



Ni fantôme ni zombie : L'émergence de la conscience subjective dans le flux des expériences

Par BRUNO LECLERCQ
Université de Liège

Résumé Sous le double slogan de la « phénoménalité intentionnelle » et l'« intentionalité phénoménale », des théories récentes de l'esprit s'efforcent de renouer avec certaines thèses Brentaniennes et/ou husserliennes pour réinsister sur l'ancrage irréductible des attitudes intentionnelles dans la dimension phénoménale, « qualitative » et subjectivement ressentie de la conscience. À l'encontre de ce programme, je rappelle ici certaines objections classiques formulées, sur le terrain de la philosophie du langage, par le behaviorisme logique, lequel invite plutôt à une hétérophénoménologie.

Dans une contribution au volume collectif sur « La nature vivante » paru dans le présent *Bulletin*, j'avais abordé la question de la naturalisation de propriétés émergentes de la matière telles que la vie ou la conscience, mais aussi la question de la spécificité anthropologique — de l'éventuelle différence de l'homme à l'égard de l'animal, du vivant ou de la machine —, à travers le modèle d'une hétérophénoménologie¹. Je m'y étais alors opposé à l'idée que la phénoménologie soit essentiellement une théorie de la conscience phénoménale au sens de l'expérience subjective, de l'« effet que ça fait » (*what is it like*) ou des *qualia* et j'avais au contraire soutenu la thèse que les phénomènes intentionnels qui constituent le cœur de l'investigation phénoménologique peuvent, dans une très large mesure et même pour

¹ Bruno Leclercq, « *What is it like to be a bat ? La phénoménologie à la troisième personne de Wittgenstein à Dennett* », dans *Bulletin d'Analyse Phénoménologique*, 2010 (vol. 6, n°2), p. 278-312 (<http://popups.ulg.ac.be/1782-2041/index.php?id=365>).

l'essentiel, être étudiés « à la troisième personne ». Or, cette thèse est assez précisément celle que, depuis quelques années, contestent farouchement les partisans de la dimension phénoménale de la conscience, dont les arguments ont été présentés et discutés à Liège lors d'un séminaire intitulé « L'Expérience subjective: Approches phénoménologiques de la conscience phénoménale »¹.

En ouverture de ce séminaire, le présent texte se donnait pour délicate mission de rappeler les principaux contours d'une position qui devait faire l'objet de critiques tout au long de la semaine suivante et, en sens inverse, de formuler un certain nombre d'objections (classiques) à l'encontre de la conception résolument subjectiviste de la conscience, objections auxquelles les intervenants du séminaire devaient ensuite s'efforcer de répondre. À vrai dire, l'ambition de cette intervention était aussi de faire plus que cela, et notamment de faire un peu bouger les « lignes de front » entre les deux camps en interrogeant, au nom des arguments mêmes des hétérophénoménologues, la thèse, soutenue par certains d'entre eux, selon laquelle la conscience phénoménale (l'expérience subjective, l'effet que ça fait, les *qualia*) ne serait qu'un « épiphénomène », qui ne serait essentiel ni sur le plan descriptif ni sur le plan explicatif à l'intentionnalité de l'esprit. C'est là ce qui justifie sa publication indépendante.

Insistons d'emblée sur le fait que les auteurs sur lesquels je m'appuierai (principalement Ludwig Wittgenstein, Gilbert Ryle et Daniel Dennett) n'entendent pas minimiser le caractère subjectif des phénomènes de conscience au nom du fait que ceux-ci pourraient essentiellement recevoir une définition *neurophysiologique*. Pour les hétérophénoménologues dont il sera ici question, il s'agit bien plutôt d'affirmer que les phénomènes de conscience sont essentiellement caractérisés par les *attitudes comportementales* des organismes (et même de certains dispositifs purement mécaniques) à l'égard de leur environnement ; à cet égard, entendue comme « stratégie de l'interprète », l'hétérophénoménologie s'inscrit dans la continuité du « behaviorisme logique », lequel, j'y insisterai, est une doctrine qui relève de la philosophie du langage avant de concerner la philosophie de l'esprit.

Or, pour annoncer mon propos et justifier d'emblée le titre du présent texte, je voudrais ici tout à la fois relayer le reproche que, avec Gilbert Ryle, les hétérophénoménologues adressent à Descartes (et tout aussi bien, sans doute, à Brentano et Husserl) de concevoir la conscience comme un « fan-

¹ Unité de recherche « Phénoménologies » : <http://www.pheno.ulg.ac.be/colloques/201204-conscience/fr/>.

tôme dans la machine »¹, et donner en sens inverse crédit à la réplique que, avec notamment David Chalmers², les partisans de la conscience phénoménale opposent au behaviorisme logique, à savoir de ne théoriser qu'un « zombie », c'est-à-dire un mécanisme qui développe des attitudes comportementales, mais ne ressent rien, une « machine sans fantôme ». Je tâcherai pour ma part de montrer que, si elle met gravement à mal l'idée d'une conscience phénoménale et d'un fantôme dans la machine, la théorie dennettienne de la conscience divisée et « multi-versions » (*multiple drafts*) condamne aussi le modèle du zombie en mettant sérieusement en question la notion même de « conscience épiphénoménale ».

Les arguments du behaviorisme logique contre le fantôme dans la machine

Commençons donc par rappeler brièvement les contours de la position hétérophénoménologique. Chez Wittgenstein et même partiellement chez Ryle³, c'est d'abord sur le terrain de la philosophie du langage que s'est développée la critique de la conception subjectiviste des phénomènes de conscience ou des phénomènes mentaux. Le modèle que dénonce Wittgenstein dans les *Recherches philosophiques*, mais aussi plus centralement encore dans ses *Remarques sur la philosophie de la psychologie*⁴, c'est celui d'un discours qui entendrait décrire les expériences internes (les sensations, mais aussi les émotions, les désirs, les espoirs, les convictions...) dans un langage phénoménologique dont les énoncés seraient rendus vrais ou faux par la présence effective ou non de ces expériences internes dans la conscience de celui à qui elles sont attribuées⁵. Dans une telle perspective, « J'ai mal », « Je vois du rouge », « J'entends un bourdonnement », « J'ai froid »,

¹ Gilbert Ryle, *The Concept of Mind*, London, Hutchinson, 1949, trad. fr. *La notion d'esprit*, Paris, Payot, 1978.

² David Chalmers, *The Conscious Mind : In Search of a Fundamental Theory*, Oxford University Press, 1996.

³ Cf. la problématique de l'« erreur de catégorie » dans *La Notion d'esprit*, *op. cit.* Cf. aussi Gilbert Ryle, « Adverbial verbs and verbs of thinking », in *On Thinking*, Oxford, Basil Blackwell, 1979, p. 17-31.

⁴ Ludwig Wittgenstein, *Recherches philosophiques*, trad. fr. Paris, Gallimard, 2004 ; *Remarques sur la philosophie de la psychologie, I et II*, trad. fr. Mauvezin, T.E.R., 1989.

⁵ Notamment énoncé dans le § 243 des *Recherches philosophiques*, ce modèle est longuement dénoncé par Wittgenstein dans les paragraphes avoisinants.

« J'ai faim », « J'ai peur », « Je déteste les choux de Bruxelles », « J'aime cette fille », « Je voudrais que ce soit réciproque », « Je crains que ce ne le soit pas », « Je crois qu'elle en aime un autre »... seraient des phrases qui énoncent des expériences essentiellement subjectives — c'est moi et moi seul qui les éprouve — et dont la valeur de vérité dépendrait de la réalisation ou non de ces expériences subjectives dans ma conscience. Ces énoncés constitueraient donc bien des jugements affirmatifs et descriptifs avec toutefois cette singularité que je suis mieux placé que vous pour savoir s'ils sont vrais, puisque j'ai directement accès à mes expériences subjectives tandis que vous ne pouvez que supputer leur présence dans ma conscience à travers leurs manifestations et conséquences dans mon comportement (en ce compris, bien sûr, ce que j'en dis). Puisque je suis mieux que vous dans les conditions de vérifier si les expériences subjectives dont il est question ont ou non lieu dans ma conscience, « je sais mieux que vous ce que je ressens et ce que je pense », « je sais mieux que vous si j'ai mal, si j'aime cette fille ou si je crois que Dieu existe »¹.

À ce modèle, Wittgenstein oppose l'objection que le jeu de langage phénoménologique — l'ensemble des règles d'usage des termes phénoménologiques — n'est pas privé mais bien intersubjectif. Si les critères de vérité de « J'ai mal », « J'aime cette fille » ou « Je crois que Dieu existe » — et donc les critères de légitimité d'usage des termes « douleur », « amour », « croyance » — étaient essentiellement privés, il ne serait plus possible à la communauté linguistique de me les enseigner ; faute d'avoir accès à mes expériences subjectives, mes « instructeurs » ne pourraient vérifier si j'utilise ces termes correctement². Qu'ils puissent m'enseigner les règles d'usage de ces termes, et parfois contester la vérité de mes énoncés psychologiques à la première personne — je dis que « Je n'ai pas mal » alors que je grimace et laisse perler des larmes sous les coups que m'inflige une brute dans la cour de récréation³ ; j'affirme que « Je n'aime pas cette fille » mais je fais manifestement tout pour me retrouver sans cesse en sa présence ; je dis que « Je ne crois pas en Dieu » mais je l'invoque et le prie d'intervenir dans la moindre situation contraire... — tout cela montre bien que les « phénomènes de conscience » dont il est ici question ont des critères essentiellement

¹ Cf. les §§ 290-297 des *Recherches philosophiques*, où Wittgenstein explore cette idée d'un savoir descriptif que chacun aurait de ses propres états mentaux.

² Cf. ce que dit Wittgenstein dans les §§ 257-261 des *Recherches philosophiques* à propos d'un éventuel langage privé.

³ Ludwig Wittgenstein, *Recherches philosophiques*, *op. cit.*, Partie II, sect. XI, p. 313-314.

intersubjectifs. Loin qu'ils doivent se contenter de formuler d'invérifiables hypothèses sur des phénomènes internes auxquels j'aurais, quant à moi, directement accès, les autres *savent* très souvent, parfois même mieux que moi, ce que je perçois, ce que je désire et ce que je crois¹.

Si, dit Wittgenstein, je ne pouvais jamais être corrigé par d'autres lorsque je m'attribue des croyances, des désirs, des perceptions (et même, sans doute, des sensations), si j'étais toujours seul à même de juger de la vérité de ces attributions, il n'y aurait tout simplement pas de jeu de langage *commun* — faute qu'il y ait des critères *intersubjectifs* d'usage de ses termes — et même tout simplement pas de jeu de langage — faute qu'il y ait de critères d'usage tout court ; car, si le seul critère reposait dans la manière dont les choses m'apparaissent au moment où je les j'éprouve, tout ce qui me paraîtrait correct serait correct, de sorte qu'il n'y aurait tout simplement plus de différence entre le correct et l'incorrect. C'est le fameux argument de Wittgenstein contre le langage privé : un langage dont les critères d'usage reposeraient essentiellement dans mes expériences subjectives manquerait d'authentiques *critères* d'usage². Et, de même, un savoir dont les critères de vérité reposeraient essentiellement dans mes expériences subjectives manquerait d'authentiques critères de vérité³. Loin d'en faire un savoir privilégié, la fameuse infaillibilité du « savoir » phénoménologique à la première personne est, pour Wittgenstein, ce qui précisément le condamne à ne pas être véritablement un savoir. Là où on ne peut pas douter, on ne peut pas non plus savoir. Si, en raison de l'immédiateté de mon accès à mes propres expériences subjectives, cela n'a pas de sens de dire que je ne sais pas si j'ai mal ou que je ne sais pas si je crois en Dieu, cela n'a pas non plus de sens de dire que je *sais* si j'ai mal ou que je *sais* si je crois en Dieu. Là où il n'y a pas possibilité d'erreur, on n'est plus dans l'ordre du savoir⁴.

Il y a là, on le voit, des objections majeures contre la phénoménologie envisagée comme une science « à la première personne »⁵. Bien sûr, Wittgenstein ne nie-t-il pas qu'on puisse dire « J'ai mal ». Mais, contrairement à ce qu'il semble, cet énoncé n'est pas essentiellement

¹ *Ibid.*, §§ 246, 303.

² *Ibid.*, §§ 202-205. Sur cette question de la possibilité de suivre une règle en privé, voir notamment le commentaire de Saul Kripke dans *Wittgenstein on Rules and Private Language*, Cambridge, Harvard University Press, 1982. Sur l'impossibilité pour des images mentales de déterminer univoquement comment suivre la une règle, cf. aussi *Recherches philosophiques*, *op. cit.*, §§ 85, 199, 239, 337.

³ Ludwig Wittgenstein, *Recherches philosophiques*, §§ 265-266, 269.

⁴ *Ibid.*, §§ 246, 288-289, 404-405.

⁵ Cf. aussi Bruno Leclercq, « *What is it like to be a bat ?* », art. cit., p. 290-291

l'affirmation ou la description d'une expérience interne. Dans un premier temps, cette production verbale — « J'ai mal » — est elle-même une des expressions corporelles de la douleur ; ma communauté m'a progressivement incité à remplacer mes cris de douleur initiaux par cette production verbale¹. Mais, donc, gémi sous l'effet d'un certain nombre de décharges neuronales, « J'ai mal » n'est pas plus déclaratif et descriptif que ne l'était « Aïe ! ». En fait, tout ce qui a la forme linguistique d'une proposition — tout ce qui a « consonance propositionnelle » —, dit Wittgenstein avant Austin, n'est pas nécessairement une proposition au sens d'une phrase déclarative et descriptive dont la valeur de vérité dépend de la question de savoir si le monde est effectivement comme elle le décrit ; bien qu'apparemment déclaratives, certaines phrases sont en fait essentiellement *performatives* (« Tu auras ton argent demain », « Vous êtes mari et femme »...) ou *expressives* (« J'en ai marre », « J'ai mal »...) ou même les deux à la fois (« Je t'emmerde », « Je t'aime »...) ². Et, en ce sens, elles ne sont pas vraies ou fausses.

Certes, comme « Vous êtes mari et femme », « J'ai mal » a cependant aussi un usage plus proprement propositionnel (déclaratif et descriptif), qui prétend à une valeur de vérité. Mais, dit Wittgenstein, pour être d'authentiques « conditions », les conditions de vérité de cette proposition ne peuvent alors simplement reposer dans l'expérience subjective de celui qui l'énonce. Lorsque « J'ai mal » est propositionnel et prétend à une valeur de vérité, c'est qu'il s'inscrit dans le jeu de langage intersubjectif au même titre que « Tu as mal » et « Il a mal » ³. La chose est bien sûr plus claire encore pour « J'aime cette fille », « Je souhaite plus de justice sociale » ou « Je crois que Dieu existe ».

¹ Ludwig Wittgenstein, *Recherches philosophiques*, § 244.

² Tout le début des *Recherches philosophiques* est consacré à critiquer la théorie « augustinienne » du langage, selon laquelle tous les termes seraient des noms dont la signification est apprise par ostension de leur référent (cf. §§ 1, 4, 10, 11, 17) et toutes les propositions du langage ne serviraient qu'à affirmer (cf. §§ 21-24, 65, 112-114, 134-137). À l'encontre du *Tractatus logico-philosophicus* (§ 4.5), Wittgenstein fait désormais valoir la diversité des jeux de langage (*Recherches philosophiques*, §§ 2-3). À cet égard, il anticipe John Langshaw Austin et son *How to Do Things with Words : The William James Lectures delivered at Harvard University in 1955*, Oxford, Clarendon Press, 1962, trad. fr. *Quand dire, c'est faire*, Paris, Le Seuil, 1970.

³ Cf. Bruno Leclercq, « *What is it like to be a bat ?* », art. cit., p. 304. L'asymétrie entre première et troisième personnes que souligne Wittgenstein pour les expressifs sera également soulignée par Austin pour les performatifs.

Le « behaviorisme logique » est, on le voit, d'abord et avant tout une thèse de philosophie du langage : il ne s'agit pas d'emblée de dire (comme le fait le behaviorisme ontologique) qu'il n'y a rien d'autre que des comportements, qu'il n'y a rien dans la « boîte noire », rien de caché sous la surface des attitudes corporelles, ni même de soutenir (comme le behaviorisme méthodologique ou épistémologique) que, par méthode, on doit en science s'en tenir à l'étude de ce qui est intersubjectivement observable, à savoir le comportement ; mais il s'agit plutôt ici de dire que les termes mêmes du langage phénoménologique requièrent des critères d'usage intersubjectifs, et les énoncés mêmes du savoir phénoménologique des conditions de vérité intersubjectives. C'est en fait la *définition* même des *concepts* phénoménologiques qui est en jeu¹. Contrairement à ce qu'il semble à première vue, les concepts de douleur, de désir ou de croyance qui sont au centre de nos jeux de langage ne sont pas caractérisés par des expériences subjectives mais bien par des traits intersubjectivement observables. Le concept de douleur — qui correspond à l'usage linguistique du terme « douleur » — n'est pas essentiellement défini par des *qualia* au sens de qualités sensorielles subjectivement vécues qui pourraient trouver leurs causes dans des sévices corporels et leurs effets dans des attitudes corporelles comme les convulsions, les grimaces ou les pleurs ; c'est ce dispositif tout entier de convulsions, grimaces et pleurs d'un organisme soumis à des sévices corporels qui régit notre usage quotidien du terme « douleur » et caractérise le concept correspondant.

La stratégie de l'interprète

De cette thèse de philosophie du langage résultent alors des thèses en philosophie de l'esprit, notamment que le phénomène de douleur² — et plus clairement encore ceux de désir ou de croyance — ne réside pas principalement dans des *qualia* mais principalement dans des réactions corporelles à des stimuli, et qu'il ne s'étudie pas prioritairement à la première personne, mais bien à la troisième personne. D'où l'idée d'hétérophénoménologie : loin de se fonder de manière privilégiée sur l'introspection, la phénoménologie

¹ Ludwig Wittgenstein, *Recherches philosophiques*, §§ 383-384.

² Pour Daniel Dennett (*Consciousness Explained*, Boston, Little, Brown and Company, 1991, trad. fr. *La Conscience expliquée*, Paris, Odile Jacob, 1993, p. 83, 475-482), la douleur est un état fonctionnel.

résulterait essentiellement d'une « stratégie de l'interprète »¹ qui attribue à un individu des douleurs, désirs ou croyances sur la base de la manière dont il se comporte dans telles ou telles circonstances². Et c'est d'ailleurs là ce qui, par extension de nos jeux de langage, permet aussi d'attribuer des sensations et des attitudes intentionnelles à des *animaux* — c'est tout l'objet de l'éthologie —, mais aussi à des *organismes élémentaires comme des amibes* — qui, par réaction chimique, « fuient » les milieux acides³ —, voire à de pures *machines*, qui se comportent de manière plus ou moins analogue à nous et satisfont donc plus ou moins bien nos critères d'usage des termes mentaux⁴. La possibilité même d'une telle extension rend particulièrement manifeste le fait que nos concepts phénoménologiques ne trouvent pas leur caractérisation dans des expériences subjectives mais bien dans des attitudes comportementales.

Il y a évidemment un certain nombre d'objections qu'on peut adresser à l'encontre d'une telle conception des phénomènes de conscience. Et, dans le style caractéristique de dialogues internes qui caractérise ses écrits, Wittgenstein s'en adresse d'ailleurs lui-même un certain nombre.

Un premier et très important contre-argument à l'égard du modèle hétérophénoménologique, c'est l'éventualité de la simulation. On voit bien en effet qu'il est — au moins théoriquement — possible de *feindre* la douleur, c'est-à-dire d'adopter les mêmes attitudes corporelles que quelqu'un qui souffre sans pour autant ressentir la douleur. Et le fait que, dans ce cas-là,

¹ Il y a, dans l'idée de stratégie de l'interprète, une part incontestable d'antiréalisme mentaliste (D. Dennett, « True believers : The intentional strategy and why it works », in W.G. Lycan (éd.), *Mind and Cognition*, Oxford, Blackwell, 1990, p. 151, 160-161). Pour sa part, cependant, Dennett parle de « semi-existence » des attitudes intentionnelles, dans la mesure où il leur reconnaît l'objectivité de tels « patterns » (D. Dennett, « De l'existence des patterns », in D. Fisette et P. Poirier (éds.), *Philosophie de l'esprit. Psychologie du sens commun et science de l'esprit*, Paris, Vrin, 2002, p. 153-161).

² Daniel Dennett, « True believers : The intentional strategy and why it works », art. cit., p. 151-157.

³ Daniel Dennett, *La Conscience expliquée*, op. cit., p. 219-220.

⁴ Ainsi, on dira que le thermostat « sait » que la température de la pièce est trop basse et qu'en conséquence il « envoie (à la chaudière) l'ordre » de se remettre en marche. Sur cette interprétation intentionaliste d'objets physiques, voir Daniel Dennett, « True believers : The intentional strategy and why it works », art. cit., p. 161-163 et déjà Ludwig Wittgenstein, *Recherches philosophiques*, op. cit., §§ 357-361, 390, 418. Wittgenstein (§§ 282-286) fait remarquer qu'il est même possible, dans des œuvres de fiction, d'attribuer des états intentionnels à de simples pots, à condition de leur donner un corps et, par ce moyen, une personnalité.

on dira que je simule et non que je souffre semble bien montrer que les critères de douleur ne tiennent pas dans les attitudes corporelles mais bien dans l'expérience subjective, qui fait ici toute la différence. Wittgenstein ne nie évidemment pas cette différence, mais, en philosophe du langage, il estime que, là encore, le sujet à qui on attribue la douleur n'est pas le seul à pouvoir déterminer s'il simule ou non ; très souvent il arrive qu'après avoir attribué à quelqu'un de la douleur sur la base de ses attitudes, nous rectifions ensuite notre jugement et posons désormais un verdict de simulation... sur la base d'autres de ses attitudes corporelles. Le concept de simulation lui-même semble donc ancré dans un jeu de langage intersubjectif¹ ; la différence entre souffrance réelle et simulation est essentiellement liée à la possibilité au moins théorique qu'à un moment, les attitudes corporelles trahissent le simulateur, c'est-à-dire qu'elles fassent la différence.

Avec l'idée de simulation — d'attitude corporelle dénuée du phénomène de conscience subjectif correspondant — se joue aussi le modèle du « zombie » (philosophique), c'est-à-dire de la créature qui est en apparence semblable à un être humain (ou à un animal) mais qui est totalement dépourvue de conscience subjective de sorte qu'elle n'éprouve aucun des phénomènes de conscience (sensations, émotions, désirs, croyances...) dont elle mime pourtant les attitudes corporelles². D'après les critères comportementaux qu'il retient, le behavioriste logique — ou l'hétérophénoménologue — semble ne pouvoir faire aucune différence entre l'homme (ou l'animal) doté d'une intentionalité intrinsèque et le zombie.

Et on trouve en effet, chez Wittgenstein, dans le cas d'au moins un phénomène de conscience (intentionnel), l'affirmation explicite qu'une différence dans ce qui est subjectivement ressenti qui ne ferait pas de différence dans les attitudes corporelles n'a aucune importance. Ce cas, c'est celui de l'« aveugle à la signification », entendu comme une sorte de zombie qui montrerait tous les signes comportementaux de compréhension de certaines expressions linguistiques — il les utiliserait dans les mêmes circonstances que les autres membres de sa communauté et réagirait de manière similaire à eux lorsque ces expressions seraient prononcées devant lui — mais qui serait pour autant dénué de toute conscience phénoménale correspondant à cette compréhension et à cette compétence. À la question de savoir si un tel zombie peut ou non se voir reconnaître des intentions de signification,

¹ Ludwig Wittgenstein, *Recherches philosophiques*, *op. cit.*, §§ 249-250

² Cf. aussi le fou de David Lewis (« Douleur de fou et douleur de martien », in D. Fiset et P. Poirier (éds.), *Philosophie de l'esprit. Psychologie du sens commun et sciences de l'esprit*, *op. cit.*, p. 289-290).

Wittgenstein répond positivement. Tout d'abord, en effet, il est évident que la conscience phénoménale qui accompagne généralement la compréhension — le « vécu » de compréhension — n'est pas une condition *suffisante* de l'intention de signification ; qui dirait éprouver ce vécu sans pour autant se montrer capable d'utiliser correctement le terme et de réagir correctement à son expression ne se verrait certainement pas attribuer l'intention de signification. Mais Wittgenstein conteste même que le vécu de compréhension soit une condition *nécessaire* de l'intention de signification ; lorsque nous attribuons des intentions de signification, nous nous contentons d'effectuer des vérifications comportementales et ne contrôlons absolument pas si le sujet ressent bien le vécu de compréhension idoine. L'expérience subjective n'entre tout simplement pas en ligne de compte lorsque nous attribuons une intention de signification¹. Et cela semble indiquer qu'elle n'est pas une composante *essentielle* de cet état intentionnel. On sait qu'Alan Turing, qui fut un élève de Wittgenstein, adoptera plus explicitement ce point de vue en suggérant d'accorder des états intentionnels à des dispositifs — êtres humains ou machines — sans se préoccuper du tout de ce qu'ils ressentent et en « testant » seulement leurs attitudes comportementales².

Dans une telle perspective, le vécu subjectif apparaît alors comme un simple épiphénomène³. Ni suffisant ni nécessaire à l'état intentionnel, il ne lui est pas essentiel et, bien qu'il accompagne souvent cet état, pourrait tout aussi bien être absent sans que cela change quoi que ce soit de fondamental à l'état en question⁴. L'idée d'épiphénomène indique aussi qu'il n'y joue

¹ Comprendre ou vouloir dire, ce n'est pas simplement avoir une certaine expérience subjective ou une certaine image dans la tête (L. Wittgenstein, *Recherches philosophiques*, *op. cit.*, §§ 35, 138-148, 156, 191).

² Alan Turing, « Computing machinery and intelligence », *Mind*, Oxford University Press, vol. 59 (236), 1950, p. 433-460, trad. fr. « Les Ordinateurs et l'Intelligence », in *La Machine de Turing*, Paris, Le Seuil, 1995, p. 133-175.

³ À noter que Dennett conteste que le vécu subjectif de douleur soit un « épiphénomène » au sens biologique d'un dispositif « accidentellement » sélectionné par l'évolution, c'est-à-dire sélectionné non pas en vertu d'un avantage de survie propre, mais du fait d'un lien contingent avec un autre dispositif qui présente une fonction de survie. Dennett fait remarquer que seule la douleur en tant que dispositif fonctionnel — et non en tant que « contenu phénoménal » — pourrait être ainsi accidentellement sélectionnée. Et, en tant que dispositif fonctionnel, la douleur n'est pas un épiphénomène ; elle a bien une valeur de survie. Cf. Bruno Leclercq, « *What is it like to be a bat ?* », art. cit., p. 294.

⁴ Ludwig Wittgenstein, *Recherches philosophiques*, *op. cit.*, § 272.

aucun rôle *explicatif*; c'est un « rouage à vide », dit Wittgenstein¹. Certes, après que j'ai multiplié les efforts pour comprendre une nouvelle règle (mathématique, par exemple), il peut se produire subitement un sentiment de compréhension (« Tilt ! Ca y est, j'ai compris ! »)². Mais, insiste Wittgenstein, que celui-ci ne soit pas condition *suffisante* de la compréhension, c'est ce que montre le fait que, même si j'ai ressenti ce vécu, on ne dira pas que je comprends si, ensuite, je me montre incapable d'appliquer correctement la règle³. Et qu'il ne soit pas condition *nécessaire*, c'est ce que montre le fait qu'il m'arrive souvent d'appliquer correctement la règle sans ressentir ce vécu⁴.

Aux côtés de la distinction entre douleur simulée et douleur effectivement ressentie, et de la distinction entre un zombie et un être proprement conscient, Wittgenstein prend encore en considération d'autres arguments en faveur de la reconnaissance de l'irréductibilité du vécu subjectif à sa manifestation comportementale. Les objections qu'il s'adresse font en effet signe vers des problèmes classiques — comme celui de Molyneux discuté par Locke, Berkeley, Leibniz ou Diderot — ou celui du spectre de couleur inversé plus récemment mis à l'ordre du jour par Sidney Shoemaker⁵. Rien,

¹ Ludwig Wittgenstein, *Recherches philosophiques*, *op. cit.*, §§ 202, 232, 270-271, Partie II sect. XI, p. 300-301 ; *Remarques sur la philosophie de la psychologie, I*, *op. cit.*, §§ 184, 215, 346 ; Daniel Dennett, *La Conscience expliquée*, *op. cit.*, p. 76-82.

² Ludwig Wittgenstein, *Recherches philosophiques*, *op. cit.*, §§ 151-155, 159, 165, 179, 316-320. « Épiphénomène », « rouage à vide », le vécu de compréhension (lequel se mêle parfois d'un sentiment de familiarité §§ 166-169 ou d'un sentiment de soulagement §§ 179-181) est parfois actionné par le reste du dispositif fonctionnel dans lequel réside la compréhension, mais lui-même n'intervient pas de manière décisive dans le dispositif, qui peut parfaitement fonctionner sans lui.

³ *Ibid.*, §§ 94, 139, 148, 149, 152, 153, 167, 183.

⁴ *Ibid.*, § 157. Faut-il nécessairement que ma conscience phénoménale se soit spécifiquement modifiée la première fois que j'ai pu lire une phrase française ? Il se peut que j'aie eu, à un moment précis, le sentiment de comprendre ce qu'était lire et que moi-même je lisais. Mais ce n'est pas forcément le cas. En fait, l'apprentissage a été très progressif et il n'est même pas certain qu'on puisse forcément dire quand j'ai lu ma première phrase...

⁵ Sydney Shoemaker, « The Inverted Spectrum », in *The Journal of Philosophy*, vol. 79 (7), 1982, p. 357-381. Cf. aussi, contre le réductionnisme neurophysiologique, l'argument de l'aveugle qui apprendrait tout ce qu'il faut savoir de la perception des couleurs d'un point de vue optique et neurophysiologique sans jamais avoir éprouvé les sensations propres aux couleurs dont il connaît les longueurs d'onde et les effets neuronaux. Dans ce cas, en effet, il semble bien qu'éprouver la sensation constitue une information supplémentaire et indépendante des informations connues à la

en effet, ne garantit que deux sujets qui se comportent de manière similaire — et utilisent par exemple de la même manière les termes de couleurs — fassent de la même façon l'expérience du rouge, du bleu ou du jaune... Si l'apprentissage du langage passe essentiellement par la normalisation des attitudes verbales dans certaines circonstances environnementales, il n'est précisément pas certain que deux utilisateurs du langage éprouvent de la même manière le contact avec les objets du monde que, sur le plan comportemental, ils s'accordent à appeler « rouges »¹. Cela semble bien montrer qu'au-delà des similarités comportementales — similarités de « surface » — il peut y avoir, dans la « boîte noire » que les behavioristes se refusent d'ouvrir, des différences essentielles dont ne rend pas compte l'hétéro-phénoménologie.

La réponse de ce behavioriste logique qu'est Wittgenstein ne consiste évidemment pas à nier la possibilité de telles différences d'expériences subjectives sous la similarité des attitudes verbales, mais à insister sur le fait que ces différences ne sont manifestement pas *essentiels* dans le jeu de langage des expériences qualitatives et dans l'usage des termes de sensation (couleurs, sons, etc.), de sorte que l'on peut dire que ce n'est pas seulement ni même principalement le vécu subjectif qui est le véritable phénomène de l'expérience du rouge².

Les partisans de la thèse contraire répliqueront sans doute en faisant valoir qu'à supposer que toute une communauté d'individus décide d'appeler « *beetle* » ce que chacun conserve précieusement dans une boîte dont il est le seul à pouvoir regarder le contenu, l'accord des usages verbaux à l'égard de ces contenus de boîte ne garantirait pas pour autant qu'il s'agit bien, dans chaque cas, des mêmes objets ; appeler « *beetle* » le contenu de la boîte quel qu'il soit permet sans doute d'accorder les pratiques verbales en surmontant l'inaccessibilité intersubjective du contenu de la boîte de chacun³, mais on voit trop bien qu'on ne rend justement pas compte de cette manière de ce qu'il y a précisément dans la boîte de chacun.

troisième personne (Frank Jackson, *Perception*, Cambridge, Cambridge University Press, 1977 ; « Epi-phenomenal qualia », in *Philosophical Quarterly*, vol. 32, 1982, p. 127-136.). Dans le même sens, voir Ned Block, « Troubles with functionalism », in W. Savage (éd.), *Perception and Cognition. Minnesota Studies in Philosophy of Science*, vol. IX, Minneapolis, University of Minnesota Press, 1978, p. 261-325.

¹ Ludwig Wittgenstein, *Recherches philosophiques*, *op. cit.*, § 271.

² *Ibid.*, § 304.

³ *Ibid.*, § 293 ; Daniel Dennett, « Quining qualia », in W.G. Lycan (éd.), *Mind and cognition*, Oxford, Blackwell, 1990, p. 524.

Or, en fait, c'est surtout cette image même de la boîte et de son contenu, de l'extérieur et de l'intérieur, mais aussi de la surface et de la profondeur, que Wittgenstein entend combattre, raison pour laquelle il se démarque explicitement du behaviorisme ontologique et épistémologique qui, à l'affirmation mentaliste selon laquelle il faut investiguer la profondeur de l'esprit sous la surface du comportement (aller voir dans la boîte et ne pas s'en tenir à ses apparences et attitudes externes), répondrait qu'il faut au contraire s'en tenir à la « surface » comportementale, aux apparences et attitudes « externes ». Un tel behaviorisme partagerait avec le mentalisme le mythe de la séparation de l'intériorité et de l'extériorité, de la surface et de la profondeur, mythe que Wittgenstein entend précisément dénoncer au nom de sa philosophie du langage¹. Le phénomène de douleur — ou celui d'expérience de rouge — qui est à l'œuvre dans nos jeux de langage n'est ni celui d'une expérience privée en deçà de tout comportement ni celui d'un comportement auquel il faudrait s'en tenir faute de pouvoir investiguer plus profondément avec les méthodes de la science. Combattre l'idée du fantôme dans la machine, ce n'est pas dire qu'il n'y a rien — ou qu'on ne peut pas savoir ce qu'il y a — à l'intérieur de la machine, mais c'est bien plutôt contester l'image même de la machine et de son poste de commande interne. Dire que les phénomènes de conscience thématés par le langage ne sont pas les expériences subjectivement ressenties dont autrui peut seulement deviner l'existence à travers leurs manifestations comportementales, ce n'est pas dire qu'ils se réduisent à ces seules manifestations comportementales d'expériences subjectives, dont on semblerait alors refuser d'interroger — ou même carrément nier — les sources subjectives. C'est plutôt dire que le jeu de langage peut et doit prendre pour critères les attitudes comportementales, lesquelles constituent donc *l'essentiel*, mais pas nécessairement le tout, des phénomènes étudiés².

Les arguments neurophysiologiques contre le théâtre cartésien

L'hétérophénoménologie s'inscrit plus nettement encore sur le terrain de la philosophie de l'esprit lorsqu'elle prolonge sa critique du mythe du fantôme dans la machine par des arguments directement empruntés aux neurosciences. C'est en particulier le cas chez Daniel Dennett, qui met en cause le

¹ Ludwig Wittgenstein, *Recherches philosophiques*, *op. cit.*, § 305-308. Voir sur ce point Bruno Leclercq, « *What is it like to be a bat ?* », art. cit., p. 304-305.

² Daniel Dennett, « Quining qualia », art. cit., p. 521, 525-526.

mythe du théâtre cartésien — ce « lieu » central de l'esprit où toutes les sensations et représentations devraient défiler pour accéder à la conscience subjective¹ — au nom du fait que l'information n'est centralisée nulle part dans le cerveau mais au contraire largement distribuée sur l'ensemble du système nerveux voire de l'organisme (notamment dans le système hormonal), et connaît toujours, pour cette raison, une multitude de « versions » successives depuis les premières décharges des neurones de perception périphériques jusqu'aux dernières décharges des neurones moteurs². Un certain nombre de phénomènes de conscience, comme les arcs réflexes — retrait immédiat des doigts qui se brûlent — sont même très largement décentralisés, étant entièrement gérés par des circuits nerveux courts autour de la moelle épinière, qui ont seulement quelques effets collatéraux dans le cerveau. Où se situe la conscience de brûlure ? Et où se situe l'intention de retirer les doigts ? Les deux phénomènes doivent-ils coïncider quelque part ? Et l'un doit-il précéder l'autre ? Il semble que l'expérience subjective de la douleur ne soit proprement consciente que bien après que les doigts ont été retirés de la source des brûlures ; loin d'être à l'origine de l'intention de retirer la main, cette expérience semble être un épiphénomène par rapport à l'arc réflexe. En fait, il y a plusieurs versions de la sensation de brûlure puisqu'une certaine intensité d'excitation neuronale de plusieurs capteurs périphériques entraîne l'excitation de neurones légèrement plus centraux auxquels les premiers sont connectés de différentes façons, et que cette excitation se répercute au long de la chaîne neuronale jusqu'à la moelle épinière où se fait le contact avec les premiers neurones moteurs, mais aussi avec des neurones qui transmettent l'excitation de la moelle épinière jusqu'au cerveau³. Cela a-t-il dès lors du sens de distinguer ce qui est antérieur et postérieur à la prise de conscience ? Et peut-on vraiment prétendre qu'une seule et même qualité sensible est bien préservée tout au long du processus ?

Bien plus : un certain nombre de ces arcs réflexes s'opèrent sans réelle prise de conscience subjective au sens fort, comme c'est le cas dans le sommeil ou lorsque l'attention est requise ailleurs. Et la même chose vaut d'ailleurs pour toute une série de phénomènes de conscience plus élaborés que les arcs réflexes comme la réaction « automatique » à des stimuli visuels, sonores et olfactifs qui n'occupent pas le centre de notre attention et ne sont pas, à cet égard, pleinement conscients. La conduite automobile, mais aussi le déplacement pédestre et toute une série d'activités quotidiennes reposent

¹ Daniel Dennett, *La Conscience expliquée*, *op. cit.*, p. 56-57, 135-178, 286-287.

² *Ibid.*, p. 146-176.

³ *Ibid.*, p. 306-309, 515-519.

sur la prise en considération de données sensorielles et sur la réaction « automatique » à leur égard sans qu'il y ait proprement expérience subjective de ces données et de l'intention d'y réagir (de sorte qu'il peut même y avoir ultérieurement surprise de découvrir qu'on avait vu telle ou telle chose et qu'on y avait réagi de telle ou telle façon). Il y a là, semble-t-il, toute une série de phénomènes de conscience intentionnelle — car, clairement, à celui qui conduit une voiture et se décale légèrement vers la gauche pour éviter une ornière, on peut et même doit attribuer des perceptions et des intentions — qui ne reposent pas sur une conscience phénoménale au sens fort.

Or, le modèle neurophysiologique que met en évidence Dennett explique bien la chose : il n'y a pas, dans le système nerveux, de théâtre cartésien où devraient converger toutes les sensations pour accéder à la conscience et pouvoir alors ensuite donner lieu à des réactions motrices proprement intentionnelles parce que initiées par la conscience ; dans l'organisme, la transmission d'informations sous forme d'excitations (et de décharges) neuronales, mais aussi de flux hormonaux... s'effectue de façon très largement *décentralisée*, dans des circuits partiellement indépendants les uns des autres et qui ne sont même pas vraiment coordonnés par un opérateur central. Assurément, les différents circuits interagissent entre eux et trouvent, dans la pratique, à se coordonner. Mais ce n'est pas là le fait d'un opérateur central, qui serait systématiquement informé de ce que font les autres sous-systèmes et d'où partiraient tous les influx endogènes. La coordination est elle-même distribuée sur plusieurs sous-systèmes et c'est essentiellement par leurs interactions entre eux ainsi qu'avec leur environnement commun que, de manière non planifiée, les sous-systèmes de l'organisme en viennent à se coordonner¹.

On voit en quoi ce modèle neurophysiologique plaide en faveur de la stratégie de l'interprète pour rendre compte des phénomènes de conscience : ces phénomènes s'attribuent à un organisme tout entier sur la base de son interaction globale à son environnement plutôt qu'à partir de la prise en considération de ce qui est subjectivement ressenti par la conscience de cet organisme, expérience subjective qui semble d'autant moins essentielle dans ce phénomène de conscience qu'elle peut venir après coup ou même être

¹ Daniel Dennett, « Why the law of effect will not go away », in W.G. Lycan (éd.), *Mind and Cognition*, Oxford, Blackwell, 1990, p. 70-72 ; *La Conscience expliquée*, *op. cit.*, p. 237.

totale¹. Lorsqu'il y a conscience phénoménale — et c'est évidemment souvent le cas —, elle ne semble être qu'un épiphénomène.

Dans cette perspective, il y a finalement, semble-t-il, très peu de place pour la subjectivité au sens d'une instance unitaire qui serait au centre de la conscience et même à l'initiative de ses « actes intentionnels ». Pour Dennett, l'unité de la conscience est essentiellement l'effet des interactions entre systèmes décentralisés d'un même organisme et de leurs interactions avec un environnement commun ; en outre, elle est elle-même attribuée à un organisme à travers la stratégie de l'interprète : à voir un organisme agir de manière pertinente dans son environnement, on lui attribue (extrinsèquement) une conscience unitaire ou une subjectivité², laquelle n'est donc clairement pas un fantôme dans la machine mais seulement un pôle narratif permettant de rendre compte de ce qui fait la cohérence globale du comportement de la machine³. Et on voit bien qu'on peut, à des degrés divers, attribuer une telle subjectivité à des dispositifs différents : êtres humains, animaux supérieurs, petits animaux⁴...

La notion d'épiphénomène en question

Tout cela met évidemment fortement en cause le modèle cartésiano-brentanien que semblent pourtant vouloir réhabiliter aujourd'hui les partisans du rôle essentiel de la conscience phénoménale dans l'intentionnalité. Mais, à vrai dire, cela me semble tout aussi bien remettre en cause le modèle du zombie. Ce que montre Dennett, c'est précisément la grande continuité entre les stimuli, les processus neurophysiologiques internes et les réactions corporelles, continuité qui n'est précisément possible que parce qu'il y a quelque chose qui sous-tend le fonctionnement de la boîte noire ; seul un ensemble complexe de sous-systèmes interconnectés peut fonder un comportement global cohérent dans un ensemble de situations très diverses. Bien plus, même s'il faut sans doute renoncer à l'idée que la conscience phénoménale serait le *terminus ad quem* de toutes les perceptions et le *terminus a quo* de tous les actes intentionnels de la conscience, il faut

¹ C'est sur ce point notamment que se joue le débat avec John Searle. Cf. Bruno Leclercq, « *What is it like to be a bat ?* », art. cit., p. 297-298.

² Daniel Dennett, « True believers : The intentional strategy and why it works », art. cit., p. 162-163.

³ Daniel Dennett, *La Conscience expliquée*, op. cit., p. 513-519, 529, 535-536.

⁴ Cf. Bruno Leclercq, « *What is it like to be a bat ?* », art. cit., p. 309 sq.

probablement aussi interroger, à l'inverse, l'idée qu'elle ne serait qu'un épiphénomène, totalement marginal dans les phénomènes de conscience. Ce que le modèle de la conscience distribuée et à versions multiples indique, c'est au contraire qu'on ne peut pas dire à quel endroit du processus a lieu exactement la conscience phénoménale ; elle est sans doute elle-même très largement distribuée, de sorte qu'il n'est pas si évident que ça de dire qu'elle vient « après coup ». S'il est sans doute faux de dire que, pour fonder l'intentionnalité, la conscience phénoménale doit nécessairement constituer quelque poste central de commande situé au carrefour de tous les trafics neuronaux qui relient les sensations et les mouvements corporels, il est aussi faux de dire qu'elle ne se trouve en fait que sur une « voie de garage ». C'est tout un réseau de trafic neuronal qui fait la conscience intentionnelle et c'est de ce même réseau tout entier qu'émerge la conscience subjective.

C'est là, me semble-t-il, ce qui ressortait déjà de la conception de l'esprit exprimée dans les *Principles of Psychology* de William James, qui avaient tant marqué Edmund Husserl¹. James s'était en effet efforcé de montrer que des processus neurophysiologiques à l'œuvre dans le cerveau ne pouvait émerger qu'un *flux* de conscience. Si, disait James, l'atomisme des idées n'est pas possible — si la même sensation ne peut être éprouvée à chaque fois à l'identique —, c'est parce que chaque sensation laisse sa trace dans le système nerveux et, en renforçant ou déforçant certaines de ses liaisons neuronales, modifie la structure plastique de ce système, de sorte que, même si exactement les mêmes stimuli s'y représentaient, ils seraient nécessairement accueillis d'une manière un peu différente². En outre, soulignait James, il est évident qu'une sensation ne peut jamais être reçue isolément ; tout passant par le même réseau neuronal, sa réception dépend forcément d'interactions avec une multitude de sensations simultanées mais aussi avec des sensations antérieures, qui ont encore des effets neuronaux quand surgissent les nouveaux stimuli, et des sensations postérieures qui commencent à exciter les capteurs périphériques³. C'est là toute l'idée

¹ Pour l'influence de James sur Husserl, cf. Bruno Leclercq, « Les données immédiates de la conscience. Neutralité métaphysique et psychologie descriptive chez James et Husserl », in *Philosophiques*, vol. 35 (2), 2008, p. 317-344. Sur la manière dont l'hétérophénoménologie peut relayer les projets de la phénoménologie husserlienne, voir Bruno Leclercq, « *What is it like to be a bat ?* », art. cit., p. 299-302.

² William James, *The Principles of Psychology*, New York, Dover Publications, 1950, chap. IX, vol. I, p. 233-234.

³ *Ibid.*, chap. IX, vol. I, p. 236, 255.

jamesienne des « franges » et du « halo » de conscience, dont Husserl fera un « horizon » de conscience¹.

Or, pour James, ce sont les recouvrements, accords, harmonies (*overtone*) ou au contraire les interférences, contrastes... que les sensations successives ne cessent d'entretenir dans le cerveau qui permettent l'apparition d'authentiques contenus de conscience — identiques dans différentes sensations — et avec eux l'émergence d'une conscience intentionnelle². Loin d'en être un moment particulier — central ou périphérique, essentiel ou épiphénoménal —, la conscience n'est donc rien d'autre que le flux tout entier.

Qu'est-ce qu'un contenu phénoménal ?

Sans nier que les états mentaux comportent généralement une dimension phénoménale — subjectivement ressentie —, l'hétérophénoménologie conteste que celle-ci leur soit essentielle au sens où elle constituerait ce dont il faudrait absolument rendre compte pour fournir une analyse satisfaisante de ces états mentaux. À cet égard, l'hétérophénoménologie semble radicalement opposée aux théories de l'intentionnalité phénoménale, qui affirment non seulement qu'intentionnalité et phénoménalité vont généralement de pair, mais qu'elles sont essentiellement indissociables, au double sens où tout état conscient — doté de caractères phénoménaux, d'un « effet que ça fait » — est nécessairement intentionnel et où, en sens inverse, tout état intentionnel (ou du moins tout état intrinsèquement intentionnel) est nécessairement conscient, doté de caractères phénoménaux, et n'est même intentionnel que dans la mesure où il est doté de ces caractères phénoménaux.

Le premier de ces sens — celui qui affirme que tout état doté de caractères phénoménaux, même une douleur, une angoisse ou la simple expérience d'une qualité sensible (un son, une odeur, une couleur...) est nécessairement intentionnel³ — pourrait paraître le plus acceptable à l'hétérophénoméno-

¹ *Ibid.*, chap. IX, vol. I, p. 251, 262, chap. XV, vol. I, p. 609, chap. XVI, vol. I, p. 644, 648, 653.

² Cf. Bruno Leclercq, « Les données immédiates de la conscience. Neutralité métaphysique et psychologie descriptive chez James et Husserl », art. cit., p. 332-334.

³ George Graham, Terence Horgan et John Tienson, « Consciousness and intentionality », in M. Velmans et S. Schneider (éds.), *The Blackwell Companion to Consciousness*, Oxford, Blackwell, 2007, p. 475 ; Terence Horgan et John Tienson, « The intentionality of phenomenology and the phenomenology of intentionality »,

logue. Cependant, contrairement au représentationnalisme d'un Dretske ou d'un Tye, les partisans de l'intentionnalité phénoménale, comme Uriah Kriegel, insistent sur la nécessité de distinguer du contenu représentationnel objectif qui se présente sous telle ou telle qualité sensible — la blancheur de ce mur, la rectangularité de ce plateau de table, le majestueux gigantisme de cet astre — un contenu proprement phénoménal et essentiellement subjectif — la fait de paraître rose, de paraître trapézoïdal ou de paraître petit¹ —, contenu phénoménal dont les critères d'identité ne peuvent manquer d'intriguer l'hétérophénoménologue. Kriegel estime en effet qu'il est possible de distinguer nettement l'*impression* que la Lune est petite du *jugement* qu'elle est gigantesque mais vue à grande distance, mais aussi l'*impression* que le plateau d'une table est trapézoïdal du *jugement* qu'il est rectangulaire mais vu sous un angle qui modifie sa perspective, ou encore l'*impression* que le mur est rose du *jugement* qu'il est blanc mais que s'y reflète une surface rouge. La chose n'est cependant évidente qu'en apparence. Ne peut-on pas tout aussi bien dire que l'on *voit* que la Lune est grande, que le plateau de table est rectangulaire et que le mur est blanc ? De l'impact des rayons de lumière sur la surface de la rétine à la prise de position judiciaire finale, n'y a-t-il pas en fait plusieurs versions de l'expérience perceptive plutôt qu'une seule qui serait celle qui marquerait le moment de la conscience phénoménale et que corrigerait ensuite le jugement ? Est-il si évident que cela qu'il y ait initialement un contenu phénoménal déterminé et que ne viennent qu'ultérieurement les inférences rationnelles qui permettent de fixer le contenu représentationnel objectif ?

D'une manière plus générale, les partisans de l'intentionnalité phénoménale entendent distinguer les contenus de conscience « étroits », entièrement caractérisés par leurs caractères phénoménaux, des contenus de conscience « larges », dont l'externalisme a raison de penser qu'ils supposent, en plus de ces caractères phénoménaux, des relations spécifiques — causales ou téléologiques — entre le sujet et les objets réels du monde². Et, renversant

in D.J. Chalmers (éd.), *Philosophy of Mind : Classical and Contemporary Readings*, Oxford University Press, p. 521-522.

¹ Uriah Kriegel, « Phenomenal content », in *Erkenntnis*, vol. 57, 2002, p. 180, 182, 188.

² George Graham, Terence Horgan et John Tienson, « Consciousness and intentionality », art. cit., p. 471 ; Terence Horgan et John Tienson, « The intentionality of phenomenology and the phenomenology of intentionality », art. cit., p. 527-528 ; Brian Loar, « Phenomenal intentionality as the basis of mental content », in M. Hahn et B. Ramberg (éds.), *Reflections and Replies : Essays on the Philosophy of Tyler Burge*. Cambridge (Mass.), MIT Press, p. 229-231.

exactement le point de vue des hétérophénoménologues, des auteurs comme Horgan et Tienson affirment que ce ne sont pas les contenus larges, mais les contenus étroits, qui sont les plus « cruciaux »¹ ou les plus « fondamentaux »² dans l'intentionnalité. Ce qui fait que je pense à Bill Clinton, que je juge qu'il ferait un meilleur candidat que John Kerry à l'investiture démocrate et que je veux voter pour lui, ce sont d'abord et avant tout des éléments internes à ma conscience, qui se présentent en elle avec une phénoménalité particulière — un effet que ça fait particulier —, même si, bien sûr, envisagés dans un sens plus large, ces contenus intentionnels dépendent aussi, pour leur détermination complète, des référents réels de ces intensions (Bill Clinton, John Kerry...)³.

Ces thèses internalistes — ou partiellement internalistes si on prend en considération la concession faite à l'externalisme en ce qui concerne les contenus larges — s'appuient, pour leur justification, sur une série d'expériences de pensées telles que celles de moi et de mon jumeau psychique sur la Terre Jumelle, de moi et de mon jumeau psychique dont le cerveau désincarné baignerait dans une cuve où il serait artificiellement stimulé par des électrodes, de moi et de mon jumeau psychique qui serait illusionné par un Malin Génie⁴..., ou, en sens inverse, l'expérience de pensée de moi et d'un autre sujet confrontés au même monde et y réagissant de façon similaire, mais dont les consciences phénoménales seraient différentes en raison, par exemple, d'une inversion des spectres de couleurs affectant nos perceptions visuelles⁵. D'après les partisans de la conscience phénoménale, moi et mon jumeau psychique éprouvons les mêmes expériences subjectives⁶ même si nos contenus larges, qui dépendent de nos environnements réels, sont distincts ; et, à l'inverse, moi et mon *alter ego* dont le spectre de couleur

¹ Terence Horgan et John Tienson, « The intentionality of phenomenology and the phenomenology of intentionality », art. cit., p. 528.

² *Ibid.*, p. 529.

³ *Ibid.*, p. 528; George Graham, Terence Horgan et John Tienson, « Consciousness and intentionality », art. cit., p. 471-472.

⁴ Terence Horgan et John Tienson, « The intentionality of phenomenology and the phenomenology of intentionality », art. cit. p. 524-529 ; Brian Loar, « Phenomenal intentionality as the basis of mental content », art. cit., p. 230 ; Katalin Farkas, « Phenomenal intentionality without compromise », in *The Monist*, vol. 91 (2), 2008, p. 273, 285.

⁵ Uriah Kriegel, « Phenomenal content », art. cit., p. 176-177.

⁶ Katalin Farkas, « Phenomenal intentionality without compromise », art. cit., p. 273, 285.

a été inversé, éprouvons des expériences subjectives distinctes même si nos environnements réels sont identiques.

Il n'est toutefois pas du tout certain qu'on sache exactement ce que l'on veut dire par là... Qu'est-ce qui fait en effet l'*identité* de ces contenus phénoménaux ?

Graham, Horgan et Tienson¹ revendiquent partager leur conception du mental avec Brentano et Husserl, mais aussi, avant eux, avec Descartes et Locke. Et, en effet, on trouve, chez ces derniers, une discussion d'hypothèses similaires à celles du cerveau dans la cuve ou du spectre inversé des couleurs. Mais ce rapprochement avec les penseurs modernes a toutes les raisons de nous inquiéter plutôt que nous rassurer. Les théories de la représentation de Descartes et Locke — c'est précisément un mérite de Husserl que de l'avoir montré² — ne tiennent leur apparente plausibilité que d'un certain nombre de confusions majeures — entre les vécus de conscience et leurs contenus, entre les composantes phénoménales et les composantes proprement intentionnelles des vécus, entre le fait de vivre immédiatement les vécus et le fait de les prendre pour objets de connaissance... — avec lesquelles on ne peut que craindre que les théories de l'intentionnalité phénoménale entendent aujourd'hui renouer. Le contenu phénoménal ne serait-il donc rien d'autre que l'« idée » des Modernes, cette composante du vécu directement présente à l'esprit et indubitablement accessible à son savoir, qui présente un contenu (le triangle, Dieu, Bill Clinton) en vertu de sa seule apparence subjective³ et qui peut non seulement réapparaître à l'identique à différents moments de la vie de l'esprit mais aussi dans les esprits d'individus différents ?

Une telle conception a quelque chose de magique : il faut et il suffit que je fasse en moi une certaine expérience — celle qui correspond à l'idée du triangle — pour que je pense au triangle ou que je vise la signification

¹ George Graham, Terence Horgan et John Tienson, « Consciousness and intentionality », art. cit., p. 468.

² Cf. Bruno Leclercq, « Les idées dans l'esprit. De Descartes à l'empirisme britannique », dans B. Collette et B. Leclercq (éds.), *L'Idée de l'« Idée »*, Bruxelles, Peeters, 2012, p. 125-147.

³ George Graham, Terence Horgan et John Tienson, « Consciousness and intentionality », art. cit., p. 471, 473-474, 476. Dans cette perspective, on ne se trompe donc jamais sur le contenu étroit, mais seulement éventuellement sur le contenu large ; ou encore on ne se trompe jamais sur ce qui est directement éprouvé mais seulement éventuellement dans le jugement objectif qu'on formule ensuite (Terence Horgan et John Tienson, « The intentionality of phenomenology and the phenomenology of intentionality », art. cit., p. 526).

« triangle ». Et lorsque je vous parle, je ne me fais comprendre de vous que si je parviens à susciter en vous une expérience semblable à celle que moi-même j'éprouve en prononçant ces mots tout en pensant à ce qu'ils veulent dire¹. Mais cette conception est précisément celle dont toute la philosophie contemporaine, Husserl en tête, a dénoncé la naïveté et le psychologisme. Dans la mesure où elle est partageable, l'idée du triangle n'est pas un *vécu* subjectif, mais un *contenu* intersubjectif. Et ce qui, dans le vécu, saisit ce contenu est une intention de signification qui ne se caractérise pas essentiellement par sa dimension *phénoménale*, mais par sa dimension *logique*, pas par sa « matière » (*Stoff*) sensible, mais par sa forme ou son sens d'appréhension². Sans doute Graham, Horgan et Tienson insistent-ils sur le fait que le « *what it is like* » de la phénoménalité intentionnelle ne se réduit pas aux sensations et aux images qui composent, pour Husserl, la couche hylétique du vécu³ ; la phénoménalité intentionnelle, disent-ils, réside « dans une qualité cognitive plutôt que dans une qualité sensorielle ou imagino-sensorielle »⁴. Et cela est fort heureux, car la théorie à l'égard de laquelle ils expriment leur distance — celle qui réduirait les intentions de signification à des images subjectivement éprouvées — a été vigoureusement dénoncée pour son psychologisme par le Husserl des première et deuxième *Recherches logiques*. Mais l'antipsychologisme de Husserl, on le sait, ne se réduit pas à faire apparaître la dimension intentionnelle et non purement sensuelle du vécu; il porte aussi et peut-être surtout sur la distinction entre le vécu et ses contenus objectifs. Or, c'est d'abord et avant tout sur le plan des contenus objectifs, qui ne sont pas des composantes *réelles* — directement éprouvées — mais seulement *intentionnelles* du vécu, que se joue la question de

¹ Outre l'expérience subjective propre à tel ou tel contenu, il y aurait aussi une expérience subjective propre à telle ou telle attitude intentionnelle à l'égard de ce contenu. Ainsi, *croire que* Bill Clinton ferait un meilleur candidat que John Kerry ne serait pas vécu de la même façon que *se demander si* Bill Clinton ferait un meilleur candidat que John Kerry (Terence Horgan et John Tienson, « The intentionality of phenomenology and the phenomenology of intentionality », art. cit. p. 522-523). Est-ce vraiment évident ? Croire doit-il forcément être éprouvé autrement qu'envisager ? et autrement que savoir ?

² Edmund Husserl, *Recherches logiques*, trad. fr. Paris, Presses Universitaires de France, 1959-1963, voir *Prolégomènes à la logique pure*, *Recherches* I, II et V.

³ George Graham, Terence Horgan et John Tienson, « Consciousness and intentionality », art. cit., p. 468, 472; Terence Horgan et John Tienson, « The intentionality of phenomenology and the phenomenology of intentionality », art. cit., p. 520.

⁴ George Graham, Terence Horgan et John Tienson, « Consciousness and intentionality », art. cit., p. 468.

l'identité, ainsi que de l'itérabilité, et même de l'intersubjectivité, de ce qui est pensé. Comme l'avait déjà souligné James, le flux du vécu change sans cesse et ne revient jamais deux fois à l'identique ; ce qui peut réapparaître à l'identique, c'est le *contenu* senti ou pensé. De même, deux flux de conscience ne peuvent jamais parfaitement coïncider l'un avec l'autre, mais ils peuvent saisir des contenus identiques, lesquels ne se singularisent pas par leurs caractères *phénoménaux* mais par leurs traits *sémantiques*.

Outre qu'elle est neurophysiologiquement implausible, la thèse selon laquelle il faudrait toujours faire — et deux individus devraient nécessairement faire — la même expérience subjective pour viser la même signification ou penser la même chose se heurte de plein fouet à la question wittgensteinienne de savoir si cela a même du sens de dire que c'est la *même* expérience subjective qui a lieu deux fois. Qu'est-ce qui fait l'identité de ces fameux contenus phénoménaux ? Est-il vraiment certain que je fais la *même* expérience — que ça me fait le *même* effet — chaque fois que je pense à Bill Clinton ou que j'estime qu'il ferait un meilleur candidat que John Kerry¹ ? À chaque signification — même « et »² —, à chaque concept³, à chaque

¹ Bien sûr, il ne s'agit pas de nier que nous ressentons un certain nombre de choses lorsque nous pensons, ni que nous ressentons les choses différemment lorsque nous comprenons une expression (orale ou écrite) et lorsque nous ne la comprenons pas (« Consciousness and intentionality », art. cit., p. 473; « The intentionality of phenomenology and the phenomenology of intentionality », art. cit., p. 523) ou lorsque nous la comprenons en un sens et lorsque nous la comprenons en un autre sens (« The intentionality of phenomenology and the phenomenology of intentionality », art. cit., p. 523; U. Kriegel, « Is intentionality dependent upon consciousness ? », in *Philosophical Studies*, vol. 16, p. 286). Mais cela veut-il dire que nous devons toujours ressentir la même chose lorsque nous comprenons l'expression dans le même sens ? Et que d'autres doivent aussi ressentir la même chose que nous pour comprendre cette expression dans le même sens que nous ?

² Brian Loar, « Phenomenal intentionality as the basis of mental content », art. cit., p. 236-237.

³ Pour combattre l'externalisme sémantique, Brian Loar (« Phenomenal intentionality as the basis of mental content », art. cit., p. 232-233) insiste sur l'indépendance des intentions de signification à l'égard de leur référence effective. Et Uriah Kriegel (U. Kriegel, « Is intentionality dependent upon consciousness ? », art. cit., p. 287-290) poursuit la démarche en opposant *grasso modo* intension et extension des descriptions définies et d'autres expressions. Mais lorsque Putnam puis Burge montrent que « les significations ne sont pas dans la tête », ils établissent précisément que l'internalisme n'est pas seulement intenable pour la référence, mais aussi pour le sens ou la teneur conceptuelle, laquelle est partagée dans une communauté scientifique.

proposition, correspond-il nécessairement un contenu phénoménal propre ? Comment puis-je m'en assurer ? Et qu'est-ce qui permet de dire que vous et moi faisons la même expérience — vivons les mêmes contenus phénoménaux — lorsque nous pensons aux mêmes individus, visons les mêmes significations ou formulons les mêmes propositions ?...

... sinon, peut-être, que nous tendons à *nous exprimer dans les mêmes mots*... Or, cela pose évidemment la question de savoir si les contenus phénoménaux sont déjà eux-mêmes linguistiquement articulés — y a-t-il un langage de la pensée et est-il universellement partagé ? — ou si c'est plutôt le langage commun qui projette rétrospectivement ses formes sur le flux de conscience et, par là, donne un sens commun à ce qui, pour le reste, ne sont qu'expériences subjectives incomparables les unes aux autres ? Si c'est le langage commun qui impose ses contraintes, les contenus sont forcément larges et non étroits¹. Brian Loar, mais aussi Graham, Horgan et Tienson insistent sur le fait qu'en particulier dans le cas de la perception, mais aussi d'attitudes intentionnelles fondées sur la perception, comme la croyance perceptive non verbalisée qu'il y a là au fond de la pièce une chaise², les contenus phénoménaux de l'intentionnalité ne sont pas forcément linguistiquement articulés. Cela, évidemment, est de grande importance et permet sans doute de résister au moins partiellement à la perspective wittgensteinienne, qui envisage l'univers sémantique tout entier à partir du langage commun. Mais il reste qu'il n'est pas du tout évident que je doive à chaque fois faire la même expérience subjective pour penser et croire — en deçà de toute formulation linguistique — qu'il y a là une chaise. Il n'est même pas évident que je doive faire quelque expérience subjective que ce soit pour entretenir cette croyance. Préoccupé par ce que, simultanément, je vous dis, je peux en effet aller chercher cette chaise — et témoigner, par mon comportement, d'un certain nombre de croyances à son égard — sans jamais éprouver quoi que ce soit de similaire à ce que je ressens lorsque je m'efforce de me concentrer exclusivement sur ma croyance qu'il y a là une chaise.

¹ Uriah Kriegel, « Is intentionality dependent upon consciousness ? », art. cit., p. 276-277 ; Terence Horgan et John Tienson, « The intentionality of phenomenology and the phenomenology of intentionality », art. cit., p. 523.

² Terence Horgan et John Tienson, « The intentionality of phenomenology and the phenomenology of intentionality », art. cit., p. 523.

Tout état intentionnel est-il phénoménal ?

On retrouve ici l'autre thèse qui caractérise la théorie de l'intentionnalité phénoménale, à savoir que tout état intentionnel est conscient. Cette thèse se heurte à toute une série d'objections évidentes : il est très manifeste que, comme la croyance qu'il y a là une chaise que je peux mobiliser sans jamais me la rendre explicitement consciente, il y a un tas d'états mentaux intentionnels qui ne semblent pas nécessairement déployer de contenus phénoménaux propres. Pour soutenir leur thèse, les partisans de la théorie de l'intentionnalité phénoménale doivent donc écarter tous ces exemples en affirmant qu'il ne s'agit pas là de cas d'états proprement intentionnels. C'est en particulier le cas des états mentaux subdoxastiques, ces états physiques des systèmes qui sont inaccessibles à la conscience mais interviennent dans des processus purement causaux de transfert et de traitement de l'information¹. Sont également exclus tous les états mentaux dispositionnels « non occur-rents » comme les croyances et les désirs qui caractérisent un système sur de longues périodes et sont susceptibles d'influencer son comportement sans devoir systématiquement être explicitement au centre de son attention consciente². Et sont évidemment aussi exclues une multitude de croyances et intentions que des sujets extérieurs attribuent au système sur la base de son comportement mais que le système lui-même n'a jamais explicitées et serait parfois même incapable d'expliciter faute de faire lui-même, voire même de pouvoir faire lui-même, cette lecture de son propre comportement (par exemple les intentions inexplicites de Shakespeare écrivant Hamlet³ ou les désirs inavoués de tel patient de psychanalyse⁴).

Pour exclure ainsi de leurs analyses une bonne part des attitudes intentionnelles — états informationnels, croyances, intentions, désirs, craintes... — qui sont pourtant au centre de nos jeux de langage psychologiques quotidiens, les partisans de la conscience phénoménale doivent introduire une distinction entre états mentaux paradigmatiques (ou prototypiques), lesquels sont intrinsèquement intentionnels et répondent à l'exigence d'être conscients ou au moins potentiellement conscients⁵, et états mentaux non

¹ George Graham, Terence Horgan et John Tienson, « Consciousness and intentionality », art. cit., p. 470-471.

² *Ibid.*, p. 478.

³ *Ibid.*, p. 477.

⁴ *Ibid.*, p. 479.

⁵ Uriah Kriegel, « Is intentionality dependent upon consciousness ? », art. cit., p. 273-278.

paradigmatiques, dont l'intentionnalité n'est pas intrinsèque mais dérivée (en particulier de l'intentionnalité de l'interprète¹) et qui ne répondent pas à cette exigence de conscience². Une telle distinction, cependant, est elle-même très problématique.

Graham, Horgan et Tienson³ pensent pouvoir la fonder sur une intuition partagée par tous les locuteurs compétents quant à ce qui est ou non un état mental. Cette stratégie, cependant, est assez peu probante : il est facile de faire l'expérience de ce que certains locuteurs compétents qualifient de « mentale » la mise en accord grammaticale (en genre et en nombre) des termes d'une phrase, opération qui se fait souvent automatiquement et de manière largement inconsciente, et que d'autres estiment par contre que la douleur n'est pas un état mental mais un état physique. Par ailleurs, l'hypothèse que les locuteurs compétents ne qualifieraient de « mentaux » que les états dotés de conscience phénoménale semble mise à mal par le constat, fait par Graham, Horgan et Tienson eux-mêmes, que « c'est un axiome virtuel de la psychologie populaire que certains cas de croyance et de désir sont inconscients »⁴.

Une autre stratégie consiste à affirmer que les états intrinsèquement intentionnels ont un contenu déterminé contrairement aux états attribués par autrui⁵. Comme le reconnaissent cependant les auteurs⁶, il faut alors expliquer comment, en dépit de leur indétermination, les états intentionnels non conscients pourraient être attribués à autrui — car il y a manifestement un savoir collectif en la matière — et ce avec une certaine réussite prédictive. Sans pour autant en faire des réalités internes à l'esprit et exclusivement accessibles par introspection, le jeu de langage qui préside à la stratégie de l'interprète confère bien une certaine « objectivité » aux croyances et désirs attribués à autrui. C'est là d'ailleurs ce qui distingue l'hétérophénoménologie de Dennett de l'éliminativisme des Churchland⁷.

¹ *Ibid.*, p. 284-285.

² *Ibid.*, p. 283. Notons que c'était déjà la réponse que faisait Searle à Dennett.

³ G. Graham, T. Horgan et J. Tienson, « Consciousness and intentionality », art. cit., p. 470.

⁴ *Ibid.*, p. 477.

⁵ *Ibid.*, p. 478.

⁶ *Ibid.*, p. 478-479.

⁷ Paul Churchland, *Scientific Realism and the Plasticity of Mind*, Cambridge, Cambridge University Press, 1979 ; « Eliminative materialism and propositional attitudes », in *The Journal of Philosophy*, vol. 78, 1981, p. 67-90, trad. fr. « Le matérialisme éliminativiste et les attitudes propositionnelles », in D. Fisette et P. Poirier (éds.), *Philosophie de l'esprit. Psychologie du sens commun et sciences de l'esprit*,

Plus prometteuse pour caractériser l'intentionnalité intrinsèque, la stratégie, proposée par Uriah Kriegel à la suite de Colin Mc Ginn, de distinction des états mentaux sur la base de la présence ou non, dans leur contenu représenté, d'un lien explicite avec la conscience qui le pense¹ : contrairement au robot qui serait capable de reconnaître Bill Clinton, je ne pense pas seulement à Bill Clinton, mais à Bill Clinton pensé par moi ou à Bill Clinton et ce qu'il représente pour moi. Intéressante, cette suggestion, qui rapporterait en quelque sorte l'intentionnalité intrinsèque à la possibilité de « réduction » phénoménologique, a cependant le désavantage d'une certaine circularité, puisque c'est en la caractérisant par la présence en elle de la conscience de soi qu'on pourrait affirmer que, contrairement à l'intentionnalité dérivée, l'intentionnalité intrinsèque est nécessairement consciente. Par ailleurs, cette suggestion ne permet pas nécessairement de définir clairement les limites de l'intentionnalité intrinsèque, notamment sur l'échelle phylogénétique². Où commence la conscience de soi ? avec les êtres humains ? les animaux « supérieurs » ? les animaux ? les organismes ? Et qu'est-ce qui permet d'en décider ? Quels sont les critères d'attribution de la conscience de soi ? Un ordinateur en est-il forcément dépourvu ?

Par ailleurs, la stratégie de l'interprète — donc l'intentionnalité dérivée — suppose-t-elle nécessairement la conscience de soi — l'intentionnalité intrinsèque ? Il est en tout cas clair qu'un certain nombre d'animaux sont parfaitement capables de déployer une stratégie de l'interprète à l'égard du comportement d'un autre animal, c'est-à-dire de lui attribuer des intentions (hostilité, bienveillance...) et d'anticiper ses actions sur la base de ses attitudes corporelles présentes³. Du moins pouvons-nous leur attribuer une telle capacité... sur la base de leur propre comportement... Mais peut-être pouvons-nous aussi attribuer cette capacité à des ordinateurs qui tiennent compte de ce que savent d'autres processeurs...

Paris, Vrin, 2002, p. 117-151 ; *Matter and Consciousness*, Cambridge (Mass.), The M.I.T. Press, 1984 ; Patricia Churchland, *Neurophilosophy*, Cambridge (Mass.), The M.I.T. Press, 1986.

¹ Uriah Kriegel, « Is intentionality dependent upon consciousness ? », art. cit., p. 286-290, 295.

² *Ibid.*, p. 289.

³ Pour Daniel Dennett (« Why the law of effect will not go away », art. cit., p. 72-73), certains des opérateurs peuvent tenir leur intelligence de leur configuration et de leur interaction avec leur environnement. C'est d'ailleurs là ce que laisse penser l'explication darwinienne de ces phénomènes (p. 64-65, 73, 75-76).

En conclusion

Sans aucun doute y a-t-il des expériences subjectives, des *qualia*, un effet que ça fait de ressentir telle ou telle chose et parfois même de désirer telle ou telle chose, de croire telle ou telle chose, etc. Ces « phénomènes » subjectifs sont-ils pour autant essentiels à — et constitutifs de — l'intentionnalité ? Le débat n'est pas clos, mais il ne me semble pas que les partisans récents de la conscience phénoménale ont été beaucoup plus loin que Searle dans leur réponse aux objections des hétérophénoménologues. L'intentionnalité semble bien être une propriété de dispositifs — pas seulement organiques — plus qu'une propriété intrinsèque à l'intimité subjective. Bien plus, dans les dispositifs organiques qui en sont capables, la phénoménalité elle-même semble distribuée sur l'ensemble du système nerveux plutôt que concentrée dans un hypothétique théâtre cartésien...