

Rencontres philosophiques de Langres 20-21-22 septembre 2013

Conférence

« Enjeux contemporains du problème du corps et de l'esprit »

Pascale Gillot, professeure agrégée de philosophie, académie de Créteil

J'examinerai la remarquable résurgence, dans la philosophie contemporaine, du problème de l'union de l'âme et du corps institué à l'âge classique ; cela nous conduira à l'étude de certains débats ou thèmes fondamentaux à l'œuvre aujourd'hui en philosophie de l'esprit, considérée dans la relation étroite qui l'unit à la psychologie, aux sciences cognitives et aux sciences sociales. L'on songe en particulier à la question d'une possible « science de la conscience » ; au problème rémanent d'une explication intégralement physicaliste des phénomènes mentaux ; et surtout, à la ligne de fracture opposant les tenants d'un *internalisme* (les tenants d'une intériorité mentale, fût-elle strictement cérébrale) et les partisans d'un *externalisme* du mental, qui tentent de réassigner l'esprit au « monde », de penser son incorporation dans la sphère publique des pratiques technologiques, symboliques et sociales (les théories de « l'esprit étendu », développées aujourd'hui par Andy Clark et d'autres philosophes comme Alva Noë).

1. La résurgence du Mind-Body Problem dans la philosophie contemporaine de l'esprit

La philosophie du XVII^e siècle, dans le contexte de la révolution scientifique et de la naissance de la physique moderne, propose une conceptualisation nouvelle de l'âme, ou de *l'esprit*. L'esprit, défini désormais comme principe de l'activité mentale, ne peut plus se concevoir, comme dans l'ancienne philosophie héritée d'Aristote, en tant que cause de l'animation du corps. Cette définition nouvelle de l'esprit dans les termes de la *pensée* est appelée par la représentation moderne (« mécaniste ») d'un univers corporel régi par des lois propres, intégralement physiques, en l'occurrence mécaniques.

Elle induit donc l'apparition d'une *question fondamentale*, celle de la relation entre le principe de l'activité mentale, et un monde matériel gouverné par les lois du mouvement et du repos.

Le problème de la relation, en l'individu humain, entre ce qui pense, l'esprit, et ce qui est étendu et appartient à l'univers physique, le corps, est une des versions les plus célèbres de cette question initiale, qui s'élabore dans la philosophie de Descartes. De quelle façon le corps-machine, la « fabrique du corps humain », pour reprendre le vocabulaire cartésien, corps dont les fonctions ordinaires ne requièrent aucune cause psychique, d'une part, et l'esprit-chose pensante d'autre part, peuvent-ils néanmoins se rapporter l'un à l'autre, jusqu'à constituer ensemble l'identité individuelle de l'homme ? Si corps et esprit relèvent de natures complètement différentes, comment comprendre l'apparence d'une détermination du corps au mouvement par cette puissance mentale que serait la volonté (phénomène du mouvement volontaire) ? Et, réciproquement, comment concevoir que l'âme puisse être déterminée à telle ou telle représentation, à la suite des affections, internes ou externes, du corps propre (phénomène de la sensation) ? La difficulté est ici considérable, et continue de hanter la philosophie contemporaine de l'esprit : le fait que l'âme puisse « être affectée par le corps, et inversement », « cela est très difficile à expliquer », écrit en effet Descartes ; et pourtant, ajoute-t-il, c'est précisément ce que nous enseignons de manière évidente « l'expérience » primordiale, qui est

celle-là même, en nous, de l'union essentielle de l'âme et du corps, en jeu par exemple « dans les passions »¹.

Mais la question de la relation entre le mental et le physique concerne également le rapport spécifique entre états psychiques et états cérébraux, entre conscience et organisation neurobiologique, ou entre intentionnalité et disposition cérébrale, par exemple. Elle a aussi trait à la possibilité même d'une « causalité mentale », c'est-à-dire à l'efficace causale de l'esprit dans un univers intégralement physique. Or, il est remarquable que cette série de questions et de problèmes, instituée avec la conceptualisation du mental à l'âge classique, demeure encore largement opérante dans la philosophie contemporaine de l'esprit. C'est ainsi qu'un auteur américain, J. Kim, dans un ouvrage intitulé *Mind in a Physical World. An Essay on the Mind-Body Problem and Mental Causation* (1998), fait de la causalité mentale un des problèmes fondamentaux auxquels continue de se heurter la *philosophy of mind*. Cet auteur évoque même une singulière « revanche de Descartes », pour qualifier l'étonnante persistance, en théorie de l'esprit, de problèmes élaborés dans la philosophie du XVII^e siècle.

De façon très schématique et très générale, donc, le problème du corps et de l'esprit (le *Mind-Body Problem*, ou MBP), tel qu'il se pose encore ou à nouveau aujourd'hui, pourrait aussi se définir comme le problème du **lieu de l'esprit**. Quelle est, quelle peut être la place de l'esprit, dans un monde intégralement physique, régi par ce qu'on appelle le principe de la « *clôture causale* » : principe épistémologique moderne en vertu duquel l'univers, au titre d'univers physique, répond à des lois autonomes, qui ne sauraient être à ce titre « suspendues » ; de sorte que la possibilité même d'une efficace causale du mental, dans cet univers, (causalité mentale qui viendrait « rompre » la série des causes physiques-mécaniques, faire intrusion dans le déterminisme causal physique) prend les traits d'une énigme (cf. sur ce point les analyses proposées par J., Kim, au début du ch. 1 de *Mind in a Physical World*).

Il semble par conséquent que la persistance de ces questions massives, au sujet de la relation entre le physique et le mental, soit l'indice de l'insistance, dans la philosophie contemporaine anglo-saxonne, de conceptions de l'esprit elles-mêmes héritées de la philosophie classique. Cet héritage de la pensée classique dans les débats contemporains autour de l'esprit est ainsi considérable, et néanmoins paradoxal : il est souvent en effet largement méconnu, sinon explicitement dénié comme tel. L'ampleur de ce paradoxe s'exprime du reste dans la curieuse fortune, en philosophie contemporaine de l'esprit, de ce que la tradition a retenu sous le nom de « cartésianisme » et de « dualisme ». Ce cartésianisme, associé à la représentation d'une dualité irréductible du corps et de l'esprit, la philosophie contemporaine n'a généralement cessé de le récuser, au prix d'une lecture parfois tronquée de la doctrine originale de Descartes. Mais la philosophie contemporaine se retrouve cependant nécessairement confrontée à ce même cartésianisme, comme à un spectre qui la hanterait, dès lors qu'elle entend rendre raison d'une certaine spécificité (une spécificité *épistémologique*) du mental ; et ce, bien après l'abandon (quasi général) du dualisme dit « des substances », du dualisme ontologique. Cette persistance paradoxale du cartésianisme se décline principalement sous deux formes, connexes : d'une part, un schème *internaliste*, qui identifie l'activité mentale à une activité interne, privée, s'effectuant en nous, dans l'esprit, la sphère mentale ; la pensée, donc, comme processus intra-mental d'une part ; et d'autre part, un certain modèle « neuropsychologique », lié à la thèse de l'interaction, qui fait du cerveau le principal siège de l'âme, le lieu où celle-ci exerce ses fonctions dans son union avec le corps humain. Sous ces deux derniers aspects, corrélatifs de la conceptualisation moderne de l'activité psychique, la philosophie contemporaine de l'esprit apparaît beaucoup plus redevable de la doctrine de Descartes que ne le laisserait entendre le rejet explicite d'un *cartésianisme* parfois réduit au « mythe », dualiste, de l'indépendance réciproque de deux types de réalité, l'une mentale, l'autre physique.

Dans un tout autre ordre, le programme cognitiviste, ou plus précisément fonctionnaliste (originellement impulsé par les travaux de Hilary Putnam au début des années 60 du XX^e siècle) associe l'esprit à une machine abstraite, à un dispositif computationnel. A ce titre, il reconduit peut-être lointainement, bien au-delà de la « machine de Turing », la figure conceptuelle de l'automate

¹ Descartes, *Entretien avec Burman* (1648), Texte présenté, traduit et annoté par Ch. Adam, Vrin, 1975, p. 71.

spirituel, proposée par Spinoza et développée par Leibniz, concept d'automate spirituel élaboré en rupture avec la définition cartésienne de l'esprit.

Je souhaiterais à présent évoquer les modalités spécifiques de la « réapparition » et de la reformulation, dans la seconde moitié du XXe siècle, du problème classique du corps et de l'esprit.

Très schématiquement, dans la première moitié du XXe siècle, et jusqu'aux années 1960, prévalent des doctrines telles que le monisme « neutre » de W. James, le béhaviorisme philosophique de G. Ryle (*The Concept of Mind*, publié en 1949, influencé par la philosophie de Wittgenstein), ou encore la théorie de l'identité esprit-cerveau (Place et Smart). (Il convient de noter que le **béhaviorisme**, radicalement opposé au « mythe cartésien des deux mondes », et à l'antithèse classique de l'intérieur et de l'extérieur, entend substituer les catégories de disposition et de comportement à la notion « internaliste » de l'esprit, conçu comme sphère mentale ; de sorte que l'esprit, en dernière instance, n'est rien d'autre qu'un certain type de dispositions, de conduites intelligentes, de savoir-faire, dans l'ordre de l'action ; comme telle, la supposée sphère mentale, comparable à une boîte noire insondable, est sans efficace spécifique, et les supposés processus intra-mentaux, en toute rigueur, relèvent d'un mythe philosophique).

Ces doctrines dominantes (le béhaviorisme, les théories de l'identité esprit/cerveau), si différentes soient-elles, visent pourtant toutes à la destitution d'un paradigme classique, celui des deux mondes, dont la philosophie de Descartes serait le lieu d'institution. Or la rupture engagée avec la conception « cartésienne » du mental, s'accompagne du programme de **dissolution du problème classique du corps et de l'esprit**. Sous ce dernier aspect, le refus du dualisme cartésien suppose également, de façon plus déterminante, le rejet des prémisses de la thèse de la distinction réelle, à savoir la partition théorique elle-même du mental et du physique, et, plus fondamentalement encore, la dualité *conceptuelle* de l'intérieur et de l'extérieur.

Or au tournant des années cinquante-soixante du XXe siècle (cf. le célèbre article de H. Feigl, « The « Mental » and the « Physical », 1958), s'engage une première rupture avec le *béhaviorisme*, comme avec les *théories de l'identité strictes*, autrement dit une rupture avec le projet même de mise hors circuit ou de réduction totale des concepts d'activité psychique et d'états mentaux. Cette rupture est fondée sur le constat d'une irréductibilité logique, ou épistémologique, du mental : constat lié à l'échec de la tentative d'établir une équivalence nomologique stricte, terme à terme, entre états mentaux et états neurophysiologiques.

De la sorte, le problème classique du corps et de l'esprit, qui semblait dépassé et comme privé de pertinence, gagne une nouvelle légitimité. Le *mind-body problem* acquiert et conserve en effet le statut, jusque dans la période contemporaine, de « problème philosophique systématique », pour reprendre une formule de Jaegwon Kim². Il semble en effet que la perspective appelée à devenir dominante en philosophie de l'esprit (jusque dans les années 80-90), celle de la nouvelle science de la cognition, ou *cognitivisme* (le modèle de l'esprit-ordinateur), en lien avec la philosophie fonctionnaliste, accorde à la causalité mentale une autonomie (épistémologique) explicite par rapport à la causalité physique. La fameuse thèse de la *réalisation multiple* du mental, promue par le fonctionnalisme, invalide l'hypothèse strictement réductionniste d'une identité des types, c'est-à-dire d'une correspondance terme à terme entre tel état mental et tel état physique, ou neurophysiologique (hypothèse caractéristique des théories de « l'identité esprit-cerveau »). La nouvelle philosophie de l'esprit représentée par le premier fonctionnalisme de Hilary Putnam (fondé sur le modèle computationnel de l'activité mentale, avec l'analogie posée entre l'esprit et l'ordinateur), rompt avec le « matérialisme classique » et avec la conception exclusivement neurophysiologique du mental. Elle entend accorder de la sorte à la « psychologie » le statut d'une science à part entière, contre le postulat physicaliste traditionnel, qui prétendait rendre raison des événements mentaux dans le vocabulaire de la science de la nature.

Simultanément, dans le contexte de la philosophie de l'esprit des années soixante, et de la réaffirmation d'une spécificité des concepts d'activité mentale et de rationalité, la théorie linguistique de Noam Chomsky promeut une représentation générale du fonctionnement de l'esprit humain en

² J. Kim, *Mind in a Physical World* [1998], MIT Paperback edition, 2000, p. 2.

opposition remarquable à la tradition empiriste et behavioriste jusqu'alors déterminante dans la philosophie anglo-saxonne. Le *mentalisme* revendiqué par N. Chomsky, dans le cadre du postulat général d'une « grammaire générative », réactive la thèse d'une innéité des fonctions psychiques. Il engage ainsi une reprise, assurément singulière, de ce cartésianisme que la philosophie du langage ordinaire, de Wittgenstein à Ryle, à travers la critique du solipsisme et de « l'antithèse de l'intérieur et de l'extérieur », avait assigné au statut de mythe philosophique. Cette réhabilitation chomskienne de la tradition rationaliste et mentaliste contribue de la sorte, en un retournement remarquable, à la destitution du behaviorisme logique, et à la résurgence, pour le moins hétérodoxe, de thématiques cartésiennes dans le champ de la philosophie de l'esprit.

Le fonctionnalisme, qui signe l'abandon de la thèse de l'identité esprit-cerveau, et la théorie linguistique « innéiste » de Chomsky, résolument opposée au behaviorisme, constituent sans doute deux conceptions distinctes du mental. Cette divergence de l'une à l'autre se trouve du reste clairement soulignée par Putnam lui-même, qui critique « l'hypothèse de l'innéité (*the innateness hypothesis*) » en jeu dans la représentation chomskienne d'« universaux linguistiques »³. Cependant ces deux perspectives philosophiques, fonctionnaliste et innéiste, si différentes soient-elles, participent toutes deux du tournant *cognitivist* en philosophie de l'esprit ; elles ont pour trait définitionnel commun le postulat d'une *autonomie explicative ou épistémologique* de l'activité mentale. Ce postulat, s'il implique un dépassement du réductionnisme traditionnel, pose une certaine *distinction* (conceptuelle, logique) de l'esprit par rapport au corps et à la causalité physique ; mais il n'engage pourtant pas de retour au dualisme ontologique classique. Telle est aussi la leçon du « monisme anomal » développé par Donald Davidson (*Actions et événements [Essays on Actions and Events*, Oxford University Press, 1980]. Celui-ci récuse la possibilité d'un isomorphisme causal de l'activité mentale et de l'activité corporelle, tout en maintenant l'axiome du caractère intégralement physique des événements du monde, contre le dualisme classique. L'essentielle « anomalie » du mental (son « holisme ») serait la marque, non d'une *réalité* particulière de l'esprit, mais de son irréductibilité théorique par rapport au déterminisme physique universel. Ce qui demeure fondamentalement en jeu, c'est la thèse, d'abord négative, de l'inadéquation des schèmes réductionniste et physicaliste pour le programme d'une *explication* des opérations de l'esprit.

2. L'énigme de la conscience : une science du mental est-elle possible ?

Qu'en est-il, précisément, de la supposée irréductibilité du mental, et celle-ci, si elle existe, doit-elle s'entendre comme l'irréductibilité de l'*expérience phénoménale* (liée à ce qu'on appelle les *qualia*, qui renvoient à la dimension subjective de l'expérience, accessible uniquement en première personne : « ce que cela fait de »). Se rencontre, ici, ce que certains philosophes contemporains (tels David Chalmers, mais aussi Colin McGinn, par exemple, et enfin John Searle) appellent le « problème difficile », l'énigme au cœur du *Mind-Body Problem*, à savoir « l'énigme de la conscience ». La question de savoir si cette énigme de la conscience est susceptible d'être finalement résolue, sur un mode scientifique, de sorte que la conscience pourrait perdre en dernière instance son statut de phénomène mental irréductible ; cette question est au centre des polémiques des débats les plus vifs aujourd'hui en philosophie de l'esprit.

Je m'attacherai plus particulièrement ici au débat structurant et paradigmatique entre Dennett et Searle, concernant cette question de la possibilité d'une explication scientifique, c'est-à-dire en troisième personne, de la *conscience* et des *qualia*. (Pour une histoire de ce long débat, cf. notamment l'ouvrage de J. Searle, *Le Mystère de la conscience, suivi d'Echanges avec Daniel C. Dennett et David J. Chalmers [The Mystery of Consciousness*, Granta Books, 1997).

La tradition dans laquelle s'inscrit Dennett, précisons-le, est celle du *fonctionnalisme* : celle-ci est opposée aussi bien aux théories de l'identité mental-cérébral qu'aux conceptions internalistes de l'esprit (qui posent le caractère privé des phénomènes mentaux, l'accès privilégié du sujet à ses états mentaux). A ce titre, il a pour adversaire privilégié John Searle qui, dans son ouvrage de 1992, *The Rediscovery of The Mind*, proposait une critique virulente du cognitivisme (la critique du modèle de

³ H. Putnam, « The « innateness hypothesis » and explanatory models in linguistics », 1967, in *Mind, Language and Reality, Philosophical Papers*, vol. 2, Cambridge University Press, 1975, pp. 107-116.

l'esprit-ordinateur), au nom précisément de l'évidence absolue de la réalité de l'expérience en première personne, de l'expérience phénoménale, en jeu dans les *qualia* (suivant, donc, un schème internaliste du mental ; cf. le célèbre argument de la chambre chinoise).

Pour cerner certains points cruciaux de ce débat, je m'attacherai plus particulièrement à l'analyse d'un livre – relativement – récent de Dennett, intitulé *De beaux rêves. Obstacles philosophiques à une science de la conscience* [traduction parue en 2008 aux Editions de l'éclat (*Sweet Dreams. A philosophy of Mindfulness and Somaesthetics*, MIT Press, 2005)], Daniel Dennett propose de poursuivre, de réviser et de renouveler la théorie fonctionnaliste de l'activité mentale, et plus précisément de la conscience, qu'il avait d'abord exposée dans son ouvrage de 1991, *La conscience expliquée* [*Consciousness explained*].

Cette reprise et ce renouvellement s'appuient sur une série de recherches scientifiques récentes, notamment en neurophysiologie, en psychologie et en sciences cognitives. Mais le postulat général qui commande la recherche de Dennett en théorie de l'esprit depuis plus de trente ans demeure fondamentalement inchangé. Il pourrait s'énoncer en ces termes : une théorie *naturaliste* et *mécaniste* de l'esprit - et en particulier de ce qui, en l'esprit, est traditionnellement considéré comme le plus difficilement objectivable, *la conscience* – cette théorie est non seulement concevable, mais déjà en voie de constitution. Fondée sur une approche épistémologique *en troisième personne*, elle représente la seule explication recevable des phénomènes mentaux, s'il est vrai que toute science s'édifie au point de vue de la troisième personne, et qu'il n'existe tout simplement pas de science en première personne. Ainsi se conçoit le projet philosophique d'une véritable *science de la conscience*, commandée par un modèle computationnel de l'esprit, et engageant par conséquent la « science des ordinateurs », mais également ouverte aux recherches conduites dans le champ des neurosciences. De façon générale, Dennett reste partisan d'un fonctionnalisme au sens large, et privilégie donc le modèle de l'esprit-ordinateur ; il récuse cependant une version dure du fonctionnalisme qui, au nom de la thèse de la réalisabilité multiple du mental, donc du caractère indifférent des modes de réalisation matérielle de l'automate logique, tend traditionnellement à négliger l'apport de la biologie et des neurosciences. La réconciliation du cognitivisme et des neurosciences autorise *de fait* une explication objective du mental ; si bien que la notion, en philosophie de l'esprit, d'un essentiel « mystère » de la conscience, l'hypothèse d'une dimension nécessairement privée du mental, sont autant d'obstacles philosophiques à une exploration scientifique et objective de l'esprit. Ce sont ces obstacles que Dennett entend dissiper, puisqu'ils entravent « la recherche scientifique en cours, celle qui peut expliquer la conscience de façon aussi profonde et complète que d'autres phénomènes naturels : le métabolisme, la reproduction, la dérive des continents, la gravitation et ainsi de suite » (ch. 2, p. 39).

Ainsi s'explique le combat engagé par Dennett contre les variantes contemporaines, en philosophie de l'esprit, de ce « Théâtre cartésien » qui fait de la conscience un phénomène intérieur, accessible exclusivement en première personne, à la différence des phénomènes physiques. Poursuivant une controverse depuis quelques temps engagée avec d'autres théoriciens de l'esprit tels que John Searle, mais aussi David Chalmers, Thomas Nagel ou encore Ned Block, Daniel Dennett s'attache à débusquer les impasses d'un anti-réductionnisme qui, sous prétexte de prendre la mesure du caractère subjectif de certains états mentaux apparemment distincts des attitudes propositionnelles (les « *qualia* »), pose un écart infranchissable entre les sciences de la nature et la philosophie de l'esprit. C'est là l'enjeu de la critique par Dennett du concept philosophique communément reçu de *qualia*, ces supposées propriétés phénoménales, qualitatives et intrinsèques liées à « l'effet que cela fait de » (contempler telle nuance de jaune, ressentir telle douleur singulière), dans l'ordre d'une expérience en première personne (Chapitre 4). Cette analyse critique, prolongeant l'entreprise de « disqualification des *qualia* » amorcée dans l'ouvrage de 1991, s'appuie sur des recherches scientifiques relevant des neurosciences et de la neurophysiologie, consacrées notamment au phénomène perceptif de la « cécité au changement » : celles-ci révèlent une possible dissociation entre la possession de *qualia* et l'expérience subjective consciente immédiate, l'hypothèse de propriétés phénoménales non conscientes manifestant ainsi toute l'ambiguïté généralement insoupçonnée du concept classique de *qualia*. Ce travail de questionnement et de dissolution d'un certain nombre de pseudo-évidences philosophiques concernant le caractère privé de la conscience s'autorise également de la discussion critique détaillée et aigüe de thèmes et de dispositifs théoriques traditionnels en philosophie de l'esprit. (Il s'agit notamment de « l'intuition du zombie » (ch. 1), qui suppose que la possession d'états mentaux dits « internes » est parfaitement distincte du comportement observable des sujets. Il s'agit aussi de l'expérience de pensée proposée par Franck

Jackson mettant en scène une scientifique, « Marie », spécialiste de la neurophysiologie de la vision des couleurs, enfermée dans un laboratoire monochromatique et travaillant sur des écrans en noir et blanc : la « libération » de Marie et son accès soudain au monde réel et à sa richesse chromatique sont censés provoquer en elle, malgré tout son savoir théorique, la surprise et l'enchantement, « preuve » de l'irréductibilité de l'expérience en première personne à toute description scientifique et objective (ch. 5).)

L'investigation conduite par Dennett autour de concepts récurrents de la *philosophy of mind* (les *qualia*, la phénoménalité, les propriétés intrinsèques), afin d'en dévoiler l'équivocité constitutive ; cette investigation peut s'entendre comme une tentative de dissolution d'un certain nombre de « mythes philosophiques » hérités de la conception cartésienne de l'esprit, et particulièrement persistants dans la philosophie contemporaine de tradition analytique. Si elle implique des thèses délibérément « contre-intuitives », aussi stimulantes que roboratives, au sujet du mental, elle ne renoue pas cependant avec un *béhaviorisme strict*, qui nierait par exemple l'existence même d'états mentaux « internes ». Il existe bien quelque chose comme une expérience subjective en première personne, concède Dennett, mais ces états « subjectifs », tant qu'ils ne sont pas l'objet d'une enquête nécessairement menée en troisième personne, ne sont qu'un « résidu » non significatif, et ne sauraient comme tels constituer les *données* d'une véritable théorie de l'esprit.

Ainsi se trouve justifiée une approche scientifique de la conscience humaine, en troisième personne, non éliminativiste ni béhavioriste pour autant, que Dennett nomme *hétérophénoménologie* (ch. 2, pp. 50-62) (Ce qui implique, par conséquent, que l'esprit ne soit pas réductible exclusivement à un fondement *biologique*, comme l'activité cérébrale). L'hétérophénoménologie traite des « données » de la conscience sur un mode *neutre* à partir principalement des « compte rendus verbaux » et des textes produits par les sujets. Sa neutralité et son objectivité en font selon D. Dennett la seule méthodologie adéquate d'une science de la conscience ; cette méthodologie scientifique n'interdit pas toutefois une « posture intentionnelle », puisqu'elle implique effectivement une analyse des structures des croyances subjectives. Le cadre théorique général demeure celui de l'attribution de croyances, attribution qui s'effectue à partir des jugements verbalisés et des actes de langage, et non par la voie de l'introspection ni par la voie de l'empathie. Le postulat fondamental de la méthode hétérophénoménologique est ainsi qu'il n'existe pas de croyances inexprimables, par quoi elle se distingue d'un internalisme qui suppose que les croyances ne peuvent être analysées qu'en première personne.

Enfin, la contribution de Dennett à l'édification positive d'une science de la conscience prend principalement les traits d'une théorie, la théorie de la « *notoriété* » ou encore de la « *célébrité cérébrale* » (ch. 6 et ch. 7), qui prolonge et modifie simultanément le modèle des « versions multiples » précédemment exposé dans *La conscience expliquée*. Associé à la notion, issue des neurosciences cognitives, d'« espace de travail global », ce modèle explicatif de la conscience fait de celle-ci, non pas un phénomène élémentaire et indivisible, mais bien la résultante d'une sorte de division du travail d'une multitude d'agents eux-mêmes non conscients et parfois concurrents les uns des autres. Il engage l'hypothèse d'une composition du travail cérébral en de multiples instances (le modèle du « Pandémonium »), qui s'oppose explicitement à celle d'un « Témoin » ou d'un « Exécutif central » caractéristique du Théâtre cartésien. La conscience peut de fait se réduire à « l'accessibilité des démons spécialistes [les multiples agents du travail cérébral] les uns envers les autres (et non envers un Exécutif ou un Ego Central imaginés comme étant d'un rang plus élevé) » (p. 156). Pour reprendre la métaphore politique proposée par Dennett, cette conception computationnelle et anti-hiérarchique s'oppose à la définition de la conscience comme « médium de représentation particulier situé dans le cerveau », de la même façon que le modèle démocratique, s'oppose au modèle monarchique d'un « Soi » qui régnerait sur les contenus mentaux et en constituerait le centre.

Cette modélisation originale engage un travail conceptuel important de redéfinition de la catégorie même de conscience : celle-ci n'est plus identifiée à « une propriété *intrinsèque*, ni même une simple propriété *dispositionnelle* », mais elle se trouve ici comprise comme « un phénomène qui exige l'actualisation d'un potentiel ». Le recours aux concepts neuroscientifiques de « réverbération » et de « courbe d'amplification soutenue » (en référence aux travaux de Dehaene et Naccache, p. 158) se conjoint de façon remarquable avec la détermination littéraire, d'obédience proustienne, de l'écho ou de la « capacité de faire écho », pour aboutir à une compréhension singulière de la *réflexivité*, désormais disjointe de toute « *phénoménalité* » (pp. 187-188). Outre donc le travail d'investigation critique et de dissolution de concepts philosophiques traditionnels en théorie de l'esprit, l'un des

principaux intérêts de la démarche de Dennett aujourd'hui consiste en une redéfinition de la conscience qui ne nie pas celle-ci, mais tente d'en produire une conceptualisation nouvelle en dehors de la dite « tradition cartésienne », avec notamment la double référence, originale, aux neurosciences et à la fiction littéraire. L'on pourrait certes regretter que cette réforme conceptuelle de la conscience demeure à l'état d'esquisse, en ce qui concerne plus particulièrement les enjeux fictionnels et littéraires des « récits de soi » des sujets de l'hétérophénoménologie.

Surtout, et c'est un point décisif, l'on peut se demander dans quelle mesure la perspective des neurosciences, fondamentalement internaliste, dans la mesure où elle rapporte les états mentaux à leurs corrélats neuraux *dans* le cerveau, est véritablement compatible avec la perspective *externaliste* également développée dans l'ouvrage. L'externalisme (revendiqué par Dennett) refuse en effet la représentation d'un Soi séparé du monde, « isolé du monde extérieur par des couches d'« input » et d'« output » », puisqu'au contraire « le Soi est massif, concret, et visible dans le monde, pas seulement « distribué » dans le cerveau mais répandu dans le monde » (ch. 6, p. 159).

Enfin, la voie hétérophénoménologique propose une redéfinition singulière du Soi (contre la figure classique d'un « Ego central »), et insiste sur la *réflexivité* immanente aux institutions symboliques et sociales, comme « capacité de faire écho (...) en grande partie imputable aux habitudes d'autostimulation que nous importons de la culture humaine » (ch. 7, p. 195).

De façon générale, la perspective ici proposée d'une théorie du sujet hétérophénoménologique, réassignant le sujet à la sphère *anthropologique* et symbolique, participe d'une des voies les plus fécondes dans le champ de la philosophie contemporaine de l'esprit. Cette voie, donc, c'est celle de l'*externalisme* (non éliminativiste), que je vais à présent plus précisément évoquer.

3. L'externalisme en question : les théories de « l'esprit étendu ».

Le problème du « lieu de l'esprit ».

Je partirai de la question – persistante – d'une localisation des phénomènes mentaux en général (cette question figurant une version contemporaine cruciale du problème du corps et de l'esprit). Il s'agit là d'une question récurrente autant que problématique, dans la mesure où elle engage simultanément le constat d'une inscription de l'esprit dans le monde, inscription « matérielle » au sens large pourrait-on dire, et l'impératif épistémologique d'une compréhension de la spécificité *conceptuelle* des événements mentaux, de l'activité mentale. Concevoir *l'esprit et le corps* à la fois comme un et comme deux, tel est le problème, qui trouve sa formulation originale dans la philosophie de Descartes, et traverse l'histoire de la philosophie, au moins depuis le XVII^e siècle.

Mais cette question, celle du « lieu » de l'esprit, s'étend bien au-delà des débats autour du dualisme classique et de ses multiples réfutations historiques. Elle constitue, pourrait-on dire, une sorte d'invariant structurel, en philosophie et en sciences sociales, puisqu'elle est reliée à un questionnement général concernant les modalités d'effectuation et d'*institution* de la pensée.

De manière générale, la question du lieu incertain, ou problématique, de l'esprit, est intéressante, dans la mesure où elle fait porter l'attention sur *les multiples manifestations du mental*, considérées en relation avec la sphère publique du monde symbolique et social, et non pas seulement en référence à une identité psychophysique de type individuel. L'esprit par conséquent, ce qui pense, perçoit, mais aussi ce qui agit, désire, ressent, n'occupe peut-être pas *nécessairement*, ou par définition, un *lieu* privilégié, comme une sphère mentale individuelle, un cerveau (ou une partie du cerveau), ou même un certain corps du monde, un corps individuel.

Tel est donc l'axiome de cette *philosophie externaliste contemporaine*, que je voudrais à présent évoquer (philosophie externaliste que l'on pourrait tenter de représenter par une ligne, dans l'histoire de la philosophie, ligne qui partirait de, Spinoza, à l'âge classique, se prolongerait vers James et Dewey, dans la tradition du pragmatisme américain, passerait par le second Wittgenstein, et irait pour l'époque contemporaine, jusqu'à Hilary Putnam et Alva Noë aux Etats-Unis, Andy Clark au Royaume Uni, ou Vincent Descombes en France).

Quoi qu'il en soit, il ne s'agit pas, suivant l'axiome externaliste, d'assigner au mental une « demeure » ou un « habitacle » particulier, pour reprendre la formule - polémique – de Spinoza lui-même (*Ethique* II, 35 Scolie); mais il s'agit bien plutôt de comprendre l'inhérence constitutive de l'esprit au monde, (son *embodiment* ou incorporation au sens large, son caractère « situé » et « distribué » dans l'environnement) : c'est-à-dire non seulement son inhérence à l'univers physique, mais également à la sphère *anthropologique* des pratiques sociales, symboliques, économiques ou politiques. S'engage alors une critique de la traditionnelle « philosophie mentale » (pour reprendre l'expression de V. Descombes dans *La denrée mentale*), philosophie mentale fondée sur le postulat d'une séparation (d'une délimitation fixe) entre l'esprit et du monde.

De manière générale, l'externalisme ici évoqué conteste le « partage » traditionnel entre philosophie de l'esprit et philosophie de l'action ; puisque dans la perspective externaliste elle-même, l'esprit est agent, nécessairement. En d'autres termes l'esprit (la pensée), loin de se « cantonner » à une sphère mentale privée, à une intériorité psychique, ou même à un certain substrat organique (neurophysiologique), se distribue dans le monde : il est susceptible de s'étendre, littéralement, à des dispositifs techniques, à des artefacts, eux-mêmes toujours plus perfectionnés, et qui sont autant d'outils indispensables à la cognition et à l'intelligence : depuis la plume, le parchemin, le stylo pour écrire, le livre, l'agenda, la machine à calculer, jusqu'à l'ordinateur, le smartphone, la tablette, etc.

De façon générale, donc, le postulat d'une philosophie non « internaliste » de l'esprit (ou « non cartésienne ») serait celui du caractère *irréremédiablement incarné et situé* (*embodiment* et *embeddedness*, comme l'écrivent les penseurs anglophones) des manifestations intentionnelles. Mais la difficulté est alors de spécifier la portée et le type d'*embodiment* et d'*embeddedness* ici en jeu.

Pour préciser cela davantage, l'on peut s'appuyer sur une distinction explicitée récemment par Andy Clark concernant la manière d'envisager la contribution du corps aux états et aux contenus mentaux⁴. A cet égard, deux modèles s'opposent. Le premier modèle affirme l'existence d'une contribution spéciale du corps, de *notre* corps, dans la détermination des contenus, états et propriétés mentaux. La présence d'esprits de type humain *dépend* alors *directement* de la possession de corps de type humain.

Suivant le second modèle en revanche (privilegié par Clark), il convient d'admettre que corps et (parties du) monde puissent compter comme éléments de systèmes plus larges, comprenant donc corps, cerveaux, mouvements des organes sensoriels, états informationnels d'artefacts de divers types, etc.; systèmes plus larges dont les états déterminent, ou aident à déterminer, les propriétés et états mentaux.

Il s'agit là de la théorie du « fonctionnalisme étendu », dont le célèbre article de 1998 (co-écrit par Andy Clark et David Chalmers) intitulé « The Extended Mind »⁵, constitue le manifeste programmatique. Cette théorie, du « fonctionnalisme étendu », ou de « l'esprit étendu », est en plein essor aujourd'hui. Il s'agit d'un « externalisme actif », qui affirme le rôle causal de l'environnement dans l'institution du « mental », et se présente comme une perspective plus radicale encore que « l'externalisme standard » revendiqué dans les années soixante-dix par des auteurs comme Hilary Putnam ou Tyler Burge.

L'externalisme actif pose l'existence d'un rôle causal immédiat (de l'environnement), avec la notion de « systèmes couplés » liant l'organisme humain et l'environnement (naturel ou artificiel), comme les artefacts et les dispositifs informationnels et technologiques : ces systèmes couplés recoupent les « systèmes plus larges » précédemment évoqués, dont la définition même récuse le partage classique de l'esprit et du « monde extérieur ».

⁴ Cf. en particulier les ouvrages d'Andy Clark, *Being There: Putting Brain, Body and World Together Again*, MIT Press, Bradford Books, 1997, et *Natural-Born Cyborgs: Minds, Technologies and the Future of Human Intelligence*, New York/ Oxford, Oxford University Press, 2003. *Supersizing the Mind. Embodiment, Action and Cognitive Extension*, Oxford University Press, 2008.

⁵ David Chalmers et Andy Clark, « The Extended Mind », *Analysis* 58, p. 10-23, 1998.

La théorie du fonctionnalisme étendu, à l'oeuvre dans cet « externalisme actif », « distribue » ainsi la cognition et l'esprit dans l'environnement et les objets qui servent de « support » et de conditions aux pratiques (intelligentes / cognitives). Or il est remarquable que le statut même du *corps* se trouve profondément modifié par ce fonctionnalisme *étendu*. Clark soutient ainsi que « rien ici ne requiert un corps unique persistant dans l'espace tridimensionnel ordinaire. Au lieu de cela, il pourrait y avoir des formes authentiques mais dispersées d'incorporation, incorporation dans des réalités virtuelles ou mixtes, et de multiples incorporations pour une intelligence unique ».

Pour résumer, l'on pourrait répertorier *cinq axiomes de l'externalisme actif*, qui contribuent à la distorsion des conceptions traditionnelles de l'esprit et de son rapport au corps et au monde physique. (Cf. l'article de Andy Clark, « Minds, Brains and Tools », 2002, qui est l'occasion d'une discussion très serrée des thèses de Daniel Dennett à propos du mental. Cf. également un ouvrage récent de Andy Clark, avec une préface très éclairante de David Chalmers, intitulé *Supersizing the Mind. Embodiment, Action and Cognitive Extension*, Oxford University Press, 2008).

1er axiome : l'esprit, loin de se définir en référence exclusive à une certaine détermination biologique innée (liée au cerveau humain) est fondamentalement le produit historique de la culture et des institutions humaines.

2e axiome : l'activité mentale, loin de s'identifier à l'expérience en première personne, est une *activité symbolique*. Penser, fondamentalement, revient à manipuler des symboles, des signes, et non à posséder des états internes « représentationnels ». A cet égard, le langage joue assurément un rôle déterminant, et détient le statut -privilegié - de « méta-outil », pourrait-on dire, dans l'extension même des outils de la pensée.

3e axiome : L'esprit peut se concevoir comme un dispositif technologique – un artefact complexe. (Reprise sur ce point de la thèse du fonctionnalisme classique).

4e axiome : La philosophie de l'esprit tend vers une philosophie de l'action. L'esprit est un agent, incorporé dans les savoir-faire, les conduites, les instruments, mais aussi dans les institutions **sociales, culturelles**.

5e axiome : Il n'existe pas de sujet privilégié de la pensée et de la représentation (cf. la fameuse proposition de Wittgenstein dans le *Tractatus logico-philosophicus*, 5. 631). L'esprit ne se conçoit pas dans les termes d'un « sujet pensant », ni dans ceux d'une conscience unifiée ou unifiante (la critique du supposé « théâtre cartésien »). Il n'est pas de l'ordre de l'intériorité.

Ainsi s'entend la définition de l'esprit comme configuration plurielle d'outils intelligents, ou d'outils pensants, sans instance centrale utilisatrice privilégiée. « *Minds are grab-bags of (userless) tools for thought* » (Andy Clark, « Minds, Brains and Tools », 2002).

Il apparaît donc que, suivant les différents auteurs partisans de l'externalisme (Andy Clark, Daniel Dennett, Alva Noë, Vincent Descombes), il faille concevoir l'esprit dans son immanence au monde symbolique et anthropologique lui-même : ainsi peuvent s'entendre, dans leur distinction même, les figures de l'esprit-artefact, de l'esprit étendu, de l'esprit « collectif » (cf. encore le récent livre de Vincent Descombes, *Les embarras de l'identité*, Gallimard, 2013). Celles-ci sont autant de variations sur le philosophème de « l'esprit dans le monde » (philosophème opposé au schème de l'intériorité mentale, « l'esprit dans la tête », auquel se rattache la représentation neuropsychologique traditionnelle, autant que problématique, de l'esprit « dans le cerveau » - ou dans une partie du cerveau).

L'on conçoit, assurément la fécondité d'une telle perspective, en l'occurrence l'externalisme : celui-ci, en posant l'existence d'institutions de l'esprit, liées à son inscription nécessairement anthropologique, sociale, technologique, permet d'échapper, non seulement au « mythe de l'intériorité », mais aussi à l'étroitesse du programme physicaliste (la voie neuropsychologique, l'esprit cantonné à « l'intériorité cérébrale », ou le « parallélisme » psychophysiologique déjà dénoncé par Wittgenstein en son temps).

Une série de problèmes insistent et persistent cependant, dans le cadre même de cette voie externaliste : celui du partage *logique* (et non empirique) entre intérieur et extérieur (avec la représentation afférente, difficilement éliminable, d'une intériorité psychique), et le problème (connexe) de l'individuation psychophysique (qui renvoie à la question classique de l'*identité personnelle*). Dans les termes de la théorie de l'*extended mind* en effet, il semble que l'esprit ne puisse plus se trouver identifié ni même assigné à un certain *corps* : dans la mesure même où l'esprit est dans le monde, et doit se concevoir dans son couplage avec des dispositifs artificiels (les outils-soutiens de l'activité mentale).

Résumons : Il existe un couplage structurel de l'esprit et de l'outil ; et l'esprit se définit lui-même comme une configuration complexe d'outils, pouvant constamment être augmentée, renforcée, étendue : ces caractérisations externalistes du « mental » engagent toutes une « désindexation » de l'esprit par rapport au corps humain individué, au corps organique de l'homme (désindexation qui se trouvait déjà impliquée, du reste, dans la thèse fonctionnaliste de la « réalisation multiple », opposée aux théories classiques de l'identité mental-cérébral).

Non seulement donc le lieu privilégié du mental n'est plus le « cerveau », mais il n'est plus même le *corps*, tout au moins considéré dans son individualité. Quels sont, dès lors, les critères d'identification de l'esprit, de la subjectivité elle-même, et quelles sont les « limites » du mental ?

Pour conclure, je m'arrêterai sur cette tension, à l'œuvre dans les théories de l'esprit étendu aujourd'hui.

L'on pourrait en effet considérer que l'externalisme contemporain engage à nouveau, dans la perspective de l'esprit étendu, ou de la nature sociale du mental, la dissolution du problème du corps et de l'esprit. Mais à ce titre précisément, l'externalisme permet-il encore de rendre raison du rapport à soi et de la subjectivation ? L'apparente éviction qu'il propose de la notion du « sujet » de la pensée et de la représentation, avec le rejet du modèle du « Théâtre cartésien », ne peut-elle s'entendre comme une manière d'esquiver la question, récurrente, de la subjectivité, plutôt que de tenter de la résoudre ? (cf. la question posée par David Chalmers, à propos de la difficile notion problématique d'une « conscience étendue », dans sa préface au livre de Andy Clark, *Supersizing the Mind*, 2008).

Telles sont les difficultés majeures à l'œuvre dans ce courant de la philosophie contemporaine de l'esprit qui s'est affranchi de la voie neuropsychologique comme de l'internalisme.

Et pour formuler autrement l'ampleur de la difficulté, nous pourrions en revenir à la question du « lieu de l'esprit » sous-jacente au MBP :

Si l'externalisme, dans son ordre spécifique, paraît en effet poser la question des lieux de l'esprit, il engage un questionnement à propos, notamment, des limites spatiales et du « champ d'action » de l'esprit (cerveau, peau, artefacts, environnement proche ou lointain). Mais peut-être suppose-t-il finalement, en dernière instance, l'absence de sens de cette question (« Où se situe l'esprit ? ») : simplement du fait de la résistance qu'opposent les « lieux » éventuels de l'esprit que sont les signes, les institutions, les artefacts, les couplages organiques-technologiques, à une localisation fixe et spécifique. L'on pourrait se demander, dès lors, dans quelle mesure la métaphore spatiale conserve une véritable pertinence, au moins dans le cadre d'une philosophie de l'esprit distincte de la « philosophie mentale » traditionnelle.

Bibliographie

Tyler Burge, *Foundations of Mind*, Oxford University Press, 2007.

Andy Clark, *Being There : Putting Brain, Body and World Together Again*, MIT Press, 1997.

Supersizing the Mind. Embodiment, Action and Cognitive Extension, Oxford University Press, 2008

David Chalmers,

The Conscious Mind, Oxford University Press, 1996

The Character of Consciousness, Oxford / New York, Oxford University Press, 2010.

Daniel Dennett, *De beaux rêves. Obstacles philosophiques à une science de la conscience*, traduction française de Claude Pichevin, Paris, Editions de l'Eclat, 2008.

Herbert Feigl, *Le « Mental » et le « Physique »*, traduction de l'anglais de Christine Lafon et Bernard Andrieu, traduction de l'allemand de Fabien Schang, Paris, L'Harmattan, 2002.

Jaegwon Kim, *L'esprit dans un monde physique. Essai sur le problème corps-esprit et la causalité mentale*, traduction française de François Athané et Edouard Guinet, Editions Syllepse, 2006. [*Mind in a Physical World. An Essay on the Mind-Body Problem and Mental Causation*, MIT Press, 1998].

Thomas Nagel, *Le point de vue de nulle part*, traduction française de Sonia Kronlund, Combas, Editions de l'Eclat, 1993.

Alva Noë, *Out of our Heads. Why you are not your Brain, and Other Essays from the Biology of Consciousness*, Hill and Wang, 2009.

Hilary Putnam, *The Threefold Cord : Mind, Body, and World*, New York, Columbia University Press, 1999.

John Searle, *La redécouverte de l'esprit*, traduction française de Claudine Tiercelin, Paris, Gallimard, 1995. [*The Rediscovery of the Mind*, 1992].

Le Mystère de la conscience, suivi d'Echanges avec Daniel C. Dennett et David J. Chalmers, traduction française de Claudine Tiercelin, Paris Odile Jacob, 1999 [*The Mystery of Consciousness*, Granta Books, 1997].