d'ossature à sa thèse du cerveau computationnel à partir de 1988.

Avec *Neurophilosophy*, P. S. Churchland commence à critiquer cette répartition qui continue, comme les modèles computationnels précédents, à séparer les hauts niveaux d'organisation des bas niveaux. L'indépendance entre le niveau de l'implémentation

physique et les deux autres niveaux ne convient pas avec le modèle sensori-moteur de P. S. Churchland qui cherche à préciser les fonctions supérieures du cerveau à partir des seuls états neurobiologiques. « Le modèle à fournir pour une théorie de certaines fonctions cérébrales telle que le patron de reconnaissance, il sera nécessaire de prendre ensemble les données physiologique et anatomique finement graminées et les données psychologiques de hauts niveaux. Dans le cas de la théorie du réseaux de tenseur, nous devons évaluer comment il explique et unifie les données, comment il peut être essayé, et s'il est assez riche pour suggérer de nouveaux essais expérimentaux. Pour être clair, il nous est apparu que les stratégies de modélisation parallèle sont une cuvée trop récente et que ce que nous voyons jusqu'ici n'est réellement qu'au début. » [\[13\]](#)

Au travers des mécanismes neurobiologiques que P. S. Churchland trouve donc les données pour développer le modèle computationnel du cerveau. La théorie des réseaux de tenseur n'aura été qu'un moyen pour P. S. Churchland de verser la neurophilosophie dans le connexionnisme neurocomputationnel [\[14\]](#)..

Notes

[\[1\]](#) F. Blayo et M. Verleysen, Les réseaux de neurones artificiels, Paris, puf, coll. « Que sais-je ? », 1996.

- [2] W. S. Mac Culloch et W. Pitts (1943), Un calcul logique des idées immanentes dans l'activité nerveuse, trad. franç. A. Péliissier, dans Sciences cognitives. Textes fondateurs (1943-1950), Éd. A. Tête, Paris, puf, 1995, p. 57-69, complété par la présentation du modèle par Marvin L. Minsky, p. 69-91
- [3] W. S. Mac Culloch (1949), Du cerveau comme calculateur, trad. franç. A. Péliissier, dans Sciences cognitives. Textes fondateurs (1943-1950), Éd. A. Tête, Paris, puf, 1995, p. 189-214.
- [4] A. Tête, Présentation de l'article d'Alan M. Turing, Machines à calculer et intelligence, trad. franç. A. Péliissier, dans Sciences cognitives. Textes fondateurs (1943-1950), Éd. A. Tête, Paris, puf, 1995, p. 253.
- [5] J. von Neumann (1948), La théorie générale et logique des automates, trad. franç. A. Péliissier, dans Sciences cognitives. Textes fondateurs (1943-1950), Éd. A. Tête, Paris, puf, 1995, p. 93-153.
- [6] J. von Neumann (1956), L'ordinateur et le cerveau, trad. P. Engel, Paris, La Découverte, 1992, p. 15.
- [7] Ibid., p. 18
- [8] Ibid., p. 48.
- [9] Ibid., p. 63.
- [10] Ibid., p. 80-81.
- [11] Ibid., p. 48.
- [12] P. S. Churchland, Neurophilosophy, op. cit., p. 359.
- [13] W. Bechtel et A. Abrahamsen (1991), Le connexionnisme et l'esprit. Introduction au traitement parallèle par réseaux, trad. franç. J. Proust, Paris, La Découverte, 1993, p. 13-14.

[14] F. Rosenblatt (1958), The perceptron. A probabilistic model for information storage and organization in the brain, *Psychological Review*, 65, p. 368-408.

Chapitre V

Vers la philosophie du corps cérébré

Plusieurs travaux existent aujourd'hui qui tentent de fonder une intrascience entre phénoménologie et neurosciences. Le cerveau peut aussi être compris à partir du vécu du corps propre dès sa constitution plastique [\[1\]](#), en raison des interaction du corps avec son environnement.

Ainsi Le cerveau psychologique (B. Andrieu, 2003) n'est plus le cerveau décrit par la neurophysiologie et la neurologie : la description des composants physiques du cerveau ne tient pas compte des états psychologiques ; car la réduction analyse ces composants en eux-mêmes et dans leurs relations. La psychologie deviendrait inutile par ce type de réduction de l'esprit au cerveau au point que l'on devrait l'éliminer. La physiopsychologie d'Auguste Comte, directement inspirée par la phrénologie de Gall, s'oppose à la psychologie traditionnelle des facultés : selon Kant la différence entre sensibilité, entendement et raison justifiait une description proprement psychologique de l'activité mentale de l'esprit humain.

Cette psychologie des facultés est le contraire de la physiopsychologie des fonctions démontrée par le positivisme, sans pour autant que ce conflit des descriptions de l'esprit humain n'établisse de communication entre elles.

Si l'on peut décrire l'esprit sans le cerveau ou, au contraire, réduire l'esprit au cerveau, nous supposons que la question de la communication entre le cerveau et l'esprit reste le postulat de ce conflit des descriptions plutôt que sa conséquence. L'étude du cerveau n'a pas toujours été, comme pourrait faire croire la naissance et le développement des neurosciences, une spécialisation naturaliste des sciences de l'esprit. La localisation des aires du langage et de la mémoire par Paul Broca et Karl Wernicke aura pu établir un lien de cause à effet à partir de 1860 entre les facultés de l'esprit et le fonctionnement du cerveau. Mais le lien entre le cerveau et l'esprit a toujours été abordé à travers des objets comme l'hystérie, la mélancolie ou la folie. Ces objets ne sont pas seulement des intermédiaires entre deux disciplines constituées comme la neurologie et la philosophie. Ils sont problématiques par essence car ils manifestent des phénomènes inexpliqués comme l'influence des humeurs, la communication des esprits animaux, ou encore la circulation des fluides.

La chair du cerveau (B. Andrieu, 2002) considère que le cerveau n'est pas un organe objectif qui serait séparé de la vie du corps, c'est-à-dire de sa constitution, de

son métabolisme et de son adaptation au milieu. Vivant, le cerveau se modifie sans cesse tant dans son organisation que dans sa spécialisation. Plutôt qu'un récepteur simple, le cerveau est soumis à ses possibilités de plasticité et de réadaptation : cette mobilité des réseaux neuronaux dynamise le cerveau tant dans la qualité de ses états mentaux que dans la communication des neurotransmetteurs. Déterminé par des facteurs génétiques, au cours de son développement et dans ses régulations, le cerveau n'est pas libre : le cerveau doit trouver une homéostasie neurofonctionnelle en synthétisant la contradiction entre la part des gènes et la part de l'histoire. Cette synthèse, nous proposons de la désigner sous le terme de la chair du cerveau.

En alliant biologie et phénoménologie, la notion de chair du cerveau ou celle de cerveau incarné rejoint le travail fondateur de Francisco Varela : en proposant de partir du corps vécu pour rendre compte de la cognition, la notion d'incarnation n'a plus ce sens dualiste de la tradition chrétienne. L'esprit est désormais présent par le corps qui le produit. Selon F. Varela, une voie moyenne doit être maintenant proposée qui étudie la cognition non comme une reconstitution d'un monde extérieur prédonné (réalisme) ou une projection d'un monde intérieur prédonné (idéalisme). La cognition est pour lui une action incarnée : « Par le mot incarnée, nous voulons souligner deux points : tout d'abord, la cognition dépend des types d'expériences qui découlent du fait d'avoir un corps doté de diverses

capacités sensori-motrices ; en second lieu, ces capacités individuelles sensori-motrices s'inscrivent elles-mêmes dans un contexte biologique, psychologique et culturel plus large. » [2]

Trouvant le lien, si recherché par Maurice Merleau-Ponty entre biologie et phénoménologie, l'Énaction démontre comment les structures cognitives émergent des schèmes sensori-moteurs récurrents qui guident l'action par la perception. La séparation entre motricité et perception conduit à une description des mécanismes cérébraux et à une interprétation, nous l'avons vu, neurocomportementale. Ainsi le point de référence de la perception n'est plus un monde prédonné mais « la structure sensori-motrice du sujet (la manière dont le système nerveux relie les surfaces sensorielles et motrices). C'est cette structure – la façon dont le sujet percevant est inscrit dans un corps –, plutôt qu'un monde préétabli, qui détermine comment le sujet peut agir et être modulé par les événements de l'environnement » [3]. L'inscription corporelle de l'esprit doit non seulement prendre en compte l'énaction, mais aussi l'individuation de la chair. La différence entre l'énaction (action incarnée) et l'individuation de la chair est de point de vue : comment les structures cognitives permettent ou non au sujet de percevoir ses incorporations constitutives.

Ce courant, celui du mouvement neurocognitif proprement dit, prouve combien le mouvement des représentations dans le corps et du corps lui-même

est bien la conséquence épistémologique de l'élaboration de modèles du mouvement. La biomécanique, inventée par Étienne Marey et Georges Demeny[4], est remplacée aujourd'hui par la simulation neurobiologique : critiquant la neurophilosophie, Alain Berthoz rappelle que le cerveau n'est pas un tenseur soulignant ainsi l'échec du neurocomputationnisme trop attaché aux modèles géométriques et non dynamique du mouvement cérébral. En supposant une chaîne continue de transformation de la sensation à la commande motrice par le moyen des systèmes sensori-moteurs ou des réflexes, la neurophilosophie, bien qu'elle prenne en compte les concepts de modèle interne comme celui du mécanisme d'anticipation, « ne fait aucune place à l'influence de l'action sur le traitement sensoriel » [5]. Le modèle interne suppose que le cerveau construit des stratégies en feedback et des intentions-protections d'action. Le principe de conservation renouvelle la perception jusqu'à la mise en action.

Comme action interne simulée, la théorie motrice de la perception de Berthoz valorise l'action directe de l'organisme sur les récepteurs ; cette perception active de la sensation suppose des préperceptions simulatrices plutôt que des représentations : le cerveau est un simulateur car l'ensemble de l'action est joué par « des modèles internes de la réalité physique qui ne sont pas des opérateurs mathématiques mais de vrais neurones dont les propriétés de forme, de résistance, d'oscillation, d'amplification » [6]. La

kinesthésie, le sens du mouvement, établirait des relations avec les cinq capteurs sensoriels traditionnels. Dans la proprioception, le cerveau utilise des motoneurones pour moduler l'information sensorielle et simuler le mouvement afin de l'adapter à l'action. À cette mobilité du cerveau s'ajoute une réorganisation des représentations corticales tant du thalamus et du cortex si un accident vient modifier le schéma corporel.

En faisant de la décision une propriété fondamentale du système nerveux, Berthoz fait de la délibération un inconscient cérébral. Il faut donc supposer deux corps : « L'action procède d'un dialogue entre le corps et son double ; la délibération et la décision expriment ce dialogue fondamental. Nous avons deux corps, le corps physique et le corps mental. Le corps mental est constitué de tous les modèles internes qui constituent les éléments du schéma corporel et permettent au cerveau de simuler, d'émuler la réalité. C'est le corps que nous percevons lorsque nous rêvons. » [7] La création au sein du cerveau d'un double cérébral qui est capable d'anticiper et de prédire les conséquences des actions en fonction de la mémoire et de l'analyse des données fournies dans la perception.

Si la perception est pour l'action, l'accès indirect au cerveau vivant par le corps vécu interdit une ontologie directe qui décrirait l'être sans l'expérience corporelle. M. Merleau-Ponty rappelle que si « nous disons : philosophie de la nature et non philosophie de l'être »

c'est « parce qu'il s'agit de rendre compte de notre expérience de l'être » [8]. Ce caractère indirect admet un « décisionisme physiologique » [9] qui reconfigure fonctionnellement les réseaux neuronaux au fur et à mesure de l'interaction du corps avec le monde avant même que la conscience en ait conscience. Cette cécité de la conscience de son cerveau in vivo et en action rend l'activité sensori-motrice centrale dans la perception en mouvement de l'environnement par l'interaction du corps cérébré.

Pas seulement cérébré le corps est aussi cérébrant par la reconfiguration dynamique des boucles internes (à la Edelman ou à la Llinas) et par le caractère autopoïétique du système nerveux (cf. Varela). La temporalité de la constitution interactive doit tenir compte des mécanismes d'anticipation et de prédiction présent dans les récepteurs sensoriels. Ce n'est pas le monde qui vient au sens, ce sont les sens qui s'immondent : la perception anticipe sur « la faisabilité d'un geste ou d'une action » [10] avant que le corps ne le mette en œuvre. Cette anticipation inconsciente compense les délais de l'excitation motrice. Comme prédicteur « l'organisme doté d'un cerveau procéderait, en toutes circonstances, à une anticipation centrale par rapport à la détermination de la réceptivité des capteurs sensoriels périphériques » [11]. Une théorie de la constitution du modèle interne du monde par l'organisme lui-même qui perçoit le monde suppose un renversement épistémologique : ce n'est plus le monde qui informe l'organisme, c'est le corps cérébré

qui définit son environnement d'action : O'Regan et Noé le font par le concept de « contingence sensori-motrice » [12] car le stimulus détermine les inputs sensoriels, pour la perception active, il détermine une loi sensori-motrice reliant les inputs sensoriels du système nerveux à ses outputs moteurs.

L'objet extérieur est bien une réalité indépendante de notre cerveau, il est un objet de la perception indépendant de nous mais dont la perception ne nous donne qu'une connaissance subjective. Le cerveau transcendantal se distingue de l'objet transcendant à nous et de l'immanence des états corporels vécus. Comme opérateur constituant, le cerveau prédictif anticipe et structure a priori l'expérience corporelle immanente dans la mesure où le flux même de cette expérience rend impossible son aperception. Le cerveau, en tant que sujet transcendantal, est « agent effectuant les actes de synthèse du divers qui lui donnent accès au monde physique » [13]. Les modèles internes, comme architecture fonctionnelle de la simulation du mouvement par les réseaux neuronaux, seraient une « implémentation neuronale de l'a priori » [14].

Notes

[1] C. Malabou, Que faire de notre cerveau ?, Paris, Bayard, 2004, p. 35-67.

- [2] F. J. Varela, E. Thompson et E. Rosch (1991), L'inscription corporelle de l'esprit, Paris, puf, 1993, p. 234.
- [3] Ibid., p. 135.
- [4] C. Pociello, La science en mouvements, Paris, puf, 1999, p. 157-199, D. de Font-Réaulx, T. Lefebvre et L. Mannoni, É. J. Marey. Actes du Colloque du centenaire, Éd. Musée d'Orsay, 2006.
- [5] A. Berthoz, Le sens du mouvement, Paris, O. Jacob, 1997, p. 57.
- [6] Ibid., p. 28.
- [7] A. Berthoz, La décision, Paris, O. Jacob, 2003, p. 169-170.
- [8] M. Merleau-Ponty, La nature ou le monde du silence, bn, vol. VI (106) v (B), 1957, cité par Emmanuel Saint-Aubert, Vers une ontologie indirecte, Paris, Vrin, 2006, p. 116.
- [9] A. Berthoz et J.-L. Petit, Physiologie et phénoménologie de l'action, Paris, O. Jacob, 2006, p. 63.
- [10] Ibid., p. 71.
- [11] Ibid., p. 35.
- [12] K. O'Regan et A. Noé, A sensorimotor account of vision and visual consciousness, Behavioral and Brain Sciences, 24, 2001, p. 939-1011.
- [13] A. Berthoz et J.-L. Petit, Physiologie et phénoménologie de l'action, op. cit., p. 101.
- [14] Ibid., p. 109.

Conclusion

La neurophilosophie n'est pas nouvelle. Elle doit son origine au rêve d'une science unifiée que poursuivent tant les sciences de la nature que les sciences humaines. Dès le fondement de la biologie au ^{xviii}^e siècle et par son renouvellement moléculaire avec la génétique au ^{xx}^e siècle (il en va de même pour la phrénologie et le renouveau épistémologique apporté par les neurosciences) est engagée l'élimination de l'Homme. L'idée n'est pas originale. Michel Foucault avait annoncé, en 1966, « peut-être la fin prochaine » de l'homme. Trente ans plus tard, et même si les travaux historiques remettent en cause le caractère fondateur du ^{xviii}^e siècle, il n'en reste pas moins que le domaine de recouvrement des sciences humaines n'aura pas été suffisant pour assurer définitivement la permanence de l'homme. Ce que nous désignons, avec d'autres, sous le vocable d'homme naturel [\[1\]](#), c'est moins sa fin interne aux sciences humaines que l'élimination des sciences humaines de la question de l'essence de l'Homme par les sciences de la vie.

La neurophilosophie est cet arbre réductionniste de la forêt actuelle des matérialismes naturalistes. La philosophie naturaliste reste le centre autour duquel une réflexion doit pourtant s'organiser, sans confondre

chaque domaine propre, la rencontre de la philosophie et des sciences de la vie. La neurophilosophie préfigure, dès 1980-1986, la recomposition du paysage intellectuel : les neurosciences et la génétique du développement, associées à une critique active du darwinisme, forceront les philosophes à penser le corps humain. Faute de cette ouverture réflexive, beaucoup de philosophes voudraient défendre un dualisme modernisé par les sciences cognitives. Or, celles-ci auront définitivement replacer la philosophie au cœur des interrogations fondamentales : Comment le corps pense ? Quelles sont les différences entre l'homme et l'animal ? Le développement de la matière suffit-il pour l'étude des formes ?

Ces interrogations, dont les réponses nous parviennent déjà, participent de la nouvelle philosophie du corps du ^{xxi}^e siècle plutôt qu'au bilan du ^{xx}^e siècle. Maladroite la neurophilosophie a cru pouvoir en réduisant la philosophie aux neurosciences fonder une réduction du corps pensant au cerveau. Elle est seulement le signe révélateur d'une tension contemporaine. La fin des idéologies favorise le naturalisme scientifique comme une évidence qu'il faudrait accepter.

Recouverts de philosophie, des apprentis voudraient expliquer la nature humaine par les seules séquences génétiques ou les flux de neurotransmetteurs. La complexité de l'homme dépassera toujours ce comportementalisme. Derrière le spectacle des

médiateurs des sciences de la vie, il convient, pour qui veut écrire une philosophie du sujet incarné, et pas seulement du cerveau, de distinguer le bon grain de l'ivraie. Reste à travailler à une philosophie du corps, plutôt qu'à une neurophilosophie.

Notes

[1] B. Andrieu, L'homme naturel de la fin promise des sciences humaines, Lyon, pul, 1998.

Bibliographie

Andrieu B., Brains in the flesh : Prospects for a neurophenomenology, Janus Head, 9. 1, 2006, p. 135-155.

Andrieu B., Herbert Feigl. De la physique au mental, Paris, Vrin, 2006.

Akins K. A., Neurophilosophy. Book review, The Journal of Philosophy, vol. LXXXVII, n° 2, février 1990, p. 93-103.

Berthoz A. et Petit J.-L., Physiologie et phénoménologie de l'action, Paris, O. Jacob, 2006.

Bickle J., Mental anomaly and the new mind-brain reductionism, Philosophy of Science, 59, 1992, p. 217-230.

Churchland P. M., Eliminative materialism and propositional attitudes, Journal of Philosophy, LXXVIII, 1979.

Churchland P. S., Neurophilosophy, Toward a Unified Science of the Mind-Brain, mit, 1986, 546 p. ; trad. franç., Paris, puf, 1999.

Churchland P. S., Brain-Wise Studies in Neurophilosophy, Bradford Book, 2002.

Churchland P. S. et Christen Y., Neurophilosophy and Alzheimer's Disease, Berlin, Springer-Verlag, 1992.

Davis W. J., Neurophilosophical reflections on central nervous pattern generators, The Behavioral and Brain Sciences, 3, 1980, p. 543-544.

Debru C., Neurophilosophie du rêve, Paris, Hermann,

1990, rééd. 2006, 398 p..

Dupuis M., Les neurones en personne. Six propositions pour une neurophilosophie, Liège, Université libre, 1989, 29 p..

Engel P., Philosophie et psychologie, Paris, Gallimard, 1996, p. 344-354.

Gallagher S., How the Body Shapes the Mind, Oxford University Press, 2005.

Gillet G. R., The neurophilosophical of pain, Philosophy, 66, 1991, p. 191-205.

Kiamie K. A., Neurophilosophy and the logic-causation argument, Dialogue, avril 1990, p. 39-42.

Kitcher P., Model of reduction, Nature, vol. 332, 31 juillet 1986, p. 413-414.

Houde O., Mazoyer B. et Tzourio-Mazoyer N., Cerveau et psychologie, Paris, puf, 2002.

Jacob P., Pourquoi les choses ont-elles un sens ?, Paris, O. Jacob, 1997.

Lecourt D., Le cerveau des philosophes, À quoi sert la philosophie. Des sciences de la nature aux sciences politiques, Paris, puf, 1993.

Llinas R. R., I of the Vortex. From Neurons to Self, Bradford Book, 2001.

McCauley R. N., The Churchlands and their Critics, Blackwell Pub, 1996.

Malabou C., Que faire de notre cerveau ?, Paris, Bayard, 2004.

Missa J.-N., L'esprit-cerveau, Paris, Vrin, 1993.

Naccache L., Le nouvel inconscient, Paris, O. Jacob, 2006.

Pacherie É., Naturaliser l'intentionnalité, Paris, puf, 1993.

Perez de Laborda A., El problema mente cerebro : a proposito de, P. S. Churchland, La Ciudad de Dios, Esp., 202, n° 2, 1989, p. 449-475.

Poirel C., Le cerveau et la pensée. Critique des fondements de la neurophilosophie, Paris, L'Harmattan, 1997.

Sereno M., A program for the neurobiology of mind, Symposium sur Neurophilosophy de, P. S. Churchland, Inquiry. An Interdisciplinary Journal of Philosophy, 29, 1986, p. 217-240.

Skarda C. A., Explaining behavior : Bringing the brain back in Symposium sur Neurophilosophy de, P. S. Churchland, Inquiry. An Interdisciplinary Journal of Philosophy, 29, 1986, p. 187-202.

Smith A., Brain-mind philosophy, Symposium sur Neurophilosophy de, P. S. Churchland, Inquiry. An Interdisciplinary Journal of Philosophy, 29, 1986, p. 203-216.