

L'esprit

Le fonctionnalisme selon Daniel Dennett ou : Dennett a-t-il perdu l'esprit ?

Xavier Kieft

Philopsis : Revue numérique
<http://www.philopsis.fr>

Les articles publiés sur Philopsis sont protégés par le droit d'auteur. Toute reproduction intégrale ou partielle doit faire l'objet d'une demande d'autorisation auprès des éditeurs et des auteurs. Vous pouvez *citer* librement cet article en en mentionnant l'auteur et la provenance.

1. Les enjeux d'une théorie

Quel statut l'explication fonctionnaliste de la conscience accorde-t-elle à l'esprit ? Les concepts mentaux, tels qu'ils sont manipulés par les théoriciens matérialistes, désignent-ils bien leur objet ou ont-ils un rôle analogique ? Autrement dit : le fonctionnement de la conscience explique-t-il la conscience elle-même ou bien se situe-t-il en amont de celle-ci comme ce qui la rend possible ou la cause ? La réponse à ces interrogations doit permettre de mieux concevoir comment certains défenseurs du monisme matérialiste cherchent à résoudre le problème corps-esprit. En effet, si les concepts mentaux auxquels il est fait référence dans leurs théories ne correspondent pas à ce dont je fais moi-même l'expérience, alors les physicalistes dissimulent sous leurs explications ce qu'est réellement le mental par opposition au physique. Ceci soulignerait son irréductibilité. Si, par contre, les concepts qu'ils envisagent désignent bien ce qui est présent dans l'activité mentale ou intentionnelle, alors cette entreprise réussira à dissoudre le problème corps-esprit par l'assimilation de l'esprit au fonctionnement du corps.

Au premier abord, si l'on parle de computation neuronale pour rendre compte de la pensée, il semble bien que lorsque je pense, je ne sens pas ma computation neuronale, et que je ne suis pas forcément conscient de cette

activité computationnelle. Mais ceci n'est peut-être qu'un problème de vocabulaire. Mon ignorance des théories scientifiques ou la traduction en langage ordinaire de ce qui se passe dans ma conscience est susceptible de m'empêcher de concevoir mes pensées comme (le produit d'une) computation neuronale. Néanmoins la différence entre mon rapport immédiat à ma conscience et les théories de l'esprit témoigne peut-être d'une distance entre le contenu de l'explication scientifique et ce qu'est effectivement la vie psychique. Est-il donc seulement possible de concevoir l'esprit par le biais des notions de mouvement, de quantité, de parties, c'est-à-dire par les notions grâce auxquelles nous concevons les choses matérielles ? La réussite du programme fonctionnaliste serait la preuve que le corps et l'esprit ne sont pas deux choses incompatibles ni deux sortes de propriétés distinctes d'une même chose, et qu'ils peuvent être conçus et décrits de façon univoque.

La pensée de Daniel Dennett, que l'on aurait trop vite fait de présenter comme le porte-parole contemporain d'un éliminativisme matérialiste, traduit le souci d'assurer au fonctionnalisme sa pertinence en ce qui concerne la résolution de ce problème. La théorie qu'il met en place s'efforce de maintenir la particularité des états mentaux, qui participent en tant que tels à l'activité consciente et n'ont plus alors un rôle uniquement surrogatoire ou d'accompagnement (de décoration) du processus cérébral. Certes, la conscience émerge de l'organisation du cerveau, mais elle n'est pas qu'un épiphénomène de l'activité cérébrale.

L'objectif de cette présentation est d'abord de montrer comment l'explication qu'il propose, qui refuse tout compromis avec un dualisme des substances ou des propriétés, permet de rendre compte de l'esprit sans l'éliminer (§ 2). Mais, en se demandant si cette théorie, pour élégante qu'elle soit, ne laisse pas de côté ce qui fait au premier abord la particularité de l'esprit tel que chacun de nous en a l'expérience (§ 3-4), il faudra se demander dans quelle mesure la connaissance de l'esprit telle qu'elle peut nous être scientifiquement présentée n'est pas simplement métaphorique et non littérale (§ 5).

2. La stratégie de Dennett

Dennett est l'inventeur d'une théorie dans laquelle les concepts mentaux sont les éléments d'une stratégie de prédiction¹. Il distingue pour cela trois niveaux d'interprétation : le point de vue du physique qui permet de prédire le comportement d'un système à partir de sa constitution ou des lois de la physique, le point de vue du plan qui concerne le comportement d'un objet relativement à la fin pour laquelle il est construit tout en ignorant les détails de sa constitution, et le point de vue intentionnel. Pour comprendre et anticiper comment un être complexe va se comporter, nous pouvons supposer en lui des intentions, des volontés, des croyances, grâce à

¹ Dennett 1978 : « Systèmes intentionnels » et 1987 : « Les vrais croyants ».

l'agencement desquelles nous allons être capables de rendre compte de ses différents états successifs et de ses actions². Si je lâche le dernier livre de Dennett, celui-ci va tomber car il est pesant : il s'agit là du point de vue physique. Si je règle mon réveil à cinq heures du matin et qu'il fonctionne correctement, celui-ci va me réveiller à cinq heures : il s'agit là du point de vue du plan, pour lequel je n'ai pas besoin d'être horloger pour pouvoir prédire le comportement du réveil. Si le petit Dan court dans la savane, c'est parce qu'il croit que s'y trouve un éléphant vêtu d'un pyjama et qu'il veut s'assurer de la vérité de sa croyance : ici l'interprétation en jeu est la stratégie intentionnelle, car la constitution de Dan telle que nous pouvons la connaître ne permet pas de prévoir ce comportement, et il n'a, semble-t-il, pas été programmé pour chasser un éléphant coquet. Dennett montre que ce dernier niveau est applicable aussi bien aux machines qu'aux animaux et aux hommes. Le plus simple pour anticiper le coup d'un ordinateur joueur d'échec est de considérer qu'il *veut* me battre, qu'il *connaît* les règles du jeu et qu'il *croit* que le coup qu'il va jouer est le meilleur possible. Il n'y a alors pas de différence radicale de nature entre les différents êtres du monde. Seul change le degré de complexité.

Mais les concepts mentaux employés ne paraissent pas, à première vue, avoir à être aussi réels que les corps du monde physique, puisqu'ils sont les instruments théoriques de la perspective mise en place pour comprendre les comportements. Ils sont, de ce fait, susceptibles d'être définis par le rôle qu'ils tiennent dans la mise en œuvre de cette stratégie : telle croyance sera la "cause" du comportement du chien, de l'ordinateur, de Dan, mais ne se trouvera nulle part. Chaque état de conscience, chaque concept mental, reste, quand il est considéré tout seul, invisible. Dans la mesure où toute définition peut être conçue comme une élimination³, ces concepts, qui sont des éléments spéculatifs, ne semblent finalement pas désigner quoi que ce soit de réel et ne sont peut-être que des fictions interprétatives, réductibles à leur fonction théorique. S'ils sont réductibles, cela signifie-t-il que la notion d'esprit est une erreur de catégorie⁴ ou qu'il n'existe pas vraiment d'esprits dans le monde ? On considère souvent que des termes ne désignant aucun objet palpable ou visible ne sont que des fictions, par opposition aux objets réels. Mais sur ce point, la position de Dennett est aussi instable qu'originale : accusé à la fois d'être un instrumentaliste et un réaliste, il revendique l'entre-deux⁵. Les croyances (archétypes des concepts mentaux, voire métonymie de l'esprit) ne sont pour lui ni des objets réels, ni de pures inventions fantaisistes.

L'esprit n'est pas une chose comme le corps dont il est l'esprit. Il n'y a d'abord que des comportements interprétables en termes de volontés et de croyances, selon la stratégie intentionnelle. L'intention de mouvoir un membre que je prête à quelqu'un, fût-ce moi-même, est d'abord un moyen

² Dennett 1987 (*ibid.*) et 1996.

³ Quine 1987 : « Définition » et « Esprit *versus* corps ».

⁴ Ryle 1949.

⁵ Dennett 1998 : « De l'existence des patterns ».

de comprendre et d'anticiper un certain comportement. Mais les intentions et l'ensemble du bagage psychologique sont aussi justifiés d'une autre façon : croire que l'on veut quelque chose, non seulement traduit, mais implique un comportement correspondant à cette croyance. La croyance devient alors l'une des causes du comportement qu'elle est susceptible d'expliquer. Si je crois que je veux quelque chose, et si cette croyance m'incite à agir d'une certaine manière, par exemple en m'efforçant d'acquérir la chose que je pense vouloir, alors cette croyance va devenir effective, même si elle est fausse ou si elle n'existe nulle part réellement dans le monde matériel – ce qui n'a aucune importance.

Il se peut que ce que je crois soit faux (il n'existe pas d'éléphant en pyjama dans la savane), cela n'empêche pas que cette croyance puisse expliquer mon comportement.

(1) Celui que j'ai vu jeter son chewing-gum est un cochon insouciant.

fonctionne aussi bien comme croyance que :

(2) Il se trouve quelque part dans la savane un éléphant vêtu d'un pyjama.

Il se peut aussi que ma croyance ne s'exerce pas sur un objet précis ou sur le bon objet, voire qu'elle soit incompatible avec d'autres croyances au moment où je la formule. Je peux croire que la personne qui a jeté son chewing-gum est un cochon insouciant sans savoir que mon meilleur ami, que je crois être un exemple de vertu, est ce cochon insouciant. Il n'existe en fait aucune différence entre des croyances portant bien sur un objet dont elles décrivent une propriété (vraie ou fausse) et des croyances sans objet réel (les éléphants coquets) ou déterminé. La distinction entre des croyances dites *de re* (concernant une chose du monde et ses propriétés) et *de dicto* (verbale ou ne désignant aucun objet individuel du monde) est, selon Dennett, inopérante⁶. Elle ne vaut pas, car nous n'avons une représentation des choses du monde qui sont visées dans la formulation de nos croyances que par un faisceau de descriptions, qui n'est rien moins que verbal, et tout à fait révisable. Une chose désignée dans une proposition est toujours dite, même si cette chose existe bien dans le monde.

La différence entre les croyances *de re* et *de dicto* ne change donc rien. Une croyance n'a que l'importance que le croyant-interprète lui accorde⁷, et je suis moi aussi témoin et interprète de mon propre comportement. Si je crois qu'il existe un éléphant vêtu d'un pyjama et que je veux le vérifier, alors ma croyance et ma volonté sont des antécédents réels de ma course à travers la savane. Elles ne sont ainsi pas qu'une manière de

⁶ Dennett 1978 : « How to Change Your Mind » et Dennett 1987 : « Au-delà de la croyance ».

⁷ Dennett 1987 : « Structures réelles, faits profonds et questions vides » et 1998 (*ibid.*).

comprendre ma course. Mais la faculté des croyances à occasionner un comportement ne permet pas de les traiter comme des choses qui existeraient en tant que telles, et que l'on pourrait distinguer par leur contenu, comme il est possible de distinguer (1) et (2). Ce qui importe, c'est seulement la fonction que les croyances et les autres états mentaux occupent. Et cette fonction, en tant que partie d'un processus cognitif ou réflexif, est réelle.

De plus, Dennett ne pense pas qu'il existe dans le cerveau un « Intentionneur central » qui dirigerait le corps comme un spectre marionnettiste dans un centre de la décision qui serait le centre de la conscience. Il n'y a pas de fantôme dans la machine⁸, pas d'homoncule, rien qui puisse faire que je veuille mouvoir un membre. Cela signifie qu'il n'existe nulle part dans nos cerveaux quelque chose qui en soi serait un esprit, qui puisse juger, croire, décider, se souvenir ou se représenter quelque chose : il n'y a qu'une activité cérébrale, finalement assez proche de l'activité computationnelle d'un ordinateur. A quoi renvoient alors tous ces concepts mentaux dont nous usons dans la stratégie intentionnelle, et pourquoi celle-ci fonctionne-t-elle ?

3. La conscience est-elle expliquée ou perdue de vue ?

En soutenant l'importance des concepts mentaux dans la mise en place de l'interprétation de nos comportements, Dennett semble assurer leur légitimité. Toutefois, en adoptant une posture résolument matérialiste, il nie l'existence distincte de quoi que ce soit qui ressemble à un esprit, un ange, ou même une pensée qui serait une chose distincte. Si nous ne croyons pas aux ectoplasmes et autres apparitions spectrales, nous accorderons volontiers qu'il ne se trouve nulle part d'esprit susceptible de nous dire : « Je suis là » quand nous lui demandons : « Esprit, où es-tu ? » Pourtant, nous avons bien du mal à concevoir que notre pensée ne fasse pas partie du monde comme une chaise qui se trouve toujours quelque part. La théorie de Dennett semble avoir fait disparaître dans un double-fond tout ce qui pourrait réellement servir de support au mental.

D'ailleurs, si le mental n'est que l'effet de l'activité cérébrale, alors l'esprit lui-même ne semblera pas susceptible d'avoir une quelconque importance dans le monde physique. Comment ce qui n'a rien de commun avec les corps pourrait-il agir comme une cause sur l'ordre des choses extérieures ? Considérer que l'esprit est le produit du cerveau, comme la vapeur est le produit d'une locomotive à charbon, c'est considérer, selon une image aussi suggestive qu'écoulée, que la conscience est un épiphénomène de l'activité cérébrale. Nous retrouvons ici une difficulté classique du problème corps-esprit, mais qui ne vaut pas pour la perspective envisagée, car aucune vapeur ne s'échappe de la locomotive de Dennett. L'esprit lui-même n'est

⁸ Ryle 1949.

rien du tout. Il est pourtant ce grâce à quoi nous agissons⁹. Mais il est impossible de trouver un moment ou un lieu où serait prise une décision, ainsi qu'il s'efforce de le montrer en réfutant le centralisme¹⁰, rebaptisé « matérialisme cartésien » dans *La conscience expliquée*¹¹.

Le centralisme comprend plusieurs caractéristiques incompatibles. Tout d'abord, la conscience y est conçue comme un flux continu et semble caractérisée par une unité pouvant faire l'objet d'une narration : il y a alors une histoire de la conscience qui peut être racontée, et dont la chronologie correspond avec ce que la conscience se représente du monde extérieur. D'autre part, le centralisme suppose, comme son nom l'indique, un centre de la conscience où sont prises les décisions, lequel est essentiellement distinct des niveaux dits périphériques, où se situent les instances sensibles et la mémoire. Il s'agit en fait d'un ensemble qui propose au jugement et à la volonté un contenu informationnel mis à leur disposition. Cette représentation de la conscience tend vers une interprétation homonculaire, car tout s'y passe comme si il y avait, dans ce que Dennett appelle le Théâtre Cartésien, un spectateur contemplant le contenu informationnel et prenant des décisions aussi bien au niveau spéculatif (en ce qui concerne le jugement théorique), qu'au niveau pratique (en ce qui concerne les actions volontaires). Cette représentation de la conscience est, selon lui, incohérente, car l'unité de la conscience qu'elle paraît supposer est immédiatement remise en question par l'existence des systèmes périphériques qui concernent aussi l'activité mentale, même si leur rôle n'est que passif. Dennett, au contraire, propose de concevoir la conscience selon le modèle d'un système agissant en parallèle et non en série, où chaque instance périphérique est susceptible d'être à l'origine de la prise de décision qui ressortira de la computation, qu'elle soit d'ordre spéculatif ou d'ordre pratique. Cette théorie cherche en outre à dépasser le substantialisme supposé de la conscience, qui fait de celle-ci une entité réelle à part dans l'homme, dont elle dirigerait les pensées et le comportement. C'est ici que le centralisme se rapproche le plus d'un dualisme classique marquant la séparation de l'esprit et du corps, que Dennett rejette comme une absurdité.

« Il n'existe pas de "flux de la conscience" unique définitif, parce qu'il n'existe pas de Quartier Général central, pas de Théâtre Cartésien "où tout se rejoint" pour le bénéfice d'un Signifieur Central.

⁹ Dennett 2003 présente les conséquences morales de sa théorie : « La liberté humaine n'est pas une illusion [...]. La liberté humaine est réelle – aussi réelle que le langage, la musique, et l'argent – elle peut donc être étudiée objectivement d'un point de vue scientifique et dépourvu de non-sens. Mais comme le langage, la musique, l'argent, et les autres produits de la société, sa persistance est affectée par ce que l'on en croit » (p. 305). Comme l'esprit, la liberté existe réellement, mais seulement dans la mesure où l'on y croit.

¹⁰ Dennett 1969.

¹¹ Dennett 1991 : deuxième partie et 2003 : « Whenever the Spirit Moves You ».

Tout le monde est d'accord pour dire qu'il n'existe pas pareil point de vue unique dans le cerveau [...].

Au lieu d'un tel flux unique (aussi vaste soit-il), il existe de multiples canaux dans lesquels des circuits spécialisés essaient, dans des Pandémoniums parallèles, d'accomplir leurs diverses tâches, en créant des Versions Multiples au fur et à mesure qu'ils opèrent.¹² »

Les Versions Multiples sont autant de perspectives possibles sur un événement mental. Elles sont comme le fruit de la perception ou de la réflexion d'un homoncule se trouvant dans un Pandémonium, où il rivaliserait avec d'autres pour établir le jugement final sur telle ou telle question. Mais évidemment, ces homoncules ne sont qu'une image et aucun d'entre eux n'existe ni ne décide réellement de ce que sera ce jugement. Seule la personne dont nous considérons la conscience peut décider de quoi que ce soit, car elle seule a un esprit.

Les discriminations correspondant aux contenus particuliers de la conscience, qui apparaissent comme autant d'éléments de l'expérience, « produisent, au bout d'un moment, quelque chose *qui ressemble assez* à un flux narratif ou une séquence »¹³. Le point important ici est que la conscience semble continue, ce qui ne veut pas dire qu'elle l'est vraiment. Dennett souligne que les contenus de conscience se trouvent divisés en plusieurs endroits du cerveau, et ne se réunissent nulle part pour constituer un récit unifié et définitif de l'expérience vécue. C'est pourquoi il n'y a pas de Théâtre Cartésien de la conscience « où tout se rejoint ». Les informations reçues par le cerveau ne touchent en fait que les zones concernées par le type de données sous la forme desquelles elles lui parviennent : sous forme d'informations auditives, visuelles, olfactives, sous la forme de souvenirs, d'habitudes de raisonnement, etc. Or, si les différentes zones du cerveau ne sont pas totalement indépendantes, aucune ne prend le pas sur les autres, aucune ne coordonne l'ensemble de la vie mentale, comme un chef d'orchestre accorderait autant d'instruments qu'il y a d'instances conscientes périphériques. La vie consciente est le fruit d'une reconstitution après coup, qui donne l'impression d'un écoulement de l'expérience vécue, mais qui n'est rien d'autre que l'image, en retard de quelques fractions de secondes, de ce qui a effectivement été perçu en tant que contenu informationnel.

En racontant une expérience, nous reconstruisons celle-ci dans une continuité spatio-temporelle qui n'est pas nécessairement la sienne. Dennett soutient cette thèse en tenant compte des expériences conduites sur le phénomène phi par Kolers et von Grünau. « Si on allume successivement et rapidement deux ou plusieurs points séparés au plus par quatre degrés d'angle visuel, on aura l'impression qu'un seul point va et vient. » Quand les points sont de couleurs différentes (en l'occurrence rouge et vert) « les sujets

¹² Dennett 1991, p. 322-323.

¹³ Dennett 1991, p. 148.

rapportent qu'ils voient la couleur du point en mouvement passer du rouge au vert au milieu de sa trajectoire »¹⁴. Il y a bien ici une différence entre la réalité objective et ce qui apparaît dans la conscience. Ce test souligne notre tendance courante à considérer les contenus de conscience comme constituant un flux, même lorsqu'il n'en est en fait rien : au niveau informationnel, la conscience (en fait les récepteurs sensoriels) ne reçoit durant ces expériences aucune information continue. L'impression de continuité homogène liée aux vécus de conscience n'est que l'effet de l'effort narratif dont chacun témoigne dès qu'il cherche à rendre compte de son expérience individuelle.

Je croise une femme dans la rue. Puis je me rappelle cette femme, mais dans mon souvenir, celle-ci porte des lunettes alors qu'en fait elle n'en avait pas. Comment cela est-il possible ? Comment nos souvenirs peuvent-ils ainsi nous tromper ? Dennett estime que deux moyens pourraient expliquer cette inadéquation avec la réalité objective, c'est-à-dire avec le contenu informationnel parvenu aux sens. Ces deux méthodes sont les révisions "staliniennes" et "orwelliennes". Leur principe est assez simple. Dans la révision orwellienne, le contenu informationnel reçu est transformé après l'arrivée dans le centre de la conscience en ajoutant aux séquences mémorielles des éléments qui leur sont étrangers. Quand le souvenir sera sollicité, il aura été modifié. Dans la révision stalinienne, les éléments informationnels arrivent au centre de la conscience, non pas tels qu'ils ont été reçus par les sens, mais déjà corrigés ou parasités par des éléments extérieurs. Selon Dennett, rien ne permet de choisir de façon définitive entre un type de révision ou un autre. Or, ces deux sortes de révisions sont incompatibles entre elles : les fausses informations sont fausses soit parce qu'elles sont mal conservées ou qu'elles sont transformées, soit parce qu'elles n'ont jamais été bien reçues. Ceci indique que la conscience ne doit pas être conçue selon le modèle d'un enregistrement d'informations conservé dans un centre mémoriel susceptible d'être sollicité à tout moment, comme nous ressortons un vieux dossier, ou nous projetons un vieux film.

Le temps de l'expérience n'est pas le même que le temps objectif, lequel est davantage le temps réel que le temps "vécu"¹⁵, puisque ce qui fait notre expérience personnelle et qui est considéré comme contenu de conscience, est toujours déjà raconté, revu et corrigé. Tout semble se passer comme si nous pouvions projeter dans notre esprit le film de nos souvenirs, ce qui marquerait la distance entre un spectateur des souvenirs et ce film lui-même (c'est-à-dire les contenus informationnels stockés dans la mémoire). Mais cette image, centraliste ou "cartésienne", suppose toute une organisation modulaire hiérarchisée permettant de passer des données brutes à leur considération et au jugement, ce que refuse Dennett. Pour lui, au contraire, « voir, c'est croire », autrement dit, il n'y a pas d'instances diverses mises en série pour passer de la perception sensorielle au jugement à propos de cette perception. Chaque perception est en même temps le

¹⁴ Dennett 1991, successivement p. 149 et p. 156.

¹⁵ Dennett 1991, p. 175.

résultat du jugement sur la perception, et il n'y a pas de différence entre voir (percevoir une chose) et voir (la reconnaître comme étant telle ou telle chose). L'intérêt de sa théorie est de proposer un modèle alternatif selon lequel l'unité de la conscience ne provient pas tant de ses propriétés *intrinsèques* que de l'unité d'orientation du discours que l'on peut tenir à son propos. C'est pourquoi il peut parler du Moi (*the self*, qui désigne aussi la conscience) comme d'un Centre de Gravité Narrative¹⁶.

« Comme je ne me lasse jamais de le répéter, tout le travail exécuté par l'homoncule imaginaire dans le Théâtre Cartésien doit être divisé et distribué dans l'espace et le temps dans le cerveau. Il est temps de répéter encore une fois ma devise : si tu te fais tout petit, tu peux tout externaliser virtuellement.¹⁷ »

Externaliser virtuellement, c'est rendre compte de l'activité mentale par une analyse sans pouvoir situer un lieu ou un moment précis où agirait notre conscience et où se trouverait notre Moi. En décrivant par exemple une prise de décision, nous pouvons chercher à expliquer un processus cognitif dans lequel la décision est toujours sur le point d'être prise ou déjà prise. Même en affinant l'analyse à l'extrême, au niveau de l'axone des dendrites neuronales, où le processus physique relève de la mécanique quantique et, de ce fait, d'un modèle probabiliste, il est impossible de dire *quand* et *où* j'agis de façon libre. Je suis donc toujours ailleurs, hors de ce centre. Pour Dennett, le Moi est l'ensemble de la structure, la conscience est le processus en entier. Sa devise signifie qu'il ne saurait y avoir en notre for intérieur quelque chose qui serait notre moi profond, comme une perle secrète qui constituerait notre nature individuelle et pourrait être le siège de nos pensées. Ce n'est donc pas comme un point réel situé dans le temps et dans l'espace que nous pourrions déterminer l'esprit. Comment la théorie de Dennett va-t-elle alors permettre d'asseoir le statut des états mentaux qu'elle mobilise ? Avant de répondre à cette question, il faut d'abord voir ce qui pour lui ne définit pas le mental, et n'a aucun rôle à jouer dans une explication scientifique de la conscience : le contenu qualitatif.

4. Ce que la conscience n'est pas

Dans *Le mystère de la conscience*, Searle note que Dennett « attaque le théâtre cartésien, non pas parce qu'à son avis, on peut trouver des états subjectifs partout dans le cerveau, mais plutôt parce qu'il ne pense pas qu'il existe quoi que ce soit qui ressemble à des états subjectifs »¹⁸. Autrement dit, la stratégie de Dennett, sous couvert de chasse aux homoncules, démons et autres fantômes dans la machine, dissimulerait un behaviorisme logique

¹⁶ Dennett 1991 : « La réalité des Moi ».

¹⁷ Dennett 2003, p. 238.

¹⁸ Searle 1997, p. 108.

hérité de Ryle dont le but serait d'éliminer la conscience constituée d'états qualitatifs, dont son "explication" ne saurait rendre compte, et qu'elle nierait simplement au nom de sa démarche scientifique à la troisième personne (l'« hétérophénoménologie »). Dans *La conscience expliquée*, Dennett note en effet : « Je ne puis prouver qu'aucune conscience de ce genre n'existe pas. Je ne puis également prouver que les gremlins n'existent pas. Le mieux que je puisse faire c'est de montrer qu'il n'y a aucune raison sérieuse d'y croire »¹⁹.

La conscience en question est celle qui serait déterminée par son contenu. Or, comme l'ont indiqué les considérations sur les croyances, pour Dennett, ce qui compte, ce n'est pas le contenu, mais la fonction. Il en va ici de même pour les perceptions (qui elles-mêmes engendrent des croyances à propos de ce qui est perçu). Il refuse d'accorder l'existence aux *qualia*, les qualités sensibles subjectives de l'expérience individuelle. Mais comment dire que ce qui nous apparaît n'existe pas, puisque nous le percevons ?

« Les philosophes ont adopté différents noms pour les choses qui sont dans le spectateur (ou propriétés du spectateur) censées accorder droit de cité aux couleurs et au reste des propriétés bannies du monde "extérieur" par les conquêtes de la physique : "sensations brutes", "perceptions sensibles", "qualités phénoménales", "propriétés intrinsèques des expériences conscientes", "contenus qualitatifs des états mentaux", et bien sûr, "qualia" [...] Je refuse l'existence de telles propriétés. Mais (et revoilà ce fameux thème) j'accorde bien volontiers qu'il existe apparemment des qualia.²⁰ »

Exister, c'est non seulement apparaître, mais aussi avoir un faisceau de propriétés susceptibles de permettre une identification. Or, les *qualia*, parce qu'ils sont toujours perçus à la première personne et à un certain moment, ne peuvent pas être identifiés objectivement. Ils sont ce qui apparaît à celui qui a une conscience, mais sont, eux, épiphénoménaux, car ils n'ont aucune fonction causale : ils sont l'aspect privé (individuel) d'un processus publiquement descriptible sans cet aspect privé. Ils ne permettent pas d'expliquer quoi que ce soit et sont donc superflus. Dennett considère qu'une conscience épiphénoménale est comme une conscience niée. Ainsi, pour lui, caractériser la conscience par des contenus qualitatifs, c'est fournir à l'éliminativisme le meilleur moyen de la supprimer.

D'autre part, les *qualia* ne sont même pas perçus réellement tels qu'ils sont au niveau qualitatif par celui qui les perçoit, car ceux-ci ne persistent pas. Le goût de la bière est un goût qui s'acquiert. La première gorgée n'a certainement pas la même saveur que la bière bue par un buveur expérimenté. Pourtant il peut tout à fait s'agir de la même bière, de même marque et de fabrication identique. Donc, la première et la dernière gorgée de bière devraient avoir les mêmes propriétés qualitatives, si celles-ci

¹⁹ Dennett 1991, p. 502-503.

²⁰ Dennett 1991, p. 461.

existaient bel et bien. Pourtant la première est mauvaise, et la dernière est bonne²¹. De même, les couleurs, qui paraissent n'exister nulle part si ce n'est dans l'œil du spectateur (ou son esprit), ne sont qu'une apparence liée au fonctionnement de notre vision, qui ne les discrimine qu'autant que cela lui sert. Elles dépendent donc des services qu'elles nous rendent, et sont subordonnées à leur fonction. Elles n'existent pas par elles-mêmes²².

Mais expliquer quel processus peut engendrer la conscience est une chose, pouvoir dire ce qu'elle est *elle-même* (ce que nous percevons ou ressentons lorsque nous sommes conscients, et pourquoi nous le percevons) est peut-être autre chose. Il existerait alors deux problèmes de la conscience : l'un concernant sa formation ou sa constitution, et l'autre rendant compte de l'expérience vécue en tant que telle. Il est ici possible de penser que la science physicaliste sera toujours incapable de rendre compte de ce qui fait le vécu de la conscience, étant donné que ce vécu est toujours une affaire privée. Mais comme on s'en doutera, Dennett refuse cette critique, et ne considère pas que ce qui apparaît à la première personne intéresse une philosophie ou une théorie de l'esprit, justement parce que ces données subjectives ne sont pas des éléments persistants ni objectifs (publics). Et contrairement à ceux qui estiment que ce point marque l'impossibilité d'expliquer réellement la conscience, il considère que prendre en compte les *qualia* et les autres données subjectives de l'expérience consciente vécue, c'est prendre le risque de se perdre dans des considérations aussi vaines que mystérieuses²³ et renoncer au devoir du théoricien.

Au point de vue objectif (à la troisième personne), le vécu de conscience subjectif et l'impression qu'il laisse ne changent rien du tout. Donc ils n'ont aucune importance. Considérons un individu réel, David, et son jumeau zombie (identique en tous points à David au niveau physiologique, mais dépourvu de cette conscience que David considère comme un fait irréductible et dont il a une perception intime évidente). Précisons que le zombie est capable de rendre compte de ses états internes en "croyant" que ce qu'il rapporte existe bien. Il peut témoigner de certains contenus qu'il distingue selon leurs qualités, mais, à la différence du contenu des états internes de David, ceux-ci ne sont pas vraiment des états conscients, puisqu'il s'agit d'un zombie.

Selon les défenseurs d'une conscience constituée par ce dont nous avons une expérience intime directe, et que la théorie de Dennett ne considère que comme une actualisation particulière du processus cognitif, il y a une différence entre David et le zombie. Si David fait l'expérience directe de sa conscience, il peut considérer qu'il manque quelque chose au zombie, qui, lui, ne fait pas la même expérience que David. Le problème, souligné par Dennett²⁴, est que pour nous, qui ne sommes ni David ni son jumeau zombie, le témoignage de David ne nous apporte rien de plus que

²¹ Dennett 1991, p. 490-491.

²² Dennett 1991, p. 464-474 et 1998 : « Instead of Qualia ».

²³ Dennett 1999.

²⁴ Dennett 2001.

celui du zombie. Ce dernier “croit” aussi qu’il est conscient, et qu’il a des états mentaux dont les contenus se distinguent les uns des autres. L’histoire nous dit que le zombie a tort, et que David a raison. Mais qu’est-ce que cela changerait pour nous si nous découvriions qu’en fait David est un zombie et que le pseudo-zombie est conscient ? – Rien. Être conscient, c’est donc avoir un cerveau qui fonctionne comme celui du zombie de l’histoire, ou comme David.

5. La théorie fonctionnaliste de l’esprit n’est-elle pas analogique ?

L’objet de la théorie de Dennett est donc bien la conscience elle-même. Mais la présentation qui en est faite n’a-t-elle pas un statut ambigu, puisqu’elle rend compte de ce qui peut engendrer une conscience en général, sans que jamais nous ne puissions retrouver en nous-même ce qui se trouve sous l’organisation décrite ? A la fin de *La conscience expliquée*, nous pouvons lire : « Je n’ai pas remplacé une théorie métaphorique, le Théâtre Cartésien, par une théorie *non* métaphorique (“littérale”, “scientifique”).²⁵ » Il semblerait donc qu’il y ait bien une distance ou une analogie entre le fruit du fonctionnalisme et la conscience elle-même, non seulement parce que le point de vue théorique est général et ne prend pas en compte telle expérience privée plutôt que telle autre, mais aussi parce que les outils théoriques utilisés ne sont que des images de ce qui se passe en nous.

Mais ceci n’est qu’une apparence. En effet, si les Versions Multiples des états mentaux, censées remplacer le flux de la conscience et le centralisme, sont bien des métaphores substituées à d’autres métaphores, ce qui fait le fonds du fonctionnalisme, l’idée selon laquelle la conscience est une organisation ou une structure, elle, est littérale. « Quiconque ou quoi que ce soit qui dispose d’une machine virtuelle de ce type comme système de contrôle est conscient au sens le plus plein du terme ; il est conscient *parce qu’il a une telle machine virtuelle.*²⁶ » Peu importe que l’être en question soit un homme, un zombie ou un robot. Ce qui est réel, ce n’est pas le “contenu” de la conscience, c’est son fonctionnement, sa structure²⁷.

Dans « The Zombic Hunch », Dennett affirme la réussite du programme fonctionnaliste :

« Serait-il, d’une façon ou d’une autre, possible que la conscience émerge de l’organisation de toutes les parties qui fonctionnent les unes avec les autres ? Et si c’était le cas, pourquoi ne pourrions-nous pas espérer la comprendre, une fois développés les bons concepts ? Telle est la voie qui a été explorée de façon enthousiaste et fructueuse durant le dernier

²⁵ Dennett 1991, p. 565.

²⁶ Dennett 1991, p. 351.

²⁷ Dennett 1998 : « De l’existence des patterns » et 1999.

quart de siècle sous la double bannière des sciences cognitives et du fonctionnalisme – l'extrapolation du naturalisme mécaniste du corps vers l'esprit.²⁸ »

Le fonctionnalisme est la bonne façon de concevoir la conscience car ce qu'il décrit est *littéralement* exact. En effet, notre corps lui-même est bien une organisation d'homoncules stupides (nos cellules), agissant de concert sans jamais savoir ce que nous savons, ni être conscients de quoi que ce soit. Seuls nous-mêmes sommes conscients. Mais il ne se trouve nulle part *en nous* un lieu où se trouverait notre conscience. Il n'existe nulle part de *res cogitans* s'ajoutant à l'activité de notre cerveau et de notre corps²⁹. « Nous sommes tous *constitués* de robots sans esprits et de rien d'autre, rien qui ne soit pas physique, aucun ingrédient qui ne soit robotique.³⁰ » C'est le fonctionnement de notre corps-machine et de son cerveau qui produit la conscience, ce qu'une perspective évolutionniste justifie pleinement³¹. Mais il ne la produit pas comme quelque chose d'autre, comme un feu produit de la fumée. Chaque conscience a l'aspect qualitatif que prend la façon particulière dont fonctionne cette organisation chez un individu particulier. C'est pour cette raison qu'il y a des apparences qualitatives, mais qu'elles ne sont *que* des apparences. Ce qui fait la conscience même ne se trouve pas là, mais dans l'explication de la structure réelle qu'est notre esprit. Nos cellules elles-mêmes fonctionnent conformément à leur constitution, selon une perspective qui rappelle le point de vue physique. Nos organes, eux, sont un peu plus complexes et fonctionnent comme des petites machines ayant pour finalité la survie du corps dans lequel le cœur est une pompe. Ce deuxième niveau correspond au point de vue du plan. Mais l'ensemble du système complexe que nous constituons peut être interprété selon le point de vue intentionnel. Ainsi, selon l'ampleur de ce que nous considérons, nous retrouvons les perspectives de prédiction du comportement présentées plus haut.

« En supposant qu'on possède un minimum de connaissance scientifique, personne ne peut aujourd'hui douter sérieusement que *nous sommes les descendants directs de robots auto-réplicatifs*. Nous sommes des mammifères, et tous les mammifères descendent d'êtres reptiliens dont les ancêtres étaient des poissons, dont les ancêtres étaient des créatures semblables à des vers, qui descendaient de créatures multicellulaires qui ont vécu il y a plusieurs centaines de millions d'années, descendant à leur tour de créatures

²⁸ Dennett 2000.

²⁹ Dennett 2003, p. 254.

³⁰ Dennett 2000.

³¹ Cf. Dennett 1995 : « Perdre la tête pour Darwin » et 1996.

unicellulaires descendant de macromolécules répliquatives ayant vécu il y a environ trois milliards d'années. [...]

Pour employer une image, votre arrière-arrière... grand-mère *était* un robot. Mais ces robots macromoléculaires ne sont pas seulement vos ancêtres, ils sont la matière dont vous êtes faits [...] – votre corps tout entier (y compris votre cerveau, bien entendu), à tous les niveaux de l'analyse, depuis la molécule, est composé de mécanismes qui accomplissent aveuglément un travail magnifique, conçu avec art.³² »

Au plus petit niveau, nous n'avons aucune intention. Aucune de nos cellules ne veut rien, ni ne sait rien. Nos organes internes ignorent tout de qui nous sommes. Pourtant, de la complexité de l'organisme considéré au niveau global émerge une unité narrative : celle dont on dit qu'elle est cet organisme, composé de cellules et d'organes, qui agit de telle manière, désire ceci et craint cela. L'image du Centre de Gravité Narrative traduit à la fois l'absence d'un trésor interne secret qui correspondrait à notre essence, et le fait que notre identité ne provient que du croisement des faisceaux narratifs grâce auxquels nous disons qui nous sommes, qu'il s'agisse d'un tiers ou de nous-mêmes³³.

De ce fait, il n'y a pas pour Dennett de propriétés inexplicables de l'esprit. Celui-ci peut toujours être reconduit, en changeant de point de vue, à une perspective mécanique. Mais, dans cette perspective, la conscience se dissout, car un mécanisme n'est conscient de rien : seule une personne l'est. Ainsi, l'esprit est à la fois réductible et irréductible au corps. Il l'est, parce que le fonctionnement et l'évolution de la machine que nous sommes permettent de l'expliquer. Il ne l'est pas, parce que le point de vue physique ou celui du plan, qui concernent la constitution des corps ou leur finalité aveugle (ce pour quoi ils ont été construits), ne permettent pas d'envisager la conscience elle-même.

Le programme théorique de Dennett permet donc de rendre littéralement compte de la conscience, c'est-à-dire de ce qui se passe dans nos cerveaux, et qui fait que nous faisons telle ou telle expérience. Cette perspective le situe nettement dans un courant "post-métaphysique", où le substantialisme est banni. On cherchera donc en vain dans ses écrits des considérations sur l'essence du corps ou celle de l'esprit, ou de quoi étayer une théorie de l'individuation radicale des êtres et des choses. Son propos n'est pas là. En reconstruisant le Moi de l'extérieur, par ce que nous pouvons en dire, Dennett retrouve une unité narrative, fictive ou théorique, mais effective, de la personne et de la conscience. Le problème corps-esprit est alors bien dissout. Mais il est inutile d'espérer trouver chez lui une mise au point sur la différence entre ce que nous concevons au point de vue théorique, que nous construisons spéculativement, et ce qui nous affecte.

³² Dennett 1996, p. 40.

³³ Dennett 1991 : « L'insoutenable légèreté de l'être ».

Nous ne sentons pas une organisation en tant qu'organisation, ni la totalité de nos cellules, mais notre corps. Il n'est cependant pas légitime d'attendre de son point de vue scientifique une justification de ce qui échappe à la reconstitution théorique des événements. Cette perspective prend en compte notre corps et notre conscience comme nous envisageons au quotidien des choses extérieures qui ne nous touchent pas : par une description de propriétés et un ensemble de relations. Bien entendu, il peut nous sembler difficilement acceptable d'imaginer qu'une structure puisse penser. Ce sont les hommes qui pensent, pas les objets théoriques. Mais c'est le niveau d'interprétation qui n'est alors pas bon. Il suffit de quitter le point de vue du plan pour la perspective intentionnelle. Nous sommes tous, chacun de nous, des structures réelles. Serait-il légitime d'attendre d'un discours théorique qui se veut scientifique une prise en compte de ce qui ne persiste pas et n'est pas publiquement observable ? Bien sûr, Dennett admet qu'il y a des apparences privées des phénomènes mentaux. Mais leur contenu même est incapable de convenir à un quelconque protocole scientifique. Ce qui importe est ce qui est effectif (non épiphénoménal) et descriptible. Tout ce qui rend compte de ces critères est explicable. Il en va de même pour nos esprits et nous-mêmes. Ce que je suis, c'est ce que je fais. Il n'y a rien à dire de moi si je ne peux pas être décrit, si mon histoire ne peut pas être racontée.

Références

Quand une traduction publiée existe, les ouvrages sont, par commodité, cités d'après l'édition française.

Daniel C. Dennett :

1969 : *Content and Consciousness*, Londres, Routledge & Kegan Paul.

1978 : *Brainstorms*, Montgomery/Cambridge, Mass., MIT/Bradford, 1984. Tr. fr. dans *Philosophie*, 1, 1984 de « Intentional Systems » (« systèmes intentionnels »).

1987 : *The Intentional Stance*, Cambridge, Mass., MIT/Bradford. *La stratégie de l'interprète. Le sens commun et l'univers quotidien*, Paris, Gallimard, 1990.

1991 : *Consciousness Explained*, Boston, Little, Brown. *La conscience expliquée*, Paris, Odile Jacob, 1993.

1995 : *Darwin's Dangerous Idea. Evolution and the Meaning of Life*, New York, Simon and Schuster. *Darwin est-il dangereux ?*, Paris, Odile Jacob, 2000.

1996 : *Kinds of Minds. The Origins of Consciousness*, New York, Brockman, Inc. *La diversité des esprits. Une approche de la conscience*, Paris, Hachette, 1998.

1998 : *Brainchildren. Essays on Designing Minds*, Cambridge, Mass., MIT/Bradford, 1998. Tr. fr. dans D. Fiset/P. Poirier (éd.), *Philosophie de*

l'esprit. I. Psychologie du sens commun et sciences de l'esprit, Paris, Vrin, 2002 de « Real Patterns » (« De l'existence des patterns »).

1999 : « The Zombic Hunch : An Extinction of an Intuition ? », Final draft for Royal Institute of Philosophy – Millennial Lecture, novembre 1999.

2000 : « Are We Explaining Consciousness Yet ? » Final draft, août 2000.

2001 : « The Fantasy of First-Person Science », 3rd version, mars 2001.

2003 : *Freedom Evolves*, New York, Viking.

Les articles non repris dans les livres publiés de Dennett sont disponibles à partir de l'adresse suivante: <http://ase.tufts.edu/pubpage.htm>.

Gilbert Ryle 1949 : *The notion of Mind*, London, Hutchinson and Co., 1949. *La notion d'esprit. Pour une critique des concepts mentaux*, Paris, Payot, 1978. [Ryle a dirigé les premières recherches de Dennett. Son influence est capitale.]

Paul A. Kolers/Michael von Grünau 1976 : « Shape and Color in Apparent Motion », in *Vision Research*, 16, p. 329-335.

Willard V. O. Quine 1987, *Quiddities. An Intermittently Philosophical Dictionary*, Cambridge, The Belknap Press/Harvard University Press. *Quiddités. Dictionnaire philosophique par intermittence*, Paris, Le Seuil, 1992. [L'autre "maître et ami" de Dennett (cf. la dédicace de Dennett 1995).]

Pascal Engel 1994, *Introduction à la philosophie de l'esprit*, Paris, La Découverte, chapitre 8 : « La conscience n'est-elle qu'un mythe ? ». [Pour une autre perspective sur les thèses de Dennett, par un fin interprète qui est aussi son principal traducteur en français.]

David J. Chalmers 1996 : *The Conscious Mind : In search of a Fundamental Theory*, Oxford/New York, Oxford University Press. [Le David de l'histoire de zombie, qui distingue deux problèmes corps-esprit : le fonctionnement (le psychologique) et l'aspect qualitatif (le phénoménal). Le gros problème est d'expliquer le phénoménal : pourquoi cela fait-il cet effet-là d'être conscient de telle ou telle chose ? Dennett se contenterait de proposer une solution au petit problème du fonctionnement psychologique.]

John R. Searle 1997 : *The Mystery of Consciousness*, New York, New York Review of Books. *Le mystère de la conscience*, Paris, Odile Jacob, 1999. [Une discussion de différentes théories de l'esprits, avec un échange avec Dennett où l'humour dissimule une incompréhension peut-être plus grande que la mauvaise foi des auteurs. Searle discute aussi Chalmers 1996.]