

La pente du naturalisme

Daniel Andler

Citer ce document / Cite this document :

Andler Daniel. La pente du naturalisme. In: Intellectica. Revue de l'Association pour la Recherche Cognitive, n°69, 2018/1-2. Les Neurosciences au sein des Sciences de la Cognition entre Neuroenthousiasme et Neuroscepticisme. pp. 149-165;

doi : <https://doi.org/10.3406/intel.2018.1876>

https://www.persee.fr/doc/intel_0769-4113_2018_num_69_1_1876

Fichier pdf généré le 25/03/2022

Résumé

Les sciences cognitives sont passées, en un demi-siècle, d'une configuration dans laquelle les neurosciences ne jouaient au mieux qu'un rôle marginal, à une configuration dans laquelle elles occupent une position centrale, voire hégémonique. Comment expliquer un pareil renversement ? L'article complète les explications courantes – difficultés du paradigme classique, progrès spectaculaire des neurosciences nourri par le développement de la neuro-imagerie... – par la prise en compte de la pression exercée par l'exigence de naturalisation : l'étude du cerveau apparaît aujourd'hui, comparée à l'approche classique, comme une voie rapide et sûre vers la naturalisation de l'esprit. Il s'agit là d'une fausse évidence, qui repose sur l'idée que nous savons en quoi consiste la naturalisation d'un domaine avant d'en avoir largement engagé la réalisation scientifique effective. Une fois dissipée cette illusion, on retrouve une perspective équilibrée sur le domaine, tant dans son développement historique que dans son état présent.

Abstract

The Naturalistic Incline. Cognitive science has gone, within a halfcentury, from a configuration in which neuroscience held at most a marginal role, to one in which it occupies center stage, if not the entire stage. How can one account for such a reversal? The paper aims at complementing the usual explanations – the shortcomings of classical cognitive science, the momentous development of neuroscience fueled by neural imagery... – by the consideration of the pressure exerted by the demand to naturalize: the study of the brain appears today, compared to the classical approach, as a quick and sure path towards naturalizing the mind. Far from being obvious, this is a mistake based on the erroneous idea that we know what it is to naturalize a given realm before we get down to actually develop the science that accomplishes it. Once this illusion evaporates, a more balanced perspective on the field, both in its historical unfolding and in its present state, is restored.

La pente du naturalisme¹

Daniel ANDLER*

RÉSUMÉ. Les sciences cognitives sont passées, en un demi-siècle, d'une configuration dans laquelle les neurosciences ne jouaient au mieux qu'un rôle marginal, à une configuration dans laquelle elles occupent une position centrale, voire hégémonique. Comment expliquer un pareil renversement ? L'article complète les explications courantes – difficultés du paradigme classique, progrès spectaculaire des neurosciences nourri par le développement de la neuro-imagerie... – par la prise en compte de la pression exercée par l'exigence de naturalisation : l'étude du cerveau apparaît aujourd'hui, comparée à l'approche classique, comme une voie rapide et sûre vers la naturalisation de l'esprit. Il s'agit là d'une fausse évidence, qui repose sur l'idée que nous savons en quoi consiste la naturalisation d'un domaine *avant* d'en avoir largement engagé la réalisation scientifique effective. Une fois dissipée cette illusion, on retrouve une perspective équilibrée sur le domaine, tant dans son développement historique que dans son état présent.

Mots-clés : Histoire, paradigme classique, neurosciences, naturalisme.

ABSTRACT. The Naturalistic Incline. Cognitive science has gone, within a half-century, from a configuration in which neuroscience held at most a marginal role, to one in which it occupies center stage, if not the entire stage. How can one account for such a reversal? The paper aims at complementing the usual explanations – the shortcomings of classical cognitive science, the momentous development of neuroscience fueled by neural imagery... – by the consideration of the pressure exerted by the demand to naturalize: the study of the brain appears today, compared to the classical approach, as a quick and sure path towards naturalizing the mind. Far from being obvious, this is a mistake based on the erroneous idea that we know what it is to naturalize a given realm *before* we get down to actually develop the science that accomplishes it. Once this illusion evaporates, a more balanced perspective on the field, both in its historical unfolding and in its present state, is restored.

Keywords: History, classical paradigm, neuroscience, naturalism.

Intellectica est né dans le milieu scientifique de l'Association pour la recherche cognitive (ARC), au début des années 1980. La charte fondatrice de l'ARC était le « rapport Lentin », qui fixait le périmètre et la mission des récemment baptisées sciences cognitives². Les sciences du cerveau en étaient

¹ Cet article reprend pour une bonne part certaines idées de mon livre récent, *La Silhouette de l'humain* (Andler, 2016), sans néanmoins en constituer un résumé ni une série d'extraits.

* Département d'études cognitives, Ecole normale supérieure, Université PSL & Sciences, normes, décision, Sorbonne Université. 33, rue Croulebarbe, 75013 Paris. daniel.andler<at>ens.fr.

² Souvent présentées comme un produit d'importation. En réalité, sous des intitulés différents, les sciences cognitives résultent d'un mouvement d'idées largement international, auquel la France a participé, parfois même en position pionnière. Voir notamment Andler, 2006a.

explicitement exclues. Nous sommes au début des années 1980. Aujourd'hui, les sciences cognitives sont sommées de changer d'intitulé : leur vrai nom serait « neurosciences cognitives ». En moins de 40 ans, nous sommes passés d'un extrême à l'autre. Il s'agit bien d'extrêmes : il aurait été possible de ne pas exclure les neurosciences il y a 40 ans, alors même que Jean-Pierre Changeux, Marc Jeannerod, Pierre Buser, Michel Imbert, Alain Berthoz et d'autres jetaient en France les bases des neurosciences cognitives ; comme il serait loisible aujourd'hui de parler de sciences neuro-cognitives, ce qui éviterait (peut-être) de les faire apparaître comme une province des neurosciences. En pratique, nous savons bien que le champ interdisciplinaire dont nous parlons, quelle que soit l'étiquette qu'on lui choisit, comprend plusieurs composantes dont l'une est les neurosciences. Nous savons que les neurosciences interagissent plus ou moins intensément avec d'autres composantes, mais qu'elles sont loin de les absorber. De fait le schéma hexagonal qui était souvent utilisé pour présenter la « galaxie » des sciences cognitives dans les années 1980 reste acceptable, à titre de première approximation... à condition d'interpréter charitablement les intitulés des six sommets³, et d'en ajouter un septième, de nature un peu différente, à savoir la biologie évolutive. Ce qui est certain, c'est que les faits ne confirment pas davantage une mainmise des neurosciences sur le domaine qu'un maintien du *statu quo ante*.

Car bien entendu, les choses ont beaucoup changé au cours de ces 40 ans, et les positions extrêmes rappelées à l'instant ne sont que l'expression exagérée d'une réalité, à savoir un déplacement du centre de gravité de la galaxie, que le présent numéro examine sous différents angles. Dans leur article introductif, Cyril Monier et Alessandro Sarti proposent plusieurs explications complémentaires, qui me semblent justes. Je me propose d'en adjoindre une autre, compatible avec elles mais se situant, en quelque sorte, en amont. Le passage d'une configuration à l'autre résulterait d'une « dynamique naturalisante » : pour « naturaliser » le mental, la voie la plus prometteuse a d'abord semblé être celle d'une « logique » de la fonction ; à quoi s'est substituée, graduellement, une « biophysique » de la structure.

1 – UNE MÉTAPHORE FLUVIALE

Qu'est-ce qu'un fleuve ? C'est, schématiquement, une masse d'eau qui obéit à la loi de la gravité sous la contrainte de la relative imperméabilité des sols traversés. Elle suit donc à tout moment la ligne de plus grande pente. C'est ainsi que sur un premier segment de son cours, le fleuve peut suivre une certaine vallée, au gradient plus favorable que les collines avoisinantes ; et que plus loin, les collines peuvent s'ouvrir et offrir à l'eau une descente plus rapide, auquel cas le fleuve quitte la vallée qu'il suivait pour s'engouffrer dans la nouvelle ouverture.

Voici alors la cible de la métaphore. Le fleuve est la recherche en mouvement, plus précisément la recherche sur l'esprit humain. La gravité est

³ Dû semble-t-il à George Miller, cet hexagone cognitif est repris et commenté dans maintes introductions, p. ex. Andler, 2006b. Ses sommets étaient à l'origine la psychologie, la linguistique, l'intelligence artificielle, les neurosciences, l'anthropologie et la philosophie.

l'exigence naturaliste. La vallée que parcourt le fleuve à partir de sa naissance dans les années 1940-1950, c'est le paradigme classique ou symbolique en sciences cognitives, associé à l'IA première manière, et fondé sur l'idée que l'objectif à atteindre, c'est la mise au jour de la logique (au sens étendu) des transactions informationnelles internes à l'esprit. Les collines qui maintiennent le fleuve dans son lit, ce sont les objections adressées à l'idée d'une science du cerveau *comme* science de l'esprit, et le dénuement (réel ou perçu) des neurosciences de l'époque. Ainsi, naturaliser l'esprit sous les espèces d'une logique informationnelle est une voie qui s'offre, le faire sous les espèces d'une biologie du cerveau une voie qui se refuse.

C'est la configuration inverse qui se met en place dans les années 1990-2000 (soit un demi-siècle plus tard). La voie logico-symbolique se trouve obstruée et le fleuve progresse avec lenteur ; son potentiel naturalisant s'affaiblit à mesure qu'on prend conscience de l'importance de ce qu'il a choisi – ou été contraint, faute de moyens – de négliger, et, sur un plan plus général, sur l'ambivalence de son engagement naturaliste. Mais cela à soi seul n'aurait pas suffi à faire dévier le fleuve de la vallée où il coulait jusque-là. Il fallait qu'en même temps les neurosciences (tout récemment baptisées) offrent une brèche profonde, à la fois en écartant l'obstacle du dénuement (elles s'enrichissent d'outils techniques – neuroimagerie – et conceptuels – modélisation, neurocalcul), et en mettant en avant, sur le plan conceptuel, leur potentiel naturalisant (pour filer la métaphore, en présentant une pente beaucoup plus forte que celle de la vallée d'origine).

Quelle que soit l'utilité de cette métaphore pour le lecteur, elle a au moins celle de justifier le propos et le plan du présent article. La question préliminaire est celle-ci : qu'est au juste la « pulsion naturalisante », la recherche d'une « naturalisation » de l'esprit ? Se pose ensuite une série de questions de nature descriptive, portant sur la dynamique engendrée par cette recherche : pourquoi la voie logico-symbolique est-elle apparue comme prometteuse ? pourquoi a-t-elle été, du moins en apparence, supplantée par la voie neuroscientifique ? ces voies sont-elles comprises, sur le terrain, comme concurrentes ou complémentaires ? La série suivante est de nature normative : quels sont les mérites des deux programmes en tant qu'entreprises de naturalisation du mental ? Enfin, il y a une leçon à tirer de cet examen quant à l'idée même de naturalisation, leçon qui à son tour éclaire la situation apparemment chaotique du domaine.

2 – NATURALISME ET NATURALISATION

Le naturalisme s'entend d'une multitude de façons, au point que certains estiment vain tout effort pour leur trouver un commun dénominateur. La seule référence à la nature est trop vague pour jouer ce rôle : que signifie que « tout appartient à la nature », que « tout est naturel » ? Tout dépend de ce à quoi la nature est opposée, et cela dépend du contexte – est-ce au surnaturel, au culturel, à l'artificiel, voire au conceptuel ? D'autre part, le naturaliste s'intéresse-t-il à la nature des choses, et affirme-t-il qu'il n'en existe, fondamentalement, que d'une sorte, qu'elles sont toutes découpées dans la même étoffe, ou bien affirme-t-il que toutes les choses sont connaissables selon un mode unique, qui est celui qui nous donne accès aux objets et phénomènes

naturels déjà identifiés, dans l'ordre physique et biologique ? En sorte, en particulier, que la philosophie elle-même, qui prétend apporter des connaissances, se ramène à ce mode unique ?

Le naturalisme qui nous concerne ici peut toutefois être circonscrit. Sur le plan ontologique, le naturel n'est pas opposé au surnaturel, mais au mental, à ses manifestations régulières et à ce qui en dépend conceptuellement : intentionnalité, conscience, libre arbitre, normativité, culture... Sur le plan épistémique, les sciences de la nature constituées sont le mode de connaissance qui est opposé soit aux sciences de l'homme et de la société, soit à des formes non scientifiques de savoir – littéraires, testimoniales, biographiques, introspectives, intuitives, narratives..., la frontière entre ces formes et les SHS étant évidemment poreuse.

Pour le naturaliste radical, ces oppositions, sur le plan ontologique comme sur le plan épistémique, n'ont pas lieu d'être : elles ne sont que le reflet temporaire de notre ignorance, renforcée par des croyances héritées d'un âge révolu. Rejetons ces croyances, faisons progresser le savoir authentique, et nous verrons les distinctions s'effacer progressivement : les entités, choses et phénomènes, qui semblaient échapper à l'ordre naturel, seront, l'une après l'autre, rapatriées dans la Nature, à mesure qu'elles entreront dans le périmètre des sciences de la nature.

On peut être naturaliste sans l'être de manière aussi radicale. On peut par exemple admettre que, sur le plan ontologique, tout est naturel au sens où tout est composé d'entités élémentaires décrites par les sciences physiques et biologiques, ou bien encore, si l'on se sent assuré de saisir la notion de matière, que tout est matériel car composé d'entités manifestement matérielles ; et en même temps contester l'idée que les sciences de la nature ont ou auront un jour les moyens de rendre compte de tout ce qui est, de l'ensemble des phénomènes. Ou encore, sans aller jusqu'à contester cette idée, on peut s'abstenir, admettre qu'on n'a d'argument convaincant ni pour la rejeter ni pour l'accepter. Inversement, on peut se désintéresser de la question ontologique, la juger mal posée, ou se déclarer incompetent, tout en adhérant au projet d'une science naturelle (c'est-à-dire conforme à la méthodologie générale des sciences de la nature actuelles) de tout ce qui est ou arrive. On peut également souhaiter laisser hors du débat certains ordres de phénomènes : la science elle-même, le libre arbitre, l'art, la normativité, les concepts...

Mais ce qui unit toutes ces variétés de naturalismes, c'est la visée – ce que j'ai appelé plus haut la « pulsion naturalisante » : être naturaliste, c'est se donner pour objectif de dissiper une illusion, celle d'un statut spécial, séparé, de certaines entités, phénomènes, connaissances ; c'est vouloir unifier, ou réunifier le réel. C'est endosser un *programme*, accepter un ordre de marche. Les motivations peuvent grandement varier : convictions philosophiques, convictions morales, pure soif de comprendre, désir d'explication d'un certain ordre de phénomènes, désir de maîtrise à des fins utilitaires, amélioratives ou non... Les objectifs intermédiaires sont également divers. Mais chacun se met en route dans la même direction générale.

Et c'est au moment de se mettre en route que le naturaliste, quelle que soit sa tendance, franchit un saut : celui qui le fait passer d'une intuition, d'une conviction, d'une aspiration à un plan d'action. Depuis l'aube des temps les

humains ont rêvé de voler ; mais c'est Icare qui le premier (admettons-le), a cherché une manière concrète de réaliser ce rêve. Pour ce qui est du naturalisme, ce saut fait passer d'une forme flottante, vaporeuse, spéculative à un ancrage dans un programme de recherche concret. C'est alors que, de même que les partisans du plus léger que l'air et ceux du plus lourd que l'air se sont séparés, tout en conservant le même objectif lointain, de même les naturalistes se sont divisés en colonnes ou en flottilles disjointes, au point parfois de se perdre de vue, et en tout cas de mettre en doute les chances de succès des armées concurrentes, ce qui n'a pas cessé d'alimenter, depuis la naissance des sciences cognitives, les controverses inter-paradigmatiques.

3 – LA VOIE LOGICO-SYMBOLIQUE ET LE FONCTIONNALISME⁴

Plaçons-nous dans les années 1930-1950. Grâce à la renaissance de la logique, il est devenu possible, d'une part de décrire formellement la dynamique d'un ensemble de pensées beaucoup plus vaste que les propositions entrant dans les syllogismes d'Aristote (« Tous les hommes sont mortels », etc.) ; d'autre part de la décrire comme une algorithmique. Penser, selon le modèle qui sera élaboré à partir des années 1950, c'est être doté des capacités représentationnelles d'un langage formel et des capacités calculatoires (au sens strict) caractérisées par la théorie de la récursivité⁵. Tel est le point de départ de la voie logico-symbolique annoncée par Hobbes, ouverte par Turing, poursuivie par Newell et Simon et l'ensemble de l'IA et par les sciences cognitives dites parfois classiques.

Reste à savoir à quoi se rapportent les symboles et sur quoi opèrent les fonctions logiques. La réponse fonctionnaliste tient en un mot : information. L'information se situe à un niveau d'abstraction qui permet de subsumer les différentes formes ou modalités que prend une pensée donnée. Ce niveau est celui où s'établissent les liens à la fois causaux et logiques entre pensées, y compris les perceptions et les actions (ou si l'on préfère, les intentions et autres déterminants mentaux des actions). L'information portée par une pensée particulière condense en quelque sorte l'ensemble de ses dispositions à influencer sur les autres. Les règles calculatoires d'inférence, quant à elles, se caractérisent par leur capacité à combiner les informations constitutives des pensées. La situation est la même que pour la mécanique élémentaire : la structure détaillée ou la composition chimique de la poulie, du levier, du plan incliné et du câble n'apportent rien à la compréhension, la prévision et le

⁴ Les générations se succèdent, et ce qui a été répété, ressassé dans les années qui ont précédé le tournant neuroscientifique n'est pas nécessairement familier pour les lecteurs plus jeunes. Les indications qui sont données dans cette courte section, à titre de rappel ou de mise au courant rapide, sont élémentaires et extrêmement sommaires. Des présentations détaillées sont légion : voir p. ex. Andler, 1992/2004, Andler, 2011a, mais il y a en a bien d'autres, telles Haugeland, 1981, Bechtel & Graham, 1999... Beaucoup de lecteurs seront tentés de sauter cette section, et je les y encourage, sauf s'ils n'y voient que de vieilles lunes vouées à l'oubli, ce qui est à mon sens une grave erreur si l'on veut comprendre l'état présent des sciences cognitives.

⁵ Le sens de « calcul », « computationnel » s'est progressivement étendu, au point aujourd'hui de renvoyer, bien souvent, au fait qu'un modèle est physiquement réalisable, de quelque manière que ce soit. Dans le cadre classique, il est au contraire de la plus haute importance que le modèle soit, à un certain niveau, à la fois descriptible dans un langage formel et réalisable dans un dispositif tel qu'un ordinateur. Sur ce point encore, voir par exemple Andler, 1992/2004.

contrôle de leurs interactions, tant du moins qu'elles leur assurent un minimum de qualités générales (rigidité, impénétrabilité, résistance à la traction, etc.). De même, la pensée « un chat est un mammifère » est commune à toutes sortes de phrases pensées, écrites, chuchotées, clamées, gravées, etc., dans différentes langues. Si l'on recherche à expliquer le rôle que joue dans la dynamique mentale du lecteur la phrase imprimée ici : « Un chat est un mammifère », la description physique de l'encre d'imprimerie formant sur cette page ladite phrase noie l'explication dans une masse d'indications qui la rend inaccessible. Quant à la transition qui fait passer du couple de pensées {« Un chat est un mammifère » , « Tout mammifère a deux reins »} à la pensée {« Un chat a deux reins »} elle est, moyennant codage, une opération arithmétique élémentaire qui peut être exécutée sur un boulier, par une suite d'écritures sur une feuille de papier, par une calculette, etc. En d'autres termes, les pensées et leur dynamique sont des fonctions et non des substances singulières : c'est leur rôle informationnel qui les caractérise, pas leurs diverses réalisations physiques ; tel est l'argument dit de réalisation multiple en faveur du fonctionnalisme⁶.

Une fois acquise, ou admise, la fidélité du modèle à son objet⁷, il reste à comprendre pourquoi il ouvre une voie à la pulsion naturalisante, en quoi il est un modèle ou l'ébauche d'un modèle *naturaliste* de la pensée ou de la cognition. C'est aujourd'hui chose à la fois plus facile et plus difficile qu'à l'époque où Turing, Newell et Simon, Fodor, Chomsky et d'autres l'ont proposé. C'est plus facile en raison de l'existence de l'ordinateur et de la pénétration du concept général du « numérique » dans l'entendement commun. Nous ne doutons généralement pas qu'en un sens au moins, un ordinateur reproduit des fragments ou des épisodes de pensée. De la même manière que la synthèse de l'urée a fait entrer en 1828 la chimie organique dans la chimie générale, abolissant la frontière ontologique entre substances fabriquées par les organismes biologiques et substances fabriquées par le chimiste, la fabrication dans un ordinateur, machine fabriquée par l'homme à partir de matériaux naturels, d'une « pensée de synthèse » abolit la frontière entre la pensée humaine et le monde des dispositifs matériels.

Remarquons au passage que dans l'analogie, la « pensée » de l'ordinateur correspond à l'urée de synthèse, et la pensée humaine à l'urée naturelle. D'un côté on a des processus inventés et réalisés par l'industrie humaine, de l'autre un couple apparemment mal assorti : l'urée que secrètent nos organes est aussi naturelle qu'on peut l'être, alors que la pensée l'est, dans la conception traditionnelle du moins, aussi peu qu'on peut l'être. Il y a donc du premier côté des entités *artificielles*, de l'autre une entité *naturelle* et une entité non

⁶ Le terme est utilisé de manière variable dans le champ des sciences cognitives : il désigne tantôt l'idée que les pensées se caractérisent par leur rôle dans un réseau d'interactions, tantôt le cadre formé par cette idée enrichie de l'hypothèse logico-symbolique, dont une spécification importante est l'hypothèse du langage de la pensée formulée par Jerry Fodor : voir Fodor, 1975, 2008. Hilary Putnam en donne l'une des premières versions par référence à l'ordinateur : les états mentaux sont à l'organe mental ce que les états internes d'un ordinateur sont à l'ordinateur : voir Putnam, 1960 (et pour une autocritique, Putnam, 1988).

⁷ Moyennant, bien entendu, d'une part un considérable travail d'explicitation et d'enrichissement du modèle, et d'autre part une idéalisation.

naturelle. En quoi la possibilité de fabriquer une pensée artificielle tend elle à prouver que la pensée est naturelle ?

La question nous éclaire sur le changement de perspective qui s'est opéré depuis les débuts des sciences cognitives. Plaçons-nous d'abord *avant* la synthèse de l'urée : la conception dominante du vivant était alors vitaliste : la distance entre l'ordre du vivant et l'ordre de l'inerte semblait infranchissable, ou du moins considérable. Considérons alors la pensée, telle qu'elle apparaissait *avant* la révolution cybernétique et la naissance des sciences cognitives classiques : elle semblait alors au moins aussi éloignée de la nature qu'avant la synthèse de l'urée l'urée biologique l'était de l'urée de synthèse. La « synthèse » de la pensée opère sur la *res cogitans* le même genre d'élimination que la synthèse de l'urée sur le principe vital. La pensée rejoint la sphère des phénomènes physiques. Aujourd'hui, cette percée nous impressionne moins, et notre objectif est de faire entrer la pensée, désormais placée dans la sphère physique, dans celle plus étroite du vivant. Est-on à mi-chemin d'une naturalisation du mental, ou bien a-t-on atteint l'objectif, moyennant la réalisation complète du programme logico-symbolique, qui est loin d'être acquise ? C'est en ce point précis que se noue un désaccord qui se formule différemment lors de la première et de la seconde phase que j'ai distinguées, et dont la métaphore fluviale nous aide – du moins je l'espère – à saisir la nature véritable.

Voyons d'abord comme il se présente au cours de la première phase, dans les années 1950-1980. D'un côté, la voie logico-symbolique est grande ouverte, d'autant qu'elle peut se prévaloir des progrès sensibles, voire impressionnants, de son volet artificiel, celui de l'intelligence artificielle symbolique ou classique : le principe du *verum-factum* de Vico (je peux connaître ce que je sais fabriquer) semble lui conférer un certificat de plausibilité. D'un autre côté, la voie neuroscientifique, on l'a rapidement évoqué, semble bloquée par de hautes collines. Sur le plan conceptuel, l'argument de réalisation multiple compte autant en défaveur de l'identification des processus mentaux à des processus neuraux qu'il compte en faveur du fonctionnalisme : si des systèmes, biologiques très différents, voire artificiels, peuvent donner naissance à un même flux mental, la description des processus neuraux réalisant ce flux dans un organisme particulier, le cerveau d'un humain, ne nous renseignera pas plus que la description des trajectoires de particules élémentaires dans un ordinateur ne nous renseignerait sur ce que l'ordinateur est en train de faire, ou sur les objectifs de ceux qui l'ont fabriqué. Historiquement, l'idée que le cerveau secrète la pensée comme le foie secrète la bile, ainsi formulée par Cabanis en 1802, mais présente dans toute une tradition matérialiste post-cartésienne, est discréditée, ne serait-ce que parce qu'elle est restée abstraite, flottante, ne donnant lieu à aucune découverte, à aucun progrès de la connaissance. La neuropsychologie qui prend naissance peu après et culmine avec les découvertes de Broca et Vernicke constitue un premier pas incontestable, qui conduira un siècle plus tard au programme modulariste en sciences cognitives, mais qui reste isolé et ne jette guère de lumière sur la notion de « siège » de la pensée. Par ailleurs, les découvertes fondamentales qui vont bouleverser ce qui ne s'appelle pas encore les neurosciences concernent la vision et d'autres fonctions qui, à l'époque,

semblent éloignées de la pensée au sens plein du terme, un sens étroitement lié au raisonnement et au langage, sur lesquels ces disciplines sont au contraire, à Broca près, muettes. Enfin, précisément en raison du désintérêt affiché par l'orientation logico-symbolique majoritaire, les neurosciences ne sont pas en mesure de contribuer pleinement au mouvement scientifique : il y a dans le rapport Lentin et ses équivalents à l'étranger un effet auto-réalisateur.

4 – LE TOURNANT NEUROSCIENTIFIQUE ET LA DOCTRINE NEURONALE

Quand on se demande, aujourd'hui, ce qui a provoqué le basculement du domaine vers les neurosciences, au tournant du 21^e siècle, on oublie que celles-ci sont une création récente. Il a fallu que s'impose, lentement, difficilement, l'idée que le cerveau puisse faire l'objet d'une science, ou d'une fédération organisée de programmes de recherche.

De manière extrêmement schématique, les sciences du cerveau, au sens moderne du terme, naissent vers le milieu du 19^e siècle, avec Broca d'une part, Ramon y Cajal d'autre part. Mais c'est seulement à partir de la fin des années 1950 que s'est cristallisée l'idée que la douzaine de spécialités « neuro-X » (avec X = anatomie, -logie, physiologie, psychologie, chimie, etc.) qui s'étaient constituées devaient se penser comme les branches d'une unique grande discipline, la science du cerveau, *brain science*⁸. Le « Brain Research Institute » est créé à UCLA en 1959, l'*International Brain Research Organisation* est fondée à Ottawa en 1961⁹. L'année suivante le MIT met sur pied un programme de recherche en neuroscience, le bientôt célèbre NRP, accompagné peu après d'un programme de formation, la Brain Research Association (BRA) britannique émerge entre 1965 et 1968, la Society for Neuroscience américaine voit le jour en 1969.

Il est difficile, dans le contexte présent, de ne pas se demander pourquoi les choses ont semblé si compliquées et exigé tout ce temps pour se mettre en place. N'est-il pas évident que si l'on part d'une définition de l'esprit comme l'ensemble des fonctionnalités cognitives du cerveau, le cerveau est l'organe de la pensée, ou pour parler comme les biologistes, la structure dont la fonction est la pensée ? Il y a donc *une* science de cette structure, les neurosciences, science qui est aussi la science de l'esprit.

On lit souvent, dans les lettres de motivation de candidats à une formation en sciences cognitives, qu'ils sont fascinés par le cerveau et veulent en comprendre le fonctionnement. On trouve des déclarations tout à fait semblables en préambule de projets de chercheurs chevronnés. Naïveté des premiers, rouerie des seconds ! Imagine-t-on Newton annoncer qu'avec les 3 millions d'euros que lui donnera l'ERC, il « comprendra le fonctionnement de l'univers » ? S'il est honnête, Newton annoncera qu'il se propose de mettre en évidence un principe d'organisation de l'univers, la gravité, qui permettra de le

⁸ Comme souvent, le français favorise le pluriel et parle *des* neurosciences, mais les Français, en l'occurrence, s'alignent sur le consensus mondial pour les concevoir comme un champ disciplinaire unique (avec en son sein, comme dans tout champ scientifique, des mathématiques (pluriel grammatical) à la physique (singulier), des spécialités.

⁹ Suite à une rencontre organisée à Marseille en 1955 par Henri Gastaut, le maître français de l'électroencéphalographie et de l'épileptologie. voir Doty, 1987.

considérer, sous un certain rapport, comme un système dont le « fonctionnement », les lois, pourront alors être graduellement mis au jour. L'univers ne se donne pas immédiatement comme objet d'une science possible ; c'est seulement au prix d'un effort prodigieux de « philosophie naturelle » qu'il le devient. Encore faut-il se garder de penser que ce qui advient ainsi est une science totale, complète, de l'univers : la physique newtonienne n'est science que de ce qui, dans l'univers, constitue son objet.

Que la situation, pour la connaissance du cerveau, soit comparable, l'histoire en fournit donc un premier indice. S'il a fallu un siècle pour que les sciences du cerveau se fondent dans les neurosciences, c'est parce qu'il leur a longtemps manqué un principe organisateur. Il est facile aujourd'hui d'identifier ce principe, moins facile sans doute de comprendre pourquoi il ne tombait pas sous le sens. Cette idée maîtresse est celle de système de traitement de l'information¹⁰. Si elle semble aujourd'hui presque triviale, elle ne s'est dégagée que progressivement, par simplification et schématisation, dans le creuset interdisciplinaire de la première cybernétique et a constitué à la fois le cadre des sciences cognitives et celui de l'IA première manière, comme on l'a vu trop vite dans la section précédente. En ce sens, les neurosciences dans leur ensemble ont été d'emblée « cognitives ». Et pourtant, il aura fallu encore une trentaine d'années pour que se constitue un champ qui se donne pour nom « neurosciences cognitives ».

C'est précisément le moment où le fleuve a changé de lit. Pour plusieurs raisons, recensées dans l'article introductif et analysées dans d'autres articles du présent numéro, les neurosciences sont apparues, au tournant du siècle, comme offrant un boulevard à l'analyse du cerveau, et en même temps une descente en forte pente vers la naturalisation de l'esprit, dès lors qu'on disposait du concept de système de traitement de l'information et de l'expérience acquise par les sciences cognitives, l'informatique, l'IA et la robotique dans le déploiement et la déclinaison fine de ce concept.

Qu'est-ce qui donne aux neurosciences cet attrait irrésistible, pourquoi offrent-elles une pente si forte à l'envie de naturalisme ? C'est d'abord que le cerveau est un objet concret, palpable, qui s'offre, semble-t-il, à l'examen sans médiation. Le contraste avec les objets de la psychologie est à son avantage : les états et processus mentaux étant inobservables, on ne peut espérer les saisir qu'indirectement via le comportement, ou via le témoignage de l'introspection. Or les deux voies semblent peu sûres. L'introspection est depuis longtemps mise en examen pour non-fiabilité, tromperie systématique voire illusion¹¹. Quant au comportement, il présente deux difficultés de taille : d'une part, remonter du comportement aux états et processus mentaux censés en constituer la cause est un problème inverse n'admettant généralement pas de solution unique ; d'autre part, un comportement n'est pas un simple épisode moteur,

¹⁰ On a beaucoup parlé, dans les années 1970-1980, de « paradigme STI » : voir Newell & Simon, 1972 ; Demailly & Lemoigne, 1986.

¹¹ Une bonne présentation historique est donnée dans Lyons, 1986 ; pour la situation présente, outre ce livre on pourra consulter l'article Schwitzgebel, 2016.

enchaînement de déplacements d'effecteurs : il est intentionnel¹², et l'intention fait intervenir non seulement les états mentaux de l'agent, mais l'environnement.

Cet avantage du cerveau sur l'esprit, c'est-à-dire sur le domaine psychologique, resterait inopérant si nous étions incapables de procéder à un examen analytique du cerveau, à en distinguer les parties et les composants, et en comprendre les principes de fonctionnement. Nous sommes longtemps restés dans cette impuissance, et ce sont les progrès des disciplines du cerveau avant leur jonction dans les neurosciences, et bien évidemment depuis l'essor de la neuroimagerie, qui ont changé la donne. L'objet concret et palpable se révèle à l'examen comme un système fonctionnel dont la complexité ne semble pas faire obstacle à notre compréhension : elle est au contraire la promesse de découvertes à venir.

Mais le cerveau possède un deuxième avantage sur le psychologique : c'est un objet naturel au sens où il est produit par des processus biologiques naturels, tant phylogénétiques qu'ontogénétiques. Sur le plan ontologique, son appartenance à l'ordre de la nature est aussi assurée qu'il est possible, et sur le plan épistémique il relève d'une science de la nature, la biologie, dont le statut et le succès sont également hors de doute.

Bref, depuis une vingtaine d'années il est devenu difficile de résister à ce que deux philosophes ont appelé dans un article retentissant de 1999¹³, la « nouvelle doctrine neuronale »¹⁴, l'idée que les sciences cognitives mûres, celles qui sont en train de s'extraire, plus ou moins rapidement, des tâtonnements initiaux, se réduiront aux neurosciences. Et que de ce fait la naturalisation de l'esprit sera sinon achevée, du moins suffisamment engagée pour ôter toute crédibilité au scepticisme, de la même manière que la naturalisation du vivant ne fait de doute pour presque personne, alors même que la biologie est loin d'avoir résolu toutes ses énigmes, personne n'imaginant qu'elle ait pour l'essentiel achevé sa tâche (comme on le dit périodiquement de la physique depuis des décennies).

Enfin, au moment même où les neurosciences, grâce à leur dynamique propre, prenaient l'avantage sur les sciences cognitives classiques, celles-ci donnaient des signes d'essoufflement. D'abord, elles avaient cru satisfaire aux exigences du naturalisme en imposant aux mécanismes qu'elles postulent d'être implémentables dans un système physique naturel, et de l'être en particulier dans le cerveau ; mais cette contrainte était faible du fait de l'instanciation multiple : elle se limitait en pratique à des bornes architecturales ou processuelles sans portée théorique. Cette indifférence aux spécificités du

¹² Le cas le plus simple est celui où je soulève ma casquette dans l'intention de m'essuyer le front. Mais même ce cas est loin d'être simple, et d'autres le sont moins encore : les rapports entre intention, mouvement et action sont complexes. Voir les travaux d'Élisabeth Pacherie, p. ex. Pacherie, 2012, et Khamassi & Pacherie, 2018.

¹³ Gold & Stoljar, 1999.

¹⁴ « Nouvelle » pour la distinguer de la théorie de Ramon y Cajal, établie par lui dans les années 1880, selon qui le tissu nerveux est composé de cellules discrètes, les neurones. Cette « nouvelle doctrine » est souvent aussi désignée par le terme ambigu de « neurophilosophie » introduit par Paul et Patricia Churchland : Churchland, 1986.

cerveau, tenue d'abord pour une force du programme de recherche, est apparue, de plus en plus, comme une faiblesse : non seulement cette indifférence à l'égard des bases naturelles semblait suspecte, mais elle privait le chercheur d'une source d'information prodigieusement riche. Ensuite, sur son terrain même, la voie logico-symbolique semblait incapable de mener à son terme la naturalisation des dimensions fondamentales du mental. La stratégie suggérée par Daniel Dennett et largement suivie, consistait à ramener les manifestations complexes de l'intentionnalité, de la conscience, de la normativité à des combinaisons de manifestations simples, et de « naturaliser » ces formes simples. Un début assez convaincant d'exécution de ce programme portait sur l'intentionnalité, premier objectif à atteindre selon Dennett¹⁵. La biosémantique, dans diverses versions¹⁶, semblait prometteuse, mais on voyait mal comment passer des formes élémentaires aux formes complexes. Les choses se présentaient moins bien encore pour ce qui est de la conscience et de la normativité.

En réalité, il s'agit là de difficultés qui, outre qu'elles sont présentées ici de manière simpliste, intéressent surtout les philosophes et sont loin d'être vues par eux comme dirimantes. Pour notre propos, ce qui compte est qu'elles ont été *perçues* de l'extérieur, pour une génération montante de chercheurs marqués par les progrès impressionnants des neurosciences, comme un signe d'échec, ou en tout cas comme une indication du fait que la voie logico-symbolique fonctionnaliste ne suffit pas à la tâche : il lui faut soit disparaître purement et simplement, soit s'adjoindre des ressources nouvelles.

Mais pour beaucoup de chercheurs plus proches du laboratoire que du fauteuil du philosophe, le défaut le plus sensible de l'approche classique était qu'elle semblait laisser hors de son champ des phénomènes qui apparaissaient pourtant comme essentiels et qu'on ne pouvait continuer à ignorer au nom d'un nécessaire étagement des tâches. La plupart de ces phénomènes sont rattachés au sigle 4EA : la cognition est incarnée (*embodied*), immergée (*embedded*), inséparable de l'action (*enacted*), instrumentée (*extended*) et émotive de part en part (*affective*) ; s'y ajoutent ou s'y combinent la considération du soi, des vécus et de la narration, et de la socialité. Il n'est nullement acquis que l'approche classique soit incapable d'intégrer ces phénomènes, mais d'une part on est en droit de lui demander des raisons de penser qu'elle le peut, d'autre part on peut s'interroger sur ce qui resterait finalement de l'approche classique une fois qu'elle aurait intégré toutes ces dimensions. Et ici encore, ce n'est pas ce qu'il en est et en sera qui compte, c'est la perception, au moment considéré, par un nombre croissant de chercheurs, des chances respectives des deux orientations.

5 – LES LIMITES DE LA DOCTRINE NEURONALE, LE PLURALISME ET LE DELTA NATURALISTE

Prendre la doctrine neuronale au sérieux, c'est ne pas la ramener d'emblée à une recommandation de bon sens, celle de poursuivre l'étude du cerveau et d'en tirer tous les enseignements pertinents pour la compréhension de l'esprit.

¹⁵ Dennett, 1969, 1978.

¹⁶ Millikan, 1984 ; Papineau, 1987 ; Neander, 2017.

C'est donc l'interpréter comme l'annonce d'un programme d'élimination ou de subordination des (autres) sciences cognitives. Il ne s'agit pas simplement d'intégrer le vocabulaire de ces autres disciplines à celui des sciences du cerveau, et de coller l'étiquette « cognitive » après l'étiquette « neurosciences ». Cette manœuvre est précisément ce qu'un grand neuroscientifique, Gerald Edelman, rejette formellement : pour lui, ce qu'il s'agit de faire advenir, c'est une « théorie scientifique de l'esprit fondée directement sur la structure et le fonctionnement du cerveau ». Il précise que par « scientifique », « [il] entend [...] une description fondée sur l'organisation neuronale et phénotypique d'un individu et dont la formulation ne fasse appel qu'aux mécanismes physiques et chimiques donnant naissance à cette organisation »¹⁷. Reste cependant la possibilité que la science du cerveau ainsi comprise laisserait aux autres disciplines historiques, du moins pendant un certain temps, un rôle heuristique ou pédagogique, permettant de faire le pont entre les concepts et descriptions scientifiques d'un côté, et les concepts de sens commun utiles pour la conduite des affaires humaines.

La doctrine neuronale doit selon moi être rejetée, mais elle n'est pas absurde, comme certains pensent pouvoir le montrer à partir d'arguments conceptuels classiques sur lesquels je ne reviens pas ici.

La première objection à la doctrine neuronale est le constat que l'essentiel des travaux en neurosciences est en fait guidé par des hypothèses formulées dans les termes de la psychologie commune ou scientifique : il s'agit le plus souvent de découvrir *comment* le cerveau accomplit une tâche informationnelle donnée, ou inversement quelle tâche informationnelle est accomplie lors d'un processus cérébral donné : la caractérisation des tâches (ou fonctions) informationnelles est par définition du ressort de la psychologie. Il est difficile aujourd'hui d'imaginer que cette dépendance puisse prendre fin. Nombre de neuroscientifiques ne s'attendent d'ailleurs à rien de tel : il leur suffit de penser qu'à terme, les neurosciences auront déterminé avec certitude (aux habituelles révisions près qui affectent toute science mûre) les mécanismes neuraux qui sous-tendent les fonctions cognitives et produisent le comportement des agents (humains voire non-humains). Or s'ils laissent à la psychologie, à la linguistique, à l'analyse conceptuelle, à l'anthropologie évolutionnaire... le soin de déterminer le répertoire des fonctions cognitives, ne retrouvent-ils pas le modèle interactif standard des sciences cognitives ?

Ils s'en démarquent du fait qu'ils accordent aux neurosciences *le dernier mot* chaque fois qu'il existe un doute ou un désaccord entre théories. Or ce privilège du dernier mot est injustifié, et c'est la deuxième objection à la doctrine neuronale. Il est motivé par une vision dichotomique dans laquelle les (autres) sciences cognitives émettent des conjectures, parmi lesquelles les neurosciences font une bonne fois le tri, de la même manière que l'observation permet une bonne fois de faire le tri entre théorie vraie (ou conjecture vivante) et théorie fautive (ou conjecture réfutée). Or même si nous acceptons pour les besoins de la discussion ce modèle beaucoup trop simple de la démarche des sciences expérimentales, nous aurions tort de faire correspondre l'observation aux neurosciences et les conjectures aux autres sciences cognitives : il y a

¹⁷ Edelman, 1989, cité par Gold & Stoljar, *op. cit.*, p. 811.

beaucoup d'observation dans ces dernières, et la part de théorie en neurosciences ne cesse de croître, ce qui est sans nul doute une bonne chose. Le « neuropositivisme » est une illusion.

Mais la porte reste ouverte à une neurophilosophie plus ambitieuse, dont on voit même s'esquisser ce qui pourrait être un début de réalisation. Il s'agit du programme modulariste généralisé consistant à déterminer, aire par aire, réseau par réseau, le répertoire des fonctionnalités de base du cerveau, grâce à la neuroimagerie. On ne peut totalement exclure qu'après s'être appuyées sur la psychologie et d'autres disciplines dans la phase de démarrage, les neurosciences puissent acquérir une compréhension suffisamment riche du cerveau pour pouvoir se passer de tout apport extérieur, un peu de la manière dont la physique fondamentale a acquis son autonomie vis-à-vis de la physique commune des objets de taille moyenne, grâce à l'outil mathématique. Cette possibilité semble néanmoins très éloignée, du fait de l'insuffisante robustesse des résultats de neuroimagerie, des difficultés conceptuelles que soulève une vision modulariste du cerveau, si éloignée qu'elle soit aujourd'hui des conceptions fixistes héritières de la phrénologie, et, *last but not least*, de l'état embryonnaire de nos connaissances générales sur le cerveau.

Si l'absorption des sciences cognitives par les neurosciences semble exclue dans un avenir prévisible, sommes-nous nécessairement ramenés à la conception standard d'une complémentarité des disciplines composantes convergeant vers une science unifiée de l'esprit et du cerveau ? C'est certainement la « ligne officielle » qui a cours dans le domaine des sciences cognitives. En réalité, celles-ci me semblent évoluer selon deux dynamiques antagonistes, à l'image des sciences en général. L'unification se produit localement, par exemple lorsqu'un désordre cognitif ou comportemental peut être attribué à un facteur génétique induisant une anomalie structurale ou métabolique dans le cerveau. Mais loin de s'accélérer et de s'étendre par généralisation à des secteurs entiers, l'unification ne progresse que pas à pas. Plus rapide me semble être le mouvement inverse : les progrès récents des neurosciences, indépendants de l'imagerie, l'éloignent de la psychologie plus qu'ils ne l'en rapprochent. Les « neurophilosophes » auraient tort d'y voir la confirmation de leurs vues, car ces progrès ne sonnent nullement la disparition de la psychologie et des autres disciplines, au contraire : ils tendent à leur rendre une certaine autonomie. En sorte que ce qui apparaît aujourd'hui comme une phase transitoire dans le développement du domaine préfigurerait non une science unifiée, mais une science pluraliste, c'est-à-dire au sein de laquelle coexistent des vocabulaires et des modèles explicatifs et prédictifs hétérogènes ne se raccordant que localement, et ne se prêtant ni à la fusion, ni à l'absorption¹⁸. Entre les différentes composantes le rapport n'est généralement pas conflictuel : l'indifférence est plus fréquente que la controverse. Chacune évolue à son rythme, sous l'effet conjoint d'une dynamique interne et des interactions avec les disciplines voisines.

¹⁸ Le pluralisme scientifique est largement discuté en philosophie des sciences depuis une quinzaine d'années : voir Kellert, Longino & Waters, 2006. Comment la consigne d'unité de la science peut être comprise dans une telle perspective est l'objet de Andler, 2011b L'idée d'un pluralisme inévitable dans les sciences cognitives commence à être prise en considération : voir p. ex. Dale, 2008 ; Dale & Chemero, 2009.

Si telle est la situation présente, ou celle qui se dessine, par où passe donc le fleuve de la science, censé suivre la ligne de plus grande pente, celle qui conduit le plus rapidement à la naturalisation du mental ? Faut-il penser qu'à certains moments, la différence de pente entre deux voies possibles est si faible que le fleuve hésite puis se divise ? Ou bien s'aviser que la science en mouvement est produite par les chercheurs, et que ce qui l'oriente est non je ne sais quelle configuration objective de la réalité, mais l'évaluation qu'ils font des chances d'aboutir, ou d'un calcul d'utilité plus large tenant compte d'autres critères, épistémiques ou sociologiques, en sorte que les uns optent pour une orientation, les autres pour une autre ? Ces deux explications me semblent très plausibles, la deuxième ayant d'ailleurs été effleurée plus haut, lorsque j'ai indiqué que ce qui compte pour notre propos n'est pas, du moins directement, la gravité objective des difficultés rencontrées par le programme classique, c'est-à-dire la probabilité objective qu'elles soient surmontées, mais l'évaluation qu'en faisaient les chercheurs tentés par la voie neuroscientifique.

C'est pourtant une troisième explication qui, sans s'opposer aux deux autres, va selon moi au cœur de la question et jette sur le naturalisme une lumière nouvelle. On a signalé, d'entrée de jeu, que le naturalisme peut s'entendre de diverses façons. En particulier, il se spécifie différemment selon le contraste que l'on envisage d'abolir par l'opération d'une naturalisation. Cela nous invite déjà à nous demander si les différents paradigmes qui se disputent la conduite des opérations au niveau le plus général, au lieu d'être, comme ils se présentent, concurrents, ne diffèrent pas par leur objectif, c'est-à-dire par le contraste qu'ils estiment, pour une raison ou une autre, infondé voire porteur d'une menace, ou bien encore (l'un n'excluant pas l'autre) qui leur apparaît comme le « verrou » qu'il faut faire sauter en priorité pour atteindre l'objectif ultime d'une naturalisation intégrale.

Mais cette latitude d'appréciation sur l'objectif premier de la démarche n'épuise pas l'ensemble des possibles. Il s'ajoute une « variable cachée », qui est la nature (ou si l'on préfère le contenu) de la naturalisation. Comme la lettre volée de Poe, elle est tellement exposée qu'on ne la voit pas. D'emblée on constate que le naturalisme engage une double conception de la nature et des sciences de la nature. On note la circularité apparente d'une nature définie comme l'objet des sciences de la nature, elles-mêmes définies comme science d'un certain domaine, la nature. On déjoue la menace de circularité en se donnant pour point de départ les sciences de la nature actuelles et en postulant un aller-retour dans lequel, par continuité, les sciences de la nature nous révèlent des pans nouveaux de la nature, et ces pans nouveaux appellent une extension des sciences de la nature, et ainsi de suite. Affaire réglée, lettre posée sur la cheminée, bien en vue. Or qu'est-ce que naturaliser le monde physique ? Nous le savons, ou croyons le savoir¹⁹, car nous nous plaçons après les faits. Mais avant Descartes, Galilée et Newton, qui eût pu dire en quoi consiste la

¹⁹ Il existe en fait un problème bien connu, celui de savoir à quel stade de développement de la physique la physicalisation d'un domaine quelconque se réfère : l'ontologie de Newton ignore les champs, trois des quatre forces fondamentales, les quarks etc. Cela donne lieu à un dilemme formulé d'abord par le philosophe des sciences Carl Hempel. Voir p. ex. Bokulich, 2011. Noam Chomsky estime que depuis la chute du mécanisme cartésien, nous avons perdu toute prise sur les notions de matière ou de physique : voir Chomsky, 2000.

naturalisation du monde physique ? Qui eût pu prévoir qu'elle consisterait à en dégager des lois prenant la forme de systèmes d'équations différentielles, alors que le calcul différentiel restait à inventer ? Et que, plus généralement, naturaliser le monde physique c'est le mathématiser, ou plus précisément, c'est montrer que tout phénomène peut être modélisé comme une interaction entre constituants élémentaires, tirés d'une liste finie et fixée, interaction régie par des lois mathématiques ? La naturalisation du monde physique a donc révélé deux choses : que le monde physique est naturel (ce qui peut sembler aujourd'hui tautologique mais ne l'est pas), et *en quoi consiste* sa naturalité, ou, ce qui revient au même, en quoi consiste le processus de naturalisation quant il s'applique au monde physique. Bref, la naturalisation n'est pas à l'image de l'exécution d'une opération préalablement connue à un domaine nouveau, quelque chose comme le transformation en farine par pilonnage de grains toujours nouveaux. Elle est la découverte d'une opération inconnue, ou plutôt l'invention d'une opération inédite.

Mais, objectera-t-on, cette découverte primordiale ne s'applique qu'au départ du processus, pour amorcer l'aller-retour évoqué à l'instant : une fois la science moderne établie, elle a pu s'étendre de proche en proche, sans nécessiter une découverte aussi fondamentale. On peut en douter : naturaliser le vivant ne semble-t-il pas avoir résulté d'un saut presque aussi considérable ? De deux sauts en réalité, la découverte de la cellule et celle de l'évolution. Quant au mental, il résiste à toute extension « par continuité » des procédés naturalisants connus. Ce n'est pas qu'il ne soit pas *partiellement* assimilable par ces procédés et par les extensions qu'ont imaginées les sciences cognitives depuis leur émergence (en particulier les sauts conceptuels de la théorie de l'information et de la computation). Mais il est loin, à ce jour, de l'être totalement. En sorte que le « gradient naturalisant », au lieu d'être le champ connu et fixé présupposé par la métaphore fluviale qui nous a servi de fil conducteur, est en réalité lui-même l'objet d'une quête. C'est un peu comme si notre paysage dynamique se déformait sous l'effet du passage du fleuve. Quant au fleuve lui-même, non seulement il se subdivise en plusieurs bras enserrant un vaste delta, mais il ne se jette dans aucune mer déjà existante, il la forme à mesure qu'il s'écoule.

RÉFÉRENCES

- Abi-Rached, J. (2012), From Brain to Neuro: The Brain Research Association and the Making of British Neuroscience, 1965-1996. *Journal of the History of the Neurosciences*, 21, 2, 189-213.
- Andler, D. (1992/2004). Calcul et représentations : les sources. In D. Andler (éd.), *Introduction aux sciences cognitives*, édition augmentée. Paris, Gallimard, pp. 13-50.
- Andler, D. (2006a). Cognitive science. In L.D. Kritzman (éd), *The Columbia History of Twentieth Century French Thought*. New York, Columbia University Press, pp. 175-181.
- Andler, D. (2006b). Sciences cognitives. In *Encyclopaedia Universalis*, (édition électronique) ; et in Dictionnaire de la philosophie, Paris, Universalis, 2006, pp. 306-341.
- Andler, D. (2011a). La philosophie des sciences cognitives. In A. Barberousse, D. Bonnay & M. Cozic, (éd), *Précis de philosophie des sciences*. Paris, Vuibert, pp. 519-571.

- Andler, D. (2011b). Unity without myths. In John Symons, Olga Pombo, & Juan Manuel Torres (eds.), *Otto Neurath and the Unity of Science*. Dordrecht, Springer, pp. 129-144.
- Andler, D. (2016). *La silhouette de l'humain. Quelle place pour le naturalisme dans le monde d'aujourd'hui ?* Paris, Gallimard.
- Bechtel, W. & Graham, G. (éd.) (1999). *A Companion to Cognitive Science*. Oxford, Blackwell.
- Bokulich, P. (2011). Hempel's Dilemma and domains of physics. *Analysis*, 71, 4, 646-651.
- Chomsky, N. (1980). *Rules and Representations*. Oxford, Blackwell. Trad. fr. *Règles et représentations*, Paris, Flammarion, 1988.
- Chomsky, N. (2000). *New Horizons in the Study of Mind and Language*. Cambridge, Cambridge University Press.
- Churchland, P.S. (1986). *Neurophilosophy. Toward a Unified Science of the Mind-Brain*. Cambridge, MA, MIT Press.
- Dale, R. (2008). The possibility of a pluralist cognitive science. *Journal of Experimental and Theoretical Artificial Intelligence*, 20, 3, 155-179.
- Dale, R., Dietrich, E. & Chemero, A. (2009). Explanatory pluralism in cognitive science. *Cognitive Science*, 33, 739-742.
- Demailly, A. & Lemoigne, J.L. (éd.) (1986). *Sciences de l'intelligence, sciences de l'artificiel, avec Herbert A. Simon*. Lyon, Presses de l'Université de Lyon.
- Dennett, D.C. (1969). *Content and Consciousness*. Londres, Routledge & Kegan Paul.
- Dennett, D.C. (1978). *Brainstorms. Philosophical Essays on Mind and Psychology*. Cambridge, MA, MIT Press.
- Doty, R.W. (1987). Neuroscience. In O. R. Brobeck, O.E. Reynolds & T.A. Appel (éd.), *The History of the American Physiological Society: The First Century, 1887-1987*. Bethesda, MD, The American Physiological Society.
- Edelman, G.M. (1989). *The Remembered Present: A Biological Theory of Consciousness*. New York, Basic Books.
- Fodor, J.A. (1975). *The Language of Thought*, Cambridge, Harvard University Press.
- Fodor, J.A. (2008). *The Language of Thought Revisited*. Oxford, Oxford University Press.
- Gold, I. & Stoljar, D. (1999). A neuron doctrine in the philosophy of neuroscience. *Behavioral and Brain Sciences*, 22, 5, 809-869.
- Haugeland, J. (1981). The nature and plausibility of cognitivism. *Behavioral and Brain Sciences*, 1, 215-226 ; et in J. Haugeland, *Having Thought*. Cambridge, MA, Harvard University Press, 1998.
- Kellert, S., Longino, H.E. & Waters, K. (éd.) (2006). *Scientific Pluralism*. Minneapolis, University of Minnesota Press.
- Khamassi, M. & Pacherie, E. (2018). L'action. In T. Collins, D. Andler & C. Tallon-Baudry (éd.), *La cognition, du neurone à la culture*. Paris, Gallimard, pp. 270-313.
- Lentin, A. (1982). Rapport sur la recherche cognitive (Rapport introductif au Colloque « Domaines et objectifs de la recherche cognitive ». Pont-à-Mousson, Association pour la recherche cognitive). *Intellectica*, 3, 3-40.
- Lyons, W. (1986). *The Disappearance of Introspection*. Cambridge, MA, MIT Press.
- Millikan, R. (1984). *Language, Thought and Other Biological Categories*. Cambridge, MA, MIT Press.
- Neander, K. (2017). *A mark of the Mental: In Defense of Informational Teleosemantics*. Cambridge, MA, MIT Press.
- Newell, A. & Simon, H.A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ, Prentice Hall.
- Newell, A. & Simon, H.A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM*, 19, 3, 113-126 ; et in Haugeland, J. (éd.)

- Mind Design: Philosophy, Psychology, Artificial Intelligence*. Cambridge, MA, MIT Press, 1981, et de nombreuses autres anthologies.
- Pacherie, E. (2012). Action. In K. Frankish & W. Ramsey (éd.), *The Cambridge Handbook of Cognitive Science*. Cambridge, Cambridge University Press, pp. 92-111.
- Papineau, D. (1987). *Reality and Representation*. Oxford, Blackwell.
- Putnam, H. (1960). Minds and Machines. In S. Hook (éd), *Dimensions of Mind*. New York, New York University Press. Et in H. Putnam, *Mind, Language and Reality, Philosophical Papers volume 2*. Cambridge, Cambridge University Press, 1975.
- Putnam, H. (1988). *Representation and Reality*. Cambridge, MA, MIT Press. Trad. fr. par C. Engel-Tiercelin, *Représentation et réalité*, Paris, Gallimard, 1990.
- Schwitzgebel, E. (2016). Introspection. In *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition).
- Turing, A.M. (1950). Computing Machinery and Intelligence. *Mind*, 59, 3, 435-450 ; et in J. Copeland (éd), *The Essential Turing*. Oxford, Clarendon Press, 2004, chap. 11 ; et dans de nombreuses anthologies.