

(Article à paraître)

## Représentationnalisme et langage privé : une défense wittgensteinienne du caractère non-représentationnel de la phénoménalité

### Introduction

Affirmer que notre théorie de l'esprit humain doit se fonder sur la notion de représentation n'est pas précisément une idée neuve en philosophie, et jusqu'à un certain point Aristote peut être considéré comme le premier défenseur d'une telle conception. En un sens, donc, le représentationnalisme constitue une approche philosophique vénérable, dont on peut suivre les évolutions depuis ses origines antiques jusqu'à ses différentes reprises, à l'âge classique et à l'époque contemporaine. On comprend ainsi parfois sous le terme de « théorie représentationnelle » de l'esprit quelque chose d'équivalent à la théorie computationnelle de l'esprit (Fodor, 1975), c'est-à-dire une théorie qui repose sur l'idée, essentielle pour une grande partie des sciences cognitives contemporaines (et pour leurs tentatives de fondation philosophique), selon laquelle l'esprit humain doit être étudié comme un système de traitement de l'information, c'est-à-dire comme un système qui forme, utilise et manipule des représentations.

L'approche à laquelle cet article s'intéresse, et que l'on nomme « représentationnalisme », constitue toutefois une tentative plus spécifique consistant à comprendre, à travers cette notion de représentation, un aspect particulier de l'esprit : cet aspect que l'on nomme « expérience consciente », « conscience phénoménale » (Block, 1995) : ou, en d'autres mots, les états mentaux qui sont tels qu'il y a un certain « effet que cela fait » (Nagel, 1974) d'être, pour un sujet, dans ces états. En ce sens plus spécifique, le représentationnalisme constitue une position philosophique relativement récente dans la philosophie de l'esprit contemporaine, dont l'apparition peut être repérée dans les années 1990<sup>1</sup> (Dretske, 1995; Harman, 1990). Cette nouveauté ne l'empêche pas d'être assez largement accepté, et de constituer peut-être la principale approche théorique actuelle en ce qui concerne la compréhension de la nature et des déterminations de nos expériences conscientes<sup>2</sup>.

Il existe de nombreuses façons de définir le représentationnalisme concernant l'expérience consciente ; et, suivant la manière dont cette approche est interprétée, plusieurs arguments distincts ont pu être fournis en sa faveur. Je projette de m'intéresser ici à une thèse, que j'identifierai au représentationnalisme, et qui constitue une interprétation assez faible de celui-ci, d'après laquelle le contenu représentationnel d'une expérience consciente fixe et détermine entièrement le caractère phénoménal de cette expérience consciente (c'est-à-dire l'effet que cela fait d'avoir cette expérience).

Dans cet article, je veux m'intéresser à certaines raisons d'accepter le représentationnalisme qui ont joué, selon moi, un rôle important dans les succès récents de celui-ci. Ces raisons sont de deux types. Il s'agit, d'un côté, d'une série d'arguments fournis par Dretske (l'un des premiers et

---

<sup>1</sup> En proposant une telle datation, je présuppose déjà une certaine acception du représentationnalisme, qui correspond à ce que je vais détailler un peu plus bas. En d'autres sens en effet, l'idée de comprendre la conscience à travers la notion de représentation et d'intentionnalité est en effet beaucoup plus ancienne, aussi bien dans la tradition analytique (Searle, 1983) que dans la tradition phénoménologique (Husserl, 1962, 1969, particulièrement la cinquième Recherche), ainsi que dans les travaux d'auteurs communs aux deux traditions (Brentano, 1944).

<sup>2</sup> A titre indicatif, dans les sondages « PhilPapers Surveys » menés auprès de plus de 3000 personnes (philosophes professionnels et étudiants avancés en philosophie), le représentationnalisme était la principale théorie de l'expérience perceptive embrassée par les sujets interrogés (avec 31,5% des personnes se déclarant en sa faveur), devant la théorie des *qualia* (12,2%) et le disjonctivisme (11%). Voir <http://philpapers.org/surveys/results.pl>

principaux défenseurs du représentationnalisme contemporain concernant l'expérience consciente) dans l'un de ses articles (Dretske, 2003) ; arguments qui s'articulent tous autour du problème de *l'individuation des expériences conscientes*, et qui me semblent figurer parmi les arguments les plus forts et les plus influents en faveur du représentationnalisme. Les autres raisons résident dans un ensemble de débats concernant la détermination du caractère phénoménal des expériences conscientes par leur contenu représentationnel, débats focalisés autour de certains cas originellement présentés comme des contre-exemples au représentationnalisme (auxquels les représentationnalistes ont répondu) et dont les représentationnalistes ont pu sembler sortir vainqueur.

Mon but est le suivant : je veux montrer qu'il existe un moyen pour l'anti-représentationnaliste de refuser ces raisons d'accepter le représentationnalisme : c'est-à-dire qu'il existe un moyen de contrer ces arguments et d'éviter de considérer les débats que j'évoquerai comme ayant véritablement tourné à l'avantage du représentationnaliste. Plus précisément, je veux présenter deux arguments inspirés de « l'argument du langage privé » (qui correspond à un ensemble de remarques faites par Wittgenstein), qui me semblent tous deux convaincants. Je souhaite montrer que, si l'anti-représentationnaliste accepte la conclusion de ces arguments, c'est-à-dire, en première approximation, s'il accepte l'idée d'après laquelle les termes de notre langage descriptif, pour être doués de sens, doivent posséder des critères publics de correction et référer à des entités publiquement accessibles, alors il peut répondre assez naturellement aux arguments représentationnalistes que je vais présenter. Mieux, il se trouve aussi en position, je crois, de formuler des contre-exemples au représentationnalisme d'une manière qui n'autorise pas des réponses du type de celles qui ont été avancées jusque-là par les tenants du représentationnalisme.

Dans cet article, je procéderai de la manière suivante : dans une première partie, je dirai quelques mots du représentationnalisme concernant l'expérience consciente (en essayant notamment de distinguer plusieurs versions de celui-ci et de différencier celle qui va m'intéresser de ses principales voisines). Dans une deuxième partie, j'exposerai trois arguments de Dretske en faveur du représentationnalisme, que l'on peut rassembler en les nommant des « arguments d'individuation des expériences ». Dans une troisième partie, je parlerai des débats autour des contre-exemples potentiels au représentationnalisme et de la manière dont les représentationnalistes sont apparemment sortis vainqueurs de ces débats. Une quatrième partie sera consacrée à un exposé des arguments que je propose, qui sont des arguments inspirés de l'argument wittgensteinien du langage privé, ainsi qu'à une défense de ceux-ci. Enfin, dans une cinquième et dernière partie, j'essaierai de montrer pourquoi l'acceptation de la conclusion de ces arguments permet à l'anti-représentationnaliste de répondre de manière simple et naturelle au représentationnaliste, en contrant les arguments de Dretske et en réinterprétant les débats précédemment exposés.

Avant de commencer le développement de ce programme, je désire procéder à quelques précisions de vocabulaire. Dans cet article, j'utiliserai les termes « expérience consciente », « états conscients » et « conscience phénoménale » pour référer à la dimension vécue et qualitative de la vie mentale. Voir du rouge, goûter du chocolat, ressentir un mal de dents, constituent des exemples d'expériences conscientes, d'états conscients, et sont des cas de conscience phénoménale. J'entendrai par « caractère phénoménal » (par lequel je traduis l'anglais *phenomenal character*) et par « effet que cela fait » ce qui différencie les expériences phénoménales entre elles d'un point de vue phénoménal : une sensation de rouge et une sensation de bleu ne font pas le même effet, elles ne possèdent pas le même caractère phénoménal. Les termes de « propriétés phénoménales » ou « *qualia* »<sup>3</sup> désignent par hypothèse les propriétés qui qualifient le caractère phénoménal d'une

---

<sup>3</sup> « *Qualia* » est un terme particulièrement chargé et polysémique dans la philosophie contemporaine. Je l'emploie ici dans un sens neutre, selon lequel « *qualia* » et « propriétés phénoménales » sont synonymes. En ce sens, la grande majorité des théoriciens s'accordent sur l'existence des *qualia*, même s'ils débattent de leur statut métaphysique et de leur nature exacte. En un autre sens, plus controversé, qui justifie l'expression de « théorie des *qualia* » par opposition à « théorie représentationnaliste », les *qualia* désignent les propriétés

expérience consciente, de sorte que l'ensemble des propriétés phénoménales (ou *qualia*, pris en un sens neutre) portées par une expérience nous donne son caractère phénoména. Par « représentation », j'entendrai tout état ou objet qui vise intentionnellement un objet auquel il attribue certaines propriétés, d'une manière telle qu'il possède des conditions de corrections (une représentation doit pouvoir être évaluée sous l'angle de la correction, c'est-à-dire être correcte ou incorrecte). Le terme de « propriété représentationnelle » désigne toute propriété d'un état ou d'un objet en vertu duquel il est une représentation. Enfin, par « contenu représentationnel », je veux désigner les conditions de correction de la représentation en question, c'est-à-dire ce que la représentation représente (l'état du monde qui doit être le cas pour que la représentation soit correcte).

### 1. Le représentationnalisme à propos de l'expérience consciente

On peut comprendre la thèse du représentationnalisme à propos de l'expérience consciente de plusieurs manières. Dans sa version la plus faible, cette thèse est relativement partagée et consensuelle : elle revient à poser qu'il existe un lien entre conscience et représentation. Par-là, on signifie que la plupart, voire la totalité des expériences conscientes possèdent des propriétés représentationnelles – en d'autres termes, qu'elles représentent que quelque chose est le cas. Ce fait est particulièrement saillant dans le cas de la conscience perceptive : lorsque je vois, par exemple, un cercle rouge, j'ai une certaine expérience consciente, dotée de certaines propriétés phénoménales, mais dans le même temps mon état mental représente l'entité perçue et me fournit de l'information quant à l'existence du cercle rouge et aux propriétés de celui-ci ; autrement dit, mon état mental *vise* intentionnellement un certain objet, en l'occurrence un cercle rouge, et le représente correctement (ou incorrectement, dans le cas par exemple d'une perception déficiente, comme une illusion). Le cas des expériences conscientes qui ne nous présentent aucun objet extérieur (comme les expériences de douleur, ou les états proprioceptifs) est un peut-être un peu différent, car ces expériences ne semblent pas à première vue représenter un objet et nous fournir de l'information sur lui. Toutefois, même ces états mentaux peuvent se voir attribuer un aspect représentationnel, au sens où ils nous présentent notre corps comme étant dans un certain état (nous fournissant ainsi des informations à son propos).

Une première thèse représentationnaliste, appuyée sur ce genre d'observation, pourrait donc être formulée ainsi : tous les états mentaux qui sont des expériences conscientes possèdent un certain contenu représentationnel, au sens au moins où ils visent certains objets externes (objets de l'environnement) ou internes (le corps). L'on peut baptiser « représentationnalisme minimal » cette thèse assez populaire<sup>4</sup>.

---

phénoménales intrinsèques et non-représentationnelles qui fixent le caractère phénoména des expériences. Il n'est pas sûr alors qu'existent des *qualia* ; et en ce sens de *qualia* le représentationnaliste refuse leur existence. Enfin, en parlant de *qualia* on sous-entend parfois que ces propriétés sont non seulement intrinsèques et non-représentationnelles, mais également non-physiques. En ce sens encore plus spécifique, le théoricien des *qualia* est nécessairement un dualiste, ou au moins un anti-physicaliste. Je ferai en sorte de spécifier le sens de « *qualia* » que j'entends utiliser à chaque fois que je ferai usage de ce terme.

<sup>4</sup> Ceux qui s'opposent au représentationnalisme minimal s'appuient généralement sur des exemples d'états conscients dont il semble qu'on ne peut leur associer aucune présentation d'objet spécifique, comme l'angoisse ou l'euphorie. Il existe de nombreuses façons de défendre le représentationnalisme minimal contre ces objections : on peut ainsi arguer de ce que l'angoisse et l'euphorie présentent bien des objets, mais que ces objets sont en très grand nombre (l'euphorie, par exemple, présentent peut-être *tous* les objets de l'environnement du sujet comme étant marqués d'une valeur positive). On peut aussi choisir de dire que l'angoisse ou l'euphorie ne sont pas vraiment des expériences conscientes en tant que telles, mais simplement une manière de nommer un certain type de succession d'expériences conscientes de joie ou de malaise (chacune de ces expériences prise individuellement visant bien un objet), ou une disposition à entrer dans ce type de succession.

La plupart de ceux qui se revendiquent du représentationnalisme à propos de l'expérience consciente vont toutefois plus loin. Ils mettent en avant le fait que cette cooccurrence des propriétés phénoménales et représentationnelles dans nos états mentaux n'est pas arbitraire ou hasardeuse, mais qu'elle se produit de manière à la fois systématique et réglée. C'est à nouveau le cas de la conscience perceptive qui semble fournir les plus forts indices en faveur d'une telle conception : voir un cercle rouge, c'est en effet tout à la fois avoir une expérience d'un cercle rouge et représenter ce cercle rouge. Mais il semble que les deux aspects soient fortement liés : il est difficile, si ce n'est impossible, d'imaginer que nous puissions entrer dans un état mental qui contienne phénoménalement un carré vert et qui pourtant représente toujours un cercle rouge. Inversement, nous avons du mal à imaginer que notre état mental puisse inclure phénoménalement un cercle rouge, tout en représentant un carré vert.

Le caractère phénoménal de nos états perceptifs serait donc lié à leur contenu représentationnel, de sorte que les deux aspects co-varient : si je suis dans un état perceptif avec un caractère phénoménal et un contenu représentationnel donnés, il semble impossible que le caractère phénoménal de mon état ne change sans que ne change également son contenu représentationnel. Si, par exemple, je fais l'expérience phénoménale d'un objet doté d'une forme X, je ne peux pas imaginer que sa forme change d'un point de vue phénoménal sans que ne change du même coup la façon dont je représente l'objet, c'est-à-dire le type de propriétés que mon état mental attribue à l'objet<sup>5</sup>.

Cette co-variation entre contenu représentationnel et contenu phénoménal des états mentaux, si elle est très intuitive dans le cas des représentations visuelles, est nettement plus discutée pour le cas d'autres types d'états mentaux (dont j'examinerai certains exemples typiques dans la suite de cet article, lors du passage qui sera consacré à l'examen de certaines objections anti-représentationnalistes). C'est toutefois l'acceptation de cette co-variation qui permet de séparer ce que l'on peut nommer le « représentationnalisme faible » du simple représentationnalisme minimal décrit précédemment. Le représentationnalisme faible, que l'on peut aussi qualifier de « représentationnalisme de la survenance » ou de « représentationnalisme de la détermination », peut s'énoncer ainsi : le caractère phénoménal d'un état mental donné *survient* sur le contenu représentationnel de cet état. En d'autres mots : il ne peut pas y avoir de changement dans le caractère phénoménal d'un état sans changement dans ses propriétés représentationnelles ; une fois fixées les propriétés représentationnelles d'une expérience phénoménale donnée, son caractère phénoménal est également fixé<sup>6</sup>.

Enfin, l'on peut réserver le nom de « représentationnalisme fort » (ou « représentationnalisme constitutif ») à la thèse qui poursuit encore le rapprochement entre expérience et représentation et va jusqu'à *identifier* les propriétés phénoménales d'un état à ses propriétés représentationnelles, en affirmant que le caractère phénoménal d'un état est *constitué* par le contenu représentationnel de cet état (ou par son contenu représentationnel d'un certain type), et que le premier existe *en vertu* de l'existence du second (Dretske, 1995; Thau, 2002; Tye, 1995b). Dans cette perspective, la détermination des propriétés phénoménales d'un état par ses propriétés représentationnelles est tout simplement due au fait que le caractère phénoménal d'un

---

<sup>5</sup> Cet argument trouve une de ses formulations contemporaines les plus influentes chez Charles Siewert (Siewert, 1998). Siewert ne défend pas une position représentationnaliste au sens le plus courant du terme mais se rapproche plutôt des tenants de « l'intentionnalité phénoménale » que j'évoquerai dans la note suivante.

<sup>6</sup> Nombreux sont les auteurs qui, dans cette compréhension du représentationnalisme, peuvent être qualifiés de représentationnalistes (Byrne, 2001; Chalmers, 2004; Dretske, 1995; Harman, 1990; Lycan, 1996; Thau, 2002; Tye, 1995b). Les auteurs qui défendent l'existence d'un lien fort entre l'intentionnalité et l'expérience, mais d'après lesquels c'est l'expérience consciente qui doit avoir priorité dans l'analyse, et que l'on range généralement sous le qualificatif de partisans de « l'intentionnalité phénoménale » (Horgan & Tienson, 2002; Kriegel, 2013; Siewert, 1998), endossent également cette thèse ; en ce sens précis et restreint, on peut donc dire d'eux qu'ils sont des représentationnalistes, même s'ils se distinguent nettement, par exemple, du représentationnalisme fort que j'expose un peu plus bas.

état n'est rien d'autre et rien de plus que son contenu représentationnel ; et ses propriétés phénoménales ne sont rien d'autre que ses propriétés représentationnelles (ou que ses propriétés représentationnelles d'un certain type)<sup>7</sup>.

Dans cet article, je vais m'intéresser uniquement à la forme faible du représentationnalisme (en laissant de côté les formes fortes et minimales). Par « représentationnalisme », j'entendrai désormais le représentationnalisme faible (ou « représentationnalisme de la survenance »), qui constitue une thèse plus faible et beaucoup plus largement acceptée que le représentationnalisme fort. Mon but est de montrer que certaines raisons de croire au représentationnalisme ainsi compris, dont l'influence a été, je crois, particulièrement importante, et qui ont joué un rôle crucial dans les succès récents du représentationnalisme, sont en réalité des raisons défectueuses ; elles ne devraient pas forcément nous mener à adopter le représentationnalisme, et l'anti-représentationnaliste peut les contrer s'il accepte la conclusion de certaines versions de l'argument du langage privé. Je veux maintenant exposer ces raisons ; dans un premier temps, en exposant plusieurs arguments représentationnalistes fournis par Fred Dretske, et dans un second temps, en donnant quelques exemples de la manière dont les idées sous-tendant ces arguments ont pu structurer la dialectique de débats récents autour du représentationnalisme.

## 2. Les arguments d'individuation des expériences pour le représentationnalisme

Il existe de nombreux arguments en faveur du représentationnalisme, mais il me semble que parmi ceux-ci, une certaine famille d'arguments a joué un rôle spécifiquement important. On peut nommer cette famille d'arguments les **arguments d'individuation des expériences**, dans la mesure où ils s'appuient sur des réflexions ayant trait à la question de l'individuation des caractères phénoménaux des états conscients. On trouve un exposé explicite et clair d'arguments de ce type dans un article de Fred Dretske<sup>8</sup>. C'est sur les versions fournies par Dretske que je vais me concentrer ; plus précisément, sur trois arguments de cette famille qui sont présents dans ce texte.

Le premier argument de ce type présenté par Dretske est un argument direct en faveur du représentationnalisme faible. Le deuxième et le troisième, eux, se trouvent plutôt mis en avant comme des formes de défense du représentationnalisme contre une famille particulière d'objections. L'idée de base qui sous-tend ces trois arguments me semble toutefois assez semblable ; et si j'entreprends de les exposer tous, c'est pour plus de clarté dans l'exposition.

Au sein de cette famille des « arguments d'individuation » pour le représentationnalisme, je propose de nommer le premier argument, qui est le plus direct et le plus explicite, **l'argument de la**

---

<sup>7</sup> Il existe de nombreuses autres distinctions permettant de classer plus finement encore les différentes variétés du représentationnalisme à propos de l'expérience consciente. On distingue ainsi le représentationnalisme pur et impur, le représentationnalisme réductif et non-réductif ; on trouve aussi des oppositions sur la question du contenu représentationnel pertinent pour comprendre le caractère phénoménaux des expériences conscientes : est-ce un contenu conceptuel ou non-conceptuel ? Est-il déterminé de manière interne ou externe ? Toutes ces distinctions ont leur importance ; il faudrait par exemple réserver un traitement particulier à la forme *réductive* du représentationnalisme, qui doit reposer sur une sémantique naturaliste et promet une explication matérialiste de l'expérience consciente. Cette version du représentationnalisme est probablement la plus populaire de toutes et tient sa popularité précisément de son ambition matérialiste. Toutefois, pour des raisons de place, je ne traiterai pas ce genre de questions ici. Voir (Chalmers, 2004; Seager & Bourget, 2007) pour une revue et une discussion de ces différentes formes de représentationnalisme.

<sup>8</sup> Il existe un argument relativement voisin que l'on trouve également dans la littérature (Harman, 1990; Tye, 2002) en faveur du représentationnalisme, et qui s'appuie sur la thèse phénoménologique selon laquelle l'expérience est « transparente », c'est-à-dire que, dans l'expérience, nous sont présentées seulement les propriétés des objets externes perçus (et non d'hypothétiques propriétés intrinsèques de nos expériences conscientes elles-mêmes). Je pense que beaucoup de ce que je vais dire sur l'argument de l'individuation des propriétés formulé par Dretske pourrait être appliqué à l'argument de la transparence, avec quelques adaptations. Toutefois, je vais me concentrer ici sur l'argument de Dretske.

**classification des expériences.** Cet argument s'appuie sur une remarque : les propriétés dont nous nous servons pour classer nos expériences phénoménales (et leur caractère phénoménales) sont les propriétés des objets extérieurs auxquels nous associons typiquement ces expériences. Nous différencions ainsi deux états phénoménaux distincts en évoquant les propriétés de leurs objets : on parle ainsi d'une expérience *de* rouge ou d'une expérience *de* vert, de la même manière que l'on individue des représentations conventionnelles et matérielles en fonction des objets qu'elles représentent. L'idée de Dretske est que la meilleure explication de ce fait est que la thèse représentationnaliste est vraie ; en ce sens, cet argument en faveur du représentationnalisme se veut une inférence à la meilleure explication.

Voilà la façon dont Dretske présente son argument :

« Nous distinguons les expériences, non pas selon leurs propriétés, mais selon les propriétés possédées par leurs objets (s'il existe de tels objets). Cela ressemble remarquablement à un mode de classification représentationnel, c'est-à-dire que cela ressemble remarquablement à la façon dont nous classons, disons, les récits et les photographies. A est une biographie d'Oscar Wilde, B est une histoire de la Guerre d'Espagne [...] C est une photographie d'un magnat turc, D est une image du maire de Chicago. Ce sont des images différentes [...], non pas parce qu'elles sont en elles-mêmes si différentes (elles pourraient être impossibles à distinguer ; le maire pourrait être le frère jumeau du magnat), mais parce que ce qu'elles représentent, ce dont elles sont les images, est si différent.

Une fois que nous remarquons que les propriétés dont nous nous servons pour classer les expériences subjectives ne sont pas les propriétés que ces expériences possèdent, mais les propriétés publiquement accessibles que les objets externes possèdent (ces objets qui sont normalement les objets des expériences), il devient, il me semble, irrésistible d'expliquer cette procédure peu usuelle de classification en termes représentationnels. Les expériences sont différentes parce que les objets dont elles sont (normalement) les expériences diffèrent de ces manières expérientiellement détectables » (Dretske, 2003, traduction par l'auteur du présent article).

Ainsi formulé, cet argument se veut un argument direct en faveur du représentationnalisme, et l'on pourrait le reformuler de la façon suivante: la première prémisse de l'argument consiste en la thèse selon laquelle nous classons nos expériences en faisant référence aux propriétés des objets auxquels nous associons ces expériences (et que nous décrivons comme les objets de ces expériences). La seconde prémisse est la thèse selon laquelle « l'explication irrésistible » (c'est-à-dire, on peut le supposer, la seule explication raisonnable et naturelle) du fait que nous classons nos expériences par les propriétés des objets auxquels nous associons ces expériences est que nos expériences sont différentes, en tant qu'expériences (c'est-à-dire sous le rapport de leur caractère phénoménales) parce qu'elles diffèrent sous le rapport de leur contenu représentationnel. La conclusion est que nos expériences diffèrent sous le rapport de leur caractère phénoménales parce qu'elles diffèrent sous celui de leur contenu représentationnel (c'est-à-dire que leur contenu représentationnel détermine leur caractère phénoménales)<sup>9</sup>.

On trouve, plus loin dans le même article, deux arguments qui s'approchent suffisamment de celui que je viens d'exposer pour pouvoir être classés dans une même catégorie. Ces deux arguments sont, eux, plus précisément dirigés *contre* l'éventualité de contre-exemples ambitionnant de montrer

---

<sup>9</sup> Il serait, je crois, possible d'interpréter cet argument de Dretske comme un argument pour le représentationnalisme fort. Ce n'est pas mon projet ici ; je veux le présenter (pour pouvoir ensuite le critiquer) uniquement en tant qu'il soutient la thèse plus faible du représentationnalisme de la survenance (par ailleurs impliqué par ce que j'ai nommé le représentationnalisme fort).

la fausseté du représentationnalisme par la mise en avant de cas d'états mentaux dont les aspects phénoménaux ne peuvent pas (ou pas complètement) être décrits de manière représentationnelle.

Voilà la façon dont Dretske formule ces deux arguments :

« Je me tourne maintenant plutôt vers l'aspect de cette théorie qui a probablement attiré le plus d'attention critique. Il s'agit de l'idée selon laquelle la façon dont les choses apparaissent phénoménalement est complètement donnée par le caractère représentationnel de l'expérience. [Cette] affirmation a attiré un torrent de contre-exemples. L'argument n'est pas alors qu'il existe des qualités phénoménales dont il s'avère que le système nerveux humain ne peut pas les représenter. Car cela constituerait un argument portant sur les capacités représentationnelles des systèmes biologique [...] L'argument est plutôt qu'il y a certaines qualités dont nous faisons l'expérience qui ne peuvent pas être comprises comme les qualités d'un objet, qui serait représenté comme ayant ces qualités [...]. L'argument procède ainsi : lorsque nous avons une expérience de type E, les choses nous apparaissent F ; mais rien en nous ne peut représenter qu'un objet a la propriété F ; donc, les expériences de type E ne sont pas, contrairement à ce que dit le représentationnalisme, complètement déterminées par leurs propriétés représentationnelles.

Une fois que l'on a ainsi décrit la situation dialectique [...] il est clair qu'il est difficile de trouver des arguments convaincants de cette forme. Comment peut-on trouver une propriété telle que rien, en nous, ne peut représenter que quelque chose la possède ? Est-ce que l'acte même de spécifier cette propriété ne constitue pas une manière de la représenter ? Est-ce qu'une propriété telle que rien ne pourrait représenter que quelque chose la possède ne serait pas, de ce fait, absolument indétectable ? [...] Il semble donc que des contre-exemples à une théorie représentationnelle de l'expérience seront contraints de faire appel à des propriétés indétectables, qui sont les seules dont cette théorie ne peut pas rendre compte » (Dretske, 2003, traduction par l'auteur du présent article)

La réflexion de Dretske se présente donc comme fournissant des arguments contre des attaques formulées à l'encontre du représentationnalisme faible (qui n'est qu'une partie de sa théorie, mais équivaut à ce que j'ai décidé d'appeler « représentationnalisme »). L'idée de Dretske peut, il me semble, se résumer ainsi : pour que le philosophe anti-représentationnaliste soit en mesure de fournir des contre-exemple *philosophiques* (et non empiriques et basés sur les capacités représentationnelles de ces systèmes biologiques que sont les cerveaux humains) à la thèse d'après laquelle le caractère phénoménaux d'un état mental est déterminé entièrement par son contenu représentationnel, il faudrait qu'il soit en mesure de trouver une expérience E telle que cette expérience présente un aspect phénoménaux qui ne peut trouver d'équivalent dans le contenu représentationnel porté par cet état. Autrement dit, il faudrait que l'anti-représentationnaliste soit capable de donner l'exemple d'un type d'expérience E tel que, dans une expérience de ce type, les choses nous apparaissent phénoménalement d'une certaine façon F, sans pour autant que cette expérience puisse nous représenter un objet comme ayant cette propriété F (dans la mesure où l'instanciation de ce genre de propriété est impossible à représenter).

Or, nous dit Dretske, cela semble extrêmement problématique, pour deux raisons. D'une part, il semble que le fait même de spécifier, dans le langage, une propriété F (dont on se demande si elle est possible à représenter ou non), constitue une façon de représenter cette propriété F. En ce sens, on voit mal comment l'anti-représentationnaliste pourrait parvenir à spécifier une propriété F qui soit impossible à représenter. Par ailleurs, il semble que les seules propriétés capables de jouer le rôle de « F » seraient des propriétés indétectables. En effet, pour toute propriété F qui qualifie la manière dont les choses nous apparaissent phénoménalement, s'il est possible de détecter cette propriété, alors il est en droit possible de la représenter, et rien ne nous permet *a priori* d'affirmer

que nous ne pouvons pas la représenter comme étant instanciée<sup>10</sup>. Or, même si Dretske ne le dit pas explicitement, il semble que la nécessité de trouver une expérience capable de nous présenter des propriétés indétectables soit si problématique qu'elle constitue presque une *reductio ad absurdum* pour l'anti-représentationnaliste<sup>11</sup>. Les difficultés de principe semblent en effet très importantes : devons-nous seulement admettre l'existence de propriétés indétectables ? Et si oui, comment pourrions-nous avoir le moindre rapport à elles, même dans l'expérience ?

On peut donc reformuler le propos de Dretske en le séparant en deux arguments, qui peuvent être explicités de la façon suivante. Le premier peut être nommé **l'argument de la spécification des propriétés**. Sa première prémisse consiste en la thèse selon laquelle, si l'anti-représentationnaliste veut fournir un contre-exemple philosophique au représentationnalisme, alors il doit décrire un cas d'expérience E dans laquelle les choses apparaissent F, où F est une propriété impossible à représenter comme étant instanciée. La deuxième prémisse consiste à dire que, pour décrire un tel cas, l'anti-représentationnaliste doit spécifier la propriété F, impossible à représenter, dont il est question. La troisième prémisse est l'idée d'après laquelle spécifier une propriété revient à la représenter. La conclusion est que l'anti-représentationnaliste ne peut pas fournir un contre-exemple *philosophique* au représentationnalisme, dans la mesure où cela l'obligerait à représenter une propriété irréprésentable – ce qui est impossible.

Le second sous-argument peut être nommé **l'argument de la détection des propriétés**. Sa première prémisse est la même que celle de l'argument précédent, et consiste en la thèse d'après laquelle, si l'anti-représentationnaliste veut fournir un contre-exemple philosophique au représentationnalisme, alors il doit décrire un cas d'expérience E dans laquelle les choses apparaissent F, où F est une propriété impossible à représenter comme étant instanciée. La deuxième prémisse consiste à dire que si une propriété est impossible à représenter comme étant instanciée, alors cette propriété est indétectable. Une troisième prémisse (implicite dans le texte de Dretske) réside dans l'idée d'après laquelle l'anti-représentationnaliste n'a pas le droit d'en appeler à des propriétés indétectables, en raison des difficultés que poserait un tel appel. La conclusion est que l'anti-représentationnaliste ne peut pas trouver un contre-exemple philosophique au représentationnalisme.

### 3. L'utilisation implicite d'arguments de ce type dans des débats récents

J'ai décrit un ensemble constitué de trois arguments utilisés par Dretske en faveur du représentationnalisme, qui me semblent manifester une certaine ressemblance : tous ces arguments mettent en avant l'idée selon laquelle l'individuation de nos expériences exige la description des propriétés qui qualifient ce qui nous apparaît dans l'expérience. Or, les propriétés que nous décrivons ainsi sont en règle générale les propriétés d'objets extérieurs, nous dit l'argument de la classification des expériences, ce qui nous donne une raison de penser que nos expériences se différencient par la manière dont elles représentent ces objets extérieurs. Par ailleurs, d'après l'argument de la spécification, le fait que nous décrivons ces propriétés implique que nous représentions ces propriétés (au moins dans le moment de leur description), ce qui rend impossible que ces propriétés, qui qualifient la phénoménalité de nos expériences, soient impossibles à représenter – empêchant ainsi l'anti-représentationnaliste de fournir les contre-exemples qu'il

---

<sup>10</sup> Il se peut, d'après Dretske, que notre système nerveux ne puisse pas représenter l'instanciation de certaines propriétés, mais ce genre d'impossibilité serait dû à notre constitution biologique particulière, dont la nature ne peut être établie qu'empiriquement. Ce n'est pas de ce genre d'impossibilité que parlent les philosophes auxquels s'attaque Dretske.

<sup>11</sup> Par ailleurs, Dretske se situe dans un cadre physicaliste, dans lequel on suppose que l'ensemble des propriétés instanciées sont des propriétés physiques (ou réalisées par des propriétés physiques), et donc détectables. Il faut bien sûr noter que tous les représentationnalistes ne sont pas physicalistes, et qu'en ce sens l'évacuation des propriétés indétectables ne serait pas acceptée aussi facilement par tous les partisans du représentationnalisme. Je reviendrai sur ce point dans la suite de ce travail.



recherche. Enfin, sauf à supposer que nous nous autorisons à décrire ces propriétés qui qualifient l'apparaître expérientiel comme des propriétés indétectables – ce que Dretske récuse implicitement – nous sommes forcés d'accepter que ces propriétés sont représentables, ce qui empêche également l'anti-représentationnaliste de fournir ses contre-exemples.

Je veux maintenant montrer, à l'aide quelques exemples, de quelle manière les idées qui sous-tendent les arguments défendus par Dretske, que je viens d'explicitier, ont été utilisées à plusieurs reprises dans les débats autour de la détermination du caractère phénoménal des états conscients par leur contenu représentationnel. En d'autres termes, je veux montrer que les idées exposées par Dretske ont joué un rôle important dans la structure de nombreux débats sur le représentationnalisme qui se sont focalisés sur des études de cas précis (le « torrent de contre-exemples » dont parlait Dretske).

Les débats dont je veux maintenant parler sont des débats qui ont pris pour objet des états mentaux dont il semblait qu'ils pouvaient constituer des contre-exemples actuels (par opposition à des contre-exemples imaginaires, proposés dans des expériences de pensées comme celle du spectre inversé ou celle des zombies) aux thèses représentationnalistes<sup>12</sup>.

Certains contre-exemples reposaient sur des états mentaux semblant échapper à l'analyse représentationnaliste, dans la mesure où ils présentaient un caractère phénoménal dont il apparaissait qu'il était impossible de l'associer à un contenu représentationnel qui lui correspondrait et le déterminerait.

Un des cas les plus célèbres de cette famille de contre-exemples a été proposé par Ned Block. Selon Ned Block, ainsi, on peut dire de l'expérience qui correspond à un plaisir intense comme celui de l'orgasme que son caractère phénoménal n'est pas épuisé par son contenu représentationnel : l'orgasme produit un certain effet phénoménal dont on ne peut pas dire qu'il puisse entièrement être capturé par le fait de représenter qu'*un orgasme est en train de se produire*. Ce fait peut être mis en lumière et appuyé sur un contraste : je peux être dans un état mental qui représente une situation dans laquelle *quelqu'un d'autre que moi* est en train d'avoir un orgasme sans pour autant que mon état mental porte ce caractère phénoménal particulièrement prégnant qui est celui de mes propres orgasmes (Block, 1995, 2003) ; cela semble bien montrer qu'il y a un aspect phénoménal de l'orgasme qui ne dépend pas seulement du contenu représentationnel de ce type d'état.

D'autres contre-exemples proposés par les anti-représentationnalistes se sont basés plus spécifiquement sur des contrastes de ce type, c'est-à-dire sur la comparaison entre deux états mentaux qui semblent pouvoir se voir attribuer le même contenu représentationnel, mais des contenus phénoménaux différents – contredisant par là le représentationnalisme<sup>13</sup>.

Christopher Peacocke (Peacocke, 1983, 1992), ainsi, a mis en avant l'idée d'après laquelle deux perceptions visuelles perspectives du même objet et selon le même angle, mais effectuées à des distances différentes, possédaient le même contenu représentationnel, sans pour autant correspondre au même vécu phénoménal. Ces deux expériences, en effet, représentent le même arbre, doté des mêmes dimensions, placé dans la même position ; toutefois le caractère phénoménal des deux états doit être distingué, dans la mesure où, dans l'expérience visuelle de l'objet que je fais lorsque je suis plus éloigné de celui-ci, la zone du champ visuel occupée par l'objet se trouve diminuée.

Enfin, Ned Block (Block, 1995, 2003) a évoqué le cas de deux états mentaux qui partagent le même contenu représentationnel, mais diffèrent par la modalité sensorielle qui les caractérise et, pour cette raison, du point de vue phénoménal : deux perceptions peuvent me représenter la même chose, par exemple « qu'il y a quelque chose au-dessus de ma tête », mais de deux manières phénoménalement distinctes, car sensoriellement différentes – par exemple, selon une modalité visuelle ou auditive.

---

<sup>12</sup> Sur cette distinction entre plusieurs types de contre-exemples, voir l'article d'Alex Byrne (Byrne, 2001).

<sup>13</sup> Pour la revue de ces cas, je m'appuie sur un chapitre de William Seager et David Bourget (Seager & Bourget, 2007).

Face à ces défis anti-représentationnalistes<sup>14</sup>, la réponse des représentationnalistes a suivi un motif toujours identique : ils ont tenté de montrer la présence, dans chacun des cas concernés, de contenus représentationnels négligés (ou de différences représentationnelles négligées) qui permettaient de rendre compte des caractères phénoménaux en question (ou des différences de caractères phénoménaux) dans le respect de la conception représentationnaliste. C'est-à-dire que les représentationnalistes ont dû relever le « défi de la caractérisation supplémentaire » (Byrne, 2001; Thalabard, 2013; Wu, 2011) : ils ont dû exhiber, pour chaque cas, des différences représentationnelles explicatives des différences phénoménales.

Ainsi, en ce qui concerne le cas de l'orgasme, on a pu dire que le contenu représentationnel de mon expérience orgasmique est en partie que « quelque chose de très plaisant est en train d'arriver là, en bas » (« *something very pleasing is happening down there* ») (Tye, 1995a, 1996a). On peut également ajouter que mon expérience représente que ce caractère plaisant augmente et diminue en intensité au cours du temps. Ceci posé, on peut analyser la différence phénoménologique entre l'épreuve d'un de mes orgasmes et la perception de l'orgasme de quelqu'un d'autre par le fait que le second état mental ne représente pas, en tant que tel, l'advenue de quelque chose d'aussi plaisant (en plus de localiser cette advenue en un endroit différent du monde). Intégrer des propriétés relevant de l'évaluation dans le domaine des propriétés que mes états mentaux peuvent représenter semble donc permettre de résoudre le problème posé au représentationnaliste par ce genre de cas.

Dans le cas des contrastes entre états mentaux, des réponses d'un style identique peuvent être trouvées. Par exemple, on peut dire des expériences visuelles et auditives qui me représentent « qu'il y a quelque chose au-dessus de ma tête » qu'elles présentent un certain nombre de différences représentationnelles qui peuvent rendre compte de leurs différences phénoménologiques : une expérience visuelle représente toujours, si ce n'est une couleur, du moins une luminosité et une teinte ; tandis que l'audition représente des propriétés des sons comme la hauteur et le timbre (Lycan, 1996). Il y a bien quelque chose de *commun* au contenu représentationnel des deux états décrits, mais il serait exagéré de les considérer comme ayant un contenu représentationnel identique ; et la mise en lumière de ces différences représentationnelles permet de comprendre pourquoi, dans un cadre représentationnaliste, ces deux expériences ont un caractère phénoménal différent.

De même, pour reprendre le contre-exemple proposé par Peacocke, on peut répondre qu'il est faux de dire que deux expériences d'un même objet et selon le même angle, mais vu à des distances différentes, ont le même contenu représentationnel mais des caractères phénoménaux différents. En effet (Tye, 1996b), les deux expériences représentent l'objet comme étant situé à des *distances différentes* du sujet ; de plus, la *résolution* avec laquelle les objets sont représentés dans chaque cas varie (c'est-à-dire que les deux expériences ne représentent pas leurs objets comme ayant exactement les mêmes propriétés, dans la mesure où l'expérience correspondant à la vision « de près » lui attribue des propriétés spatiales plus précisément déterminées que l'autre expérience). Cette fois, c'est notamment la prise en compte de propriétés relationnelles dans le champ de ce qui peut être représenté, ainsi que la mise en avant des différences de finesse de grain entre les représentations, qui permettent au représentationnaliste de formuler sa réponse.

Il est possible d'attribuer une structure dialectique commune à toutes ces réponses. Tout se passe comme si, face aux prétentions des anti-représentationnalistes, qui déclarent avoir découvert des cas de caractères phénoménaux impossibles à comprendre comme étant déterminés par le contenu représentationnel des états conscients, ou des différences phénoménales qui ne correspondent pas à des différences représentationnelles entre ces états, les représentationnalistes répondaient d'abord en s'intéressant à la spécification de ces différences. Qu'a l'orgasme de si particulier ? Il est plaisant. Qu'est-ce qui sépare l'expérience du même objet vu à dix mètres et à trente mètres ? La distance et la précision de la vision. En quoi se distinguent la vision et l'audition de

---

<sup>14</sup> De nombreux autres exemples pourraient être cités, mais je me limite à ceux-ci pour des raisons de place.

quelque chose au-dessus de moi ? Dans un cas, j'entends de l'audible (un son) et dans l'autre je vois du visible.

Une fois cette spécification engagée, il devient possible de postuler une propriété représentée qui rende compte de cette distinction, dans la mesure où elle correspond à cette spécification. Par exemple, on supposera que l'orgasme est plaisant parce qu'il consiste à représenter que quelque chose de plaisant arrive. Des perceptions visuelles effectuées à distance différentes se distinguent phénoménologiquement dans la mesure où chaque expérience représente les objets comme étant plus ou moins distants, et de manière plus ou moins précise. Enfin, les modalités sensorielles donnent lieu à des expériences différentes car elles ne représentent pas les mêmes propriétés comme étant instanciées (le visible a une luminosité et une teinte, l'audible a un timbre et une hauteur).

Ces réponses représentationnalistes correspondent donc au modèle que proposait Dretske dans ses arguments : c'est en exigeant l'individuation des caractères phénoménaux censés n'être pas déterminables par un contenu représentationnel que le représentationnaliste obtient des propriétés comme la distance ou le caractère plaisant. Libre à lui de supposer alors que les états concernés (l'orgasme, la perception visuelle perspective) comprennent, dans leur contenu représentationnel, la représentation de ces propriétés, ce qui lui permet alors de répondre au représentationnaliste. Et l'anti-représentationnaliste aurait bien des difficultés à établir, sur une base *a priori*, que ces propriétés ne *peuvent pas* être représentées, car (a) L'acte même de les spécifier (dans le moment où l'on individue les caractères phénoménaux) est déjà une façon de les représenter de (b) Si ces propriétés sont détectables, elles peuvent être représentées (et il serait très étrange qu'elles soient indétectables).

La plupart des auteurs s'accordent à dire que les représentationnalistes sont sortis vainqueurs de ces débats (Byrne, 2001; Chalmers, 2004; Speaks, 2010); et l'ensemble de ces réponses aux contre-exemples putatifs ont joué, il me semble, un rôle psychologique et historique important dans le succès du représentationnalisme à propos de l'expérience consciente. En montrant que le représentationnalisme pouvait répondre de manière acceptable aux cas apparemment les plus problématiques pour lui, les représentationnalistes ont fourni ce qui pouvait apparaître comme des *confirmations* de la théorie.

Ce que je veux toutefois montrer, c'est que l'anti-représentationnaliste, pour peu qu'il accepte la conclusion de certains arguments inspirés de l'argument du langage privé de Wittgenstein, qui reposent sur l'idée d'après laquelle les termes de notre langage descriptif doivent, pour pouvoir être utilisés d'une manière douée de sens, posséder des critères publics d'application, est en mesure tout à la fois de résister aux arguments de Dretske et de montrer que l'apparent succès des réponses des représentationnalistes aux contre-exemples qu'on leur a adressés ne constitue en rien une confirmation de leurs conceptions.

#### **4. Le langage privé et l'anti-représentationnalisme**

Mon but est à présent d'argumenter en faveur de deux idées. D'une part, je veux présenter des arguments, inspirés de l'argument wittgensteinien du langage privé, en faveur de l'idée d'après laquelle les termes de notre langage, pour pouvoir être doués de sens, doivent posséder des critères publics d'application, et référer pour cela à des entités publiquement accessibles. D'autre part, je veux montrer que si l'anti-représentationnaliste accepte les conclusions de ces arguments, il acquiert une méthode pour tout à la fois déminer les arguments représentationnalistes de Dretske, et arguer de ce que les apparents succès représentationnalistes (dans leurs réponses aux contre-exemples des anti-représentationnalistes) ne constituent en rien une confirmation de la théorie représentationnaliste.

Ce que l'on nomme classiquement « l'argument du langage privé » correspond à un ensemble de réflexions proposées par Wittgenstein dans les *Recherches Philosophiques* (Wittgenstein, 2004), qui prennent place essentiellement aux paragraphes 244-271 (le paragraphe 258 étant sans doute le plus discuté) des *Recherches*. L'interprétation exacte de cet argument est

l'objet de débats très nombreux et très ouverts (voir Candlish & Wrisley, 2012 pour une première approche de ces débats), et il me semble extrêmement risqué de se prononcer quant à l'interprétation correcte de l'argument, ou sur ce que Wittgenstein avait exactement l'intention d'affirmer en proposant ces remarques. Pour éviter ces écueils interprétatifs (inévitables en histoire de la philosophie, et plus particulièrement lorsqu'il s'agit d'un auteur tel que Wittgenstein), je veux ici présenter deux arguments simplement *inspirés* par l'argument du langage privé, dont je n'affirme pas qu'ils correspondent à ce que Wittgenstein a voulu affirmer, mais qui me semblent suffisamment convaincants en eux-mêmes pour être pris en compte.

Je nomme le premier de ces arguments « l'argument de la privauté de droit », et le second argument « l'argument de la privauté de fait ». Le deuxième argument s'éloigne plus de la lettre wittgensteinienne ; j'espère toutefois parvenir à montrer pourquoi les deux arguments me semblent acceptables. Le premier argument vise à montrer que si une entité est privée « de droit », c'est-à-dire si elle est d'une nature intrinsèquement subjective et privée, telle qu'un seul sujet peut y avoir accès (telle que le seraient, par exemple, les *qualia* subjectifs et non-physiques dont l'existence est supposée par certains philosophes), alors les termes de notre langage descriptif ne peuvent pas y faire référence. Le second argument appuie l'idée d'après laquelle si une entité est privée « de fait », c'est-à-dire si elle est d'une nature intrinsèquement objective **mais** que la configuration des faits est telle que, de manière systématique (mais pas nécessairement interchangeable), seul un sujet peut y avoir accès (telles que le seraient, par exemple, des propriétés intrinsèques de certains états cérébraux, accessibles seulement par introspection), alors les termes de notre langage descriptif élaborés **sur la base de cet accès privé** ne peuvent pas y faire référence<sup>15</sup>. Le premier des arguments vise à armer, contre le représentationnalisme, le théoricien anti-représentationnaliste des *qualia* non-physiques ; le second intéressera le théoricien anti-représentationnaliste des *qualia* physiques (comme propriétés intrinsèques et non-représentationnelles de nos états cérébraux). Voilà les arguments.

**L'argument de la privauté de droit.** La première prémisse (P1) de l'argument est la suivante : si des termes de notre langage pouvaient référer à des entités privées *de droit*, c'est-à-dire à des entités *par principe accessibles à un seul sujet* (dont le mode d'existence est intrinsèquement subjectif et privé), alors il serait impossible d'associer à ces termes des critères publics de correction. Par « critère public de correction », on entend un ensemble d'observations que doivent nécessairement pouvoir faire les autres sujets pour que l'usage du terme soit correct. Par exemple, à la phrase « il y a un chien ici » sont associés des critères publics de correction, dans la mesure où d'autres sujets doivent pouvoir faire certaines observations<sup>16</sup> (par exemple observer la présence d'un animal doté de telles et telles propriétés sensibles et capacités typiques, dans l'environnement du locuteur) pour que l'usage du terme soit correct. Or, s'il existe des entités purement privées *de droit* (par exemple, certaines propriétés qualitatives, intrinsèquement subjectives, de nos expériences conscientes), alors on ne voit pas quels critères publics de correction pourraient être associés à l'usage des termes qui y réfèrent. « Je fais l'expérience d'une tache rouge en face de moi » ne nécessite pas, pour être prononcé correctement, que d'autres sujets observent qu'effectivement une tache rouge se trouve devant moi. Et si mon expérience est vraiment une entité purement privée, alors il n'existe en réalité *aucun* fait publiquement accessible qui pourrait être observé, et dont l'observation pourrait servir de critère de correction à mon occurrence de cette phrase.

La deuxième prémisse (P2) de l'argument est la suivante : si l'on admet qu'il est impossible d'associer à un terme des critères publics de correction, alors il n'est plus possible de distinguer un usage correct du terme par un sujet et un usage du terme qui *semble simplement correct* au sujet. En effet, si seul ce qu'observe le sujet, ce à quoi il accède, permet de déterminer la correction d'un usage, alors toute situation dans laquelle le sujet observe que son usage est correct, ou a l'impression que son usage est correct, sera *ipso facto* une situation d'usage correct.

---

<sup>15</sup> Avec certaines spécifications qui seront détaillées.

<sup>16</sup> « Observation » étant ici pris en un sens non-factif.

*La troisième prémisse (P3) de l'argument se formule ainsi : le fait, pour un terme, d'avoir un usage correct différenciable d'un usage qui semble simplement correct sans l'être au sujet qui l'emploie, est une condition nécessaire pour que ce terme ait un sens dans un langage descriptif. Cette prémisse est plausible si l'on présuppose que la possibilité de l'erreur est une condition nécessaire pour que quelque chose puisse être une description douée de sens. Une description est un acte linguistique qui possède, pour utiliser le vocabulaire de John Searle (Searle, 1983), une direction d'ajustement des mots vers le monde, au sens où notre usage des termes doit se modeler sur l'objet qu'il décrit, et au sens où notre description est incorrecte si elle échoue à se conformer à cet objet. Une description qui ne pourrait pas échouer, qui serait telle que l'ajustement serait par principe effectué à partir du moment où il est tenté, ne pourrait plus être dite avoir cette direction d'ajustement, par exemple, plutôt que l'autre (du monde vers les mots), ou qu'aucune d'entre elles ; elle ne satisferait plus en réalité aux règles du jeu de langage descriptif. Un terme à propos duquel il n'existe pas de différence entre un usage correct et un usage qui semble correct à l'utilisateur ne peut donc pas être un terme doué de sens dans un langage descriptif<sup>17</sup>.*

*La conclusion de l'argument de la privauté de droit est qu'il ne peut pas exister de termes doués de sens, appartenant à un langage descriptif, qui permettent de référer à des entités purement privées (et privées en droit : dont le mode d'existence est tel qu'elles sont purement subjectives, accessibles seulement à un seul sujet). Ce qu'il est important de noter, c'est que l'acceptation de la conclusion de cet argument n'implique pas de nier l'existence et la réalité d'entités privées, auxquelles un sujet peut avoir un certain accès (à condition que l'on admette qu'une telle relation « d'accessibilité » ne peut pas s'analyser comme une prise de connaissance permettant la description des entités concernées). La seule conclusion est que, si de telles entités privées existent, et donc par exemple si des *qualia* non-physiques et intrinsèquement subjectifs existent, il est impossible de nous servir de notre langage dans sa fonction descriptive pour en détailler les aspects. Il semble par ailleurs qu'une telle position puisse être interprétée comme étant celle de Wittgenstein qui, dans certaines remarques des *Recherches*, met en avant le fait que l'impossibilité d'un langage décrivant des entités purement privées n'implique pas l'inexistence ou l'irréalité de ces mêmes entités<sup>18</sup>.*

**L'argument de la privauté de fait.** Ce deuxième argument, également inspiré des remarques wittgensteiniennes sur le langage privé, s'écarte toutefois plus de la lettre de Wittgenstein. Il vise à établir que nous ne pouvons pas utiliser, dans notre langage descriptif, de termes doués de sens permettant de référer à des entités privées *de fait* que nous aurions élaborés sur la base de cet accès privé. Des entités privées de fait sont des entités telles qu'elles sont, de fait (et de manière systématique, quoique non nécessaire ; cela peut, par exemple, provenir de certains aspects contingents de la structure matérielle du monde), accessibles à un seul sujet.

*La première prémisse (P1\*) de l'argument est la suivante : si des termes de notre langage descriptif pouvaient référer à des entités privées de fait sur la base de cet accès privé, alors il serait impossible d'associer à ces termes des critères publics de correction. Imaginons en effet qu'une entité donnée soit, de fait, et de manière systématique, seulement observable par un sujet S (même s'il s'agit en droit d'une entité dont le statut intrinsèque est objectif et donc *public*). Rentreraient dans cette catégorie, par exemple, les propriétés physiques de nos états cérébraux que l'on supposerait accessibles par introspection mais qui seraient cachées à l'observation d'autrui en raison de leur situation physique, à l'intérieur de la boîte crânienne<sup>19</sup>. Serait également concerné le*

---

<sup>17</sup> Ce terme pourrait par contre, dans un vocabulaire wittgensteinien, avoir une fonction expressive.

<sup>18</sup> On peut par exemple citer le paragraphe 304 des *Recherches* (Wittgenstein, 2004) à l'appui : « 'Mais tu admettras tout de même qu'il y a une différence entre un comportement de douleur accompagné de douleur et un comportement de douleur en l'absence de douleur.' – L'admettre? Pourrait-il y avoir une différence plus grande! – 'Et pourtant tu en reviens toujours à ce résultat : La sensation elle-même est un rien' – Certainement pas. Elle n'est pas un quelque chose, mais elle n'est pas non plus un rien! ».

<sup>19</sup> Je mets de côté ici les conséquences des techniques d'imagerie cérébrale, que je discuterai un peu plus loin.

« scarabée » qui, dans la situation imaginée par Wittgenstein<sup>20</sup>, se trouve dans une boîte, et que seul le possesseur de la boîte est autorisé, de manière systématique, à observer. « J'ai une expérience de rouge » (qui référerait par hypothèse à une propriété intrinsèque d'un état cérébral) ou « Mon scarabée est plus gros qu'hier » ne peuvent se voir associés à aucune observations que doivent pouvoir faire les autres sujets afin que ces énoncés soient corrects, étant donné que par hypothèse les autres sujets ne *peuvent pas* faire d'observations à propos de ces entités.

*La deuxième prémisse (P2\*)* de l'argument est identique à P2.

*La troisième prémisse (P3\*)* de l'argument est identique à P3.

*La première conclusion de l'argument de la privauté de fait* est la suivante : il ne peut pas exister de termes doués de sens doués de sens, appartenant à un langage descriptif, qui permette de référer à des entités privées de fait *sur la base de cet accès privé*.

Cette conclusion ne signifie pas bien sûr qu'*aucun terme* de notre langage descriptif ne puisse référer à ces entités. Par exemple, il existe sans aucun doute des descriptions physiques (du type : « l'état cérébral doté de telles propriétés physiques » pour les propriétés intrinsèques de nos états cérébraux ; ou, pour le scarabée de Wittgenstein : « la créature dotée de telle propriété biologique ») qui se trouvent référer à ces entités ; mais il est impossible d'utiliser ces séquences linguistiques au sein d'une entreprise de description de ces entités qui s'appuierait sur cet accès privé. Et comme ces entités sont privées *de fait*, cela signifie qu'il n'existe aucune forme d'accès « public » (d'accès disponible pour les autres sujets) qui permettrait de les atteindre : de telles descriptions physiques ne peuvent donc être utilisées sur aucune base observationnelle. Certes, on pourra toujours manipuler ces descriptions physiques, et il se peut que je décrive mon état cérébral actuel par un terme physique et que, « tombant juste » par un coup du sort, j'arrive à y référer correctement. Mais je ne pourrai pas utiliser ces termes dans une véritable entreprise descriptive, au cours de laquelle je serai capable d'associer des critères publics de correction à l'usage de mes termes et où je serai en mesure (1) D'atteindre, lors de l'utilisation de ces termes dans des énoncés, la vérité plus souvent que ce que permettrait la simple chance ; (2) D'être « corrigible » lorsque je fais une erreur dans les énoncés mobilisant ces termes.

Une objection pourrait alors être formulée avant que cet argument puisse être utilisé par l'anti-représentationnaliste physicaliste : en admettant que l'argument de la privauté de fait ainsi présenté soit convaincant, il se pourrait toutefois les propriétés physiques intrinsèques des états cérébraux qui, selon le théoricien anti-représentationnaliste physicaliste, déterminent le caractère phénoménal des expériences, ne soient *pas* ou du moins ne soient *plus* « privées de fait ». Leur situation particulière, à l'intérieur de la boîte crânienne, ne signifie plus, depuis l'invention de techniques fiables d'observation du fonctionnement du cerveau (comme le scanner ou l'IRM), que ces propriétés soient inaccessibles à d'autres sujets que le sujet dont c'est le cerveau. Les propriétés physiques intrinsèques qui déterminent hypothétiquement le caractère phénoménal des expériences sont probablement des propriétés que je peux observer par imagerie cérébrale, dont je peux décrire l'instanciation, et à la description desquelles je peux associer des critères publics de correction (par exemple, « La propriété cérébrale X est instanciée » a pour critères publics de correction le fait que l'IRM du cerveau du sujet fasse voir tel et tel motif observable publiquement). Ces propriétés sont donc, dorénavant, publiques, y compris *dans les faits*.

Il n'en reste pas moins, et ce point est suffisant pour la question qui nous occupe, que le langage « introspectif », utilisé par les sujets lorsqu'ils parlent de leurs propres expériences, est un langage qui, s'il réfère aux propriétés des expériences, y réfère sur la base de l'hypothétique accès privé des sujets à celles-ci. Ce langage, pour l'instant du moins, n'est pas normé par les méthodes d'accès public aux états cérébraux, telles que les techniques d'imagerie cérébrale. En d'autres termes, le langage introspectif n'a pas substantiellement changé depuis l'invention de l'imagerie cérébrale et/ou à cause de cette invention. Si ce type d'actes linguistiques n'a pas vu, pour l'instant, ses critères d'application, son contexte et son mode d'utilisation changer, l'on peut considérer qu'il n'y a pas de raison pour que sa capacité à entrer ou non dans des descriptions douées de sens ait

---

<sup>20</sup> Voir le paragraphe 293 des *Recherches philosophiques* (Wittgenstein, 2004).

changée. Je propose de nommer un tel langage, élaboré du temps de la privauté de fait d'une entité et qui n'a pas vu ses règles changer depuis la « publication » de l'entité en question, un langage « classique ».

La thèse d'après laquelle l'absence de changement de règles d'un langage élaboré du temps de la privauté d'une entité, lors de la « publication » de cette dernière, signifie que ce langage est un langage « classique » dont les capacités à entrer ou non dans une description ne sont pas modifiées par cette publication constitue la *quatrième prémisse de l'argument de la privauté de fait* (P4\*). Considérer que le langage par lequel les sujets tentent de décrire eux-mêmes leurs expériences (hors contexte de laboratoire, par exemple) est un langage « classique » en ce sens me semble constituer un fait suffisamment obvie pour fournir un outil argumentatif à l'anti-représentationnaliste. Je vais donc formuler cet outil sous la forme d'une deuxième conclusion à l'argument de la privauté de fait.

*Deuxième conclusion de l'argument de la privauté de fait* : si une entité a longtemps été privée de fait mais ne l'est plus, si les termes du langage qui ont été forgés du temps de sa privauté n'ont pas vu leurs règles d'usages changer lors de sa « publication », alors les termes de ce langage « classique » ne peuvent pas plus qu'autrefois être des termes doués de sens, prenant place dans un langage descriptif, et référant à ces entités privées (quand bien même seraient apparus, parallèlement, d'autres termes permettant de décrire ces entités sur la base d'un accès public à celles-ci, et dotés pour leur part de critères publics de correction).

Ces deux arguments ayant été présentés, je pense que l'on peut admettre que l'anti-représentationnaliste, entendu comme celui qui s'oppose à la thèse d'après laquelle les propriétés représentationnelles d'une expérience consciente fixent et déterminent complètement son caractère phénoménal, peut endosser ces versions de l'argument du langage privé. Que cet anti-représentationnaliste soit un théoricien physicaliste d'après lequel les expériences conscientes ont leur caractère phénoménal déterminé par des propriétés physiques (cérébrales) intrinsèques et non-représentationnelles de l'expérience (Block, 1996, 2003, 2005), ou bien par des propriétés intrinsèques purement mentales et qualitatives (les *qualia* compris en leur sens le plus restreint, comme des propriétés intrinsèques, qualitatives et non-physiques des expériences conscientes) (Jackson, 1982, 1986), l'anti-représentationnaliste peut soutenir une position d'après laquelle ces propriétés existent et déterminent le caractère phénoménal de nos expériences sans que nous soyons capables de les décrire adéquatement à l'aide de termes doués de sens. Dans le cas d'un théoricien des *qualia* non-physiques, la raison en serait que ces entités sont privées de droit, ayant un mode d'existence intrinsèquement subjectif (et il s'appuierait sur l'argument de la privauté de droit). Dans le cas d'un théoricien des *qualia* physiques, une telle position serait justifiée par l'argument de la privauté de fait : ces entités étant des propriétés intrinsèques et non-représentationnelles de nos états cérébraux, elles n'ont longtemps été accessibles qu'au sujet dont le cerveau instancie ces propriétés, tandis qu'elles restaient inaccessibles aux autres sujets en raison de leur situation matérielle ; et notre langage introspectif, élaboré dans le contexte de cette privauté de fait, n'a pas pu se voir attribuer de critères publics de correction, et pour cette raison ne permet pas d'entreprendre la description de ces entités.

Ce que je veux maintenant montrer, c'est que l'anti-représentationnaliste qui endosse ces arguments est en position de répondre aux arguments représentationnalistes de Dretske que je viens de présenter ; il est également en mesure d'intervenir dans les débats sur le représentationnalisme dont la dialectique s'est construite autour de contre-exemples supposés, pour montrer que les apparents succès des représentationnalistes, faces à ces contre-exemples putatifs, ne constituent en rien des confirmations du représentationnalisme. C'est ce que je vais m'employer à montrer maintenant.

## **5. Contre les arguments pour le représentationnalisme ; refuser ses succès dialectiques**

Commençons par examiner le premier des arguments représentationnalistes, **l'argument de la classification des expériences**. La première prémisse de cet argument, rappelons-le, est la thèse

d'après laquelle nous classons nos expériences en faisant référence aux propriétés des objets auxquels nous associons ces expériences. L'anti-représentationnaliste n'a pas à récuser cette première prémisse, qui semble bien appuyée par nos pratiques linguistiques ; en revanche, il peut refuser d'accepter la seconde prémisse, d'après laquelle « l'explication irrésistible » du fait que nous classons nos expériences par les propriétés des objets auxquels nous associons ces expériences est que les différences phénoménales de nos expériences sont déterminées par leurs différences représentationnelles.

En effet, l'anti-représentationnaliste qui accepte l'un des deux arguments que j'ai présentés dispose d'une explication tout à fait raisonnable et naturelle au fait remarqué par la première prémisse. Car si les arguments présentés précédemment sont corrects, alors nous ne pouvons pas utiliser des termes doués de sens permettant de décrire des entités privées de droit ; et nous ne pouvons pas non plus utiliser des termes doués de sens permettant de décrire des entités privées de fait (ou anciennement privées de fait) dans un langage « classique » élaboré sur la base de cet accès privé<sup>21</sup>. Pour cette raison, que nos expériences aient leur caractère phénoménal déterminé ou non par d'hypothétiques entités privées (de droit, ou de fait), il faut s'attendre à ce que nous décrivions et classions de toute façon nos expériences, non en référence à ces entités, mais seulement en faisant appel aux propriétés publiques des objets publics qui causent typiquement ces expériences – les seuls objets et les seules propriétés que nous pouvons décrire à l'aide d'un langage classique correctement formé, doué de sens, dotés de critères véritables (c'est-à-dire publics) de correction.

Autrement dit, pour l'anti-représentationnaliste qui accepte les arguments de privauté, il est *tout à fait naturel* que nous classions nos expériences en référence aux entités publiques qui semblent causer ou accompagner leurs occurrences (et donc en « expérience d'un cercle bleu » et « expérience d'un fa dièse »), pour la simple et bonne raison que nous serions incapables de décrire leurs propriétés intrinsèques et privées (de droit ou de fait), même si elles existaient (et l'anti-représentationnaliste suppose par ailleurs qu'elles existent). Imaginons en effet qu'il existe des *qualia* non-physiques qui déterminent le caractère phénoménal de nos expériences : aucun terme de notre langage descriptif ne permettrait d'y faire référence, et donc de détailler ce type d'entités. Imaginons que des propriétés physiques intrinsèques, pendant très longtemps publiquement inaccessibles, déterminent ce même caractère phénoménal : aucun terme de notre langage descriptif « classique » ne permettrait d'y faire référence et de détailler ces entités.

De cette manière, l'anti-représentationnaliste, qu'il soit physicaliste ou anti-physicaliste, dispose d'une explication naturelle du fait remarqué par la première prémisse, ce qui lui permet de refuser la deuxième prémisse de l'argument de la classification des expériences. Cela empêche l'argument de fonctionner en faveur du représentationnalisme : le représentationnalisme n'est pas l'explication irrésistible du fait linguistique remarqué par Dretske, puisque la théorie concurrente dispose d'une explication de ce fait parfaitement naturelle.

Venons-en maintenant au second argument, **l'argument de la spécification des propriétés**. Rappelons-le, la première prémisse de cet argument consistait en l'idée d'après laquelle, si l'anti-représentationnaliste veut fournir un contre-exemple philosophique au représentationnalisme, alors il doit décrire un cas d'expérience E dans laquelle les choses apparaissent F, où F est une propriété impossible à représenter comme étant instanciée. La suite de l'argument appuie l'idée d'après laquelle il est impossible de fournir un tel contre-exemple, puisque spécifier une propriété F comme étant irréprésentable est impossible (spécifier une propriété étant déjà une façon de la représenter).

Cette fois, c'est la première prémisse de l'argument que peut et doit refuser l'anti-représentationnaliste qui accepte les arguments de privauté. En effet, un anti-représentationnaliste de ce type admet que, pour autant que les expériences voient leur caractère phénoménal déterminé par autre chose que des propriétés représentationnelles, et notamment par des propriétés intrinsèques *privées* (en droit ou en fait), nous ne disposons pas de termes dans notre langage descriptif qui soient capables d'y faire correctement référence ; dans le cas de propriétés seulement

---

<sup>21</sup> Je tiens pour acquis que les pratiques linguistiques auxquelles fait référence Dretske participent d'un tel langage « classique », apparu bien avant les techniques d'imagerie cérébrale ou même les neurosciences.



*anciennement privées de fait*, nous ne disposons pas de termes dans notre langage descriptif « classique » qui permette d'y faire correctement référence.

Pour cette raison, l'anti-représentationnaliste serait bien en peine de fournir un contre-exemple dans lequel il pourrait spécifier une certaine façon F qu'ont les choses d'apparaître dans notre expérience, telle que F n'est pas représentable. Au contraire, l'anti-représentationnaliste doit maintenir que fournir un bon contre-exemple philosophique au représentationnalisme, loin d'impliquer que nous détaillions une expérience E telle que les choses y apparaissent F (où F n'est pas représentable), doit plutôt dire qu'un tel contre-exemple ne peut être ainsi présenté, et que les cas corrects de contre-exemples au représentationnalisme sont plutôt des cas d'expérience E dans lesquelles le caractère phénoménal de l'expérience, le « vécu » que constitue l'expérience, ne *semble pas être uniquement déterminé et épuisé par ce que nous représentent ces expériences*. Mais, et c'est le point important, l'anti-représentationnaliste ne doit surtout pas admettre l'exigence de *spécification* des aspects de ce caractère phénoménal. Pour l'anti-représentationnaliste dualiste, une telle spécification, énoncée dans un langage dont les termes ont des critères publics de correction, *manquerait* nécessairement la nature réelle (privée et indescriptible) des propriétés qui déterminent le caractère phénoménal des expériences. Pour l'anti-représentationnaliste physicaliste, une telle spécification, si elle est fournie dans le langage « classique » (et c'est bien le seul langage qui est disponible pour le philosophe qui, dans son fauteuil, tente de détailler la nature de ses expériences), ne peut également que manquer son objet, puisque si cette spécification est une description douée de sens alors elle ne peut référer aux entités qui déterminent réellement le caractère phénoménal des expériences – mais elle finira sans aucun doute par référer à des propriétés « classiquement » publiques d'objets classiquement publics.

Quant au troisième argument de Dretske en faveur du représentationnalisme, **l'argument de la détection des propriétés**, son statut est un peu différent. A première vue, il peut être contré de la même façon. Sa première prémisse est en effet la même que celle de l'argument précédent, et concerne la nature des contre-exemples que devrait pouvoir fournir un anti-représentationnaliste. Elle doit à nouveau être refusée, en récusant l'exigence de spécification des aspects du caractère phénoménal qui échappent à la détermination par le contenu représentationnel de l'état en question. Un représentationnaliste pourrait tenter de reformuler ce troisième argument sous une autre forme, en arguant de ce que, si son opposant est physicaliste, alors les propriétés qui déterminent le caractère phénoménal des expériences doivent *de toute façon* être des propriétés physiques (en l'occurrence, cérébrales), donc des propriétés détectables – et pour cette raison représentables (que nous soyons actuellement capables de les spécifier ou pas). L'anti-représentationnaliste physicaliste pourrait certes de son côté répondre qu'une propriété peut être représentable sans être une propriété représentationnelle (ou sans qu'on puisse lui faire correspondre une propriété représentationnelle), et que ce serait typiquement la situation d'hypothétiques propriétés intrinsèques des expériences qui détermineraient leur caractère phénoménal. Toutefois, une telle réponse ne ferait pas fond sur la prise en compte des arguments de privauté ; en ce sens, l'on peut dire que les arguments inspirés de l'argument du langage privé ne permettraient qu'au théoricien dualiste des *qualia* comme entités *purement et ontologiquement privées* de répondre à une telle reformulation du troisième argument de Dretske<sup>22</sup>, et que le physicaliste anti-représentationnaliste devrait y faire face par d'autres moyens.

Voilà comment l'anti-représentationnaliste peut s'appuyer sur les arguments inspirés des remarques sur le langage privé pour contrer les arguments représentationnalistes précédemment exposés. On voit que le chemin argumentatif est légèrement différent selon que l'anti-représentationnaliste choisit une option dualiste ou physicaliste ; toutefois, ces deux voies ont une inspiration commune, en ce qu'elles s'appuient toutes deux sur une réflexion quant aux conditions que doivent remplir nos pratiques linguistiques pour pouvoir être des pratiques descriptives. Il se

---

<sup>22</sup> Sachant bien sûr que l'anti-représentationnaliste dualiste qui défend l'existence de *qualia* purement privés accepte, en un sens, la thèse que Dretske semble considérer comme aberrante, d'après laquelle il existe des propriétés qui sont bel et bien *indétectables* (pour tout autre que le sujet qui les instancie).

peut qu'un philosophe de l'opinion qui était celle de Dretske ne prête guère attention à l'argumentation du dualiste, étant donné que le physicalisme peut sembler avoir par ailleurs en sa faveur de *bons arguments* indépendants des questions liées au caractère représentationnel des expériences ; toutefois, il me semble pertinent de fournir des outils argumentatifs anti-représentationnalistes *également pour le dualiste*, étant donné que certains dualistes embrassent actuellement le représentationnalisme (Chalmers, 2004).

Je désire à présent de montrer de quelle manière les arguments de privauté permettent également à l'anti-représentationnaliste de refuser d'interpréter la capacité de réponse du représentationnalisme aux contre-exemples qui lui ont été présentés comme des succès et des confirmations de celui-ci. Pour ce faire, je reprendrai le modèle que j'avais appliqué pour contrer les arguments de Dretske.

En effet, la dialectique du débat focalisé sur les contre-exemples putatifs à la théorie représentationnaliste reposait, du côté représentationnaliste, sur un procédé de spécification de ces fameux aspects non-représentationnels des expériences. Qu'a, par exemple, l'orgasme de si spécifique ? Il est plaisant. En quoi se distinguent la vision et l'audition de quelque chose au-dessus de moi ? Ce que je perçois est, dans un cas, visible, dans l'autre cas audible. A la suite de quoi, le représentationnaliste pouvait supposer la présence, dans l'ensemble de ce qui est représenté par l'état en question, de propriétés correspondant à ces spécifications.

Mais le fait que le représentationnaliste soit alors en mesure de supposer un contenu représentationnel correspondant à ces spécifications ne constitue pas nécessairement un succès du représentationnalisme. En effet, d'après l'anti-représentationnaliste qui endosse les arguments de privauté, il est *normal* et *naturel* que l'étape de *spécification* du caractère phénoménal d'une expérience donnée (ou de la différence phénoménale entre deux expériences données), exécutée à l'aide d'un langage « classique » doté de critères publics d'application, s'effectue en référence à des objets et à des propriétés publiquement accessibles de manière classique – dont le représentationnaliste aura ensuite beau jeu de supposer qu'ils sont *représentés* par les expériences en question. Autrement dit : une fois que vous avez accepté de spécifier les caractères phénoménaux pertinents, il sera toujours aisé de les rapporter à des propriétés d'objets publiquement accessibles de manière classiques, et particulièrement d'objets externes – propriétés dont on peut ensuite toujours supposer que l'état mental concerné les représente.

En d'autres termes, l'anti-représentationnaliste qui accepte les arguments de privauté prédit, tout autant que le représentationnaliste, qu'une fois passée l'étape de la spécification (dans un langage classique) des aspects supposément non-représentationnels des expériences en question, le représentationnaliste sera en mesure de déterminer un contenu représentationnel qui pourra rendre compte des aspects phénoménaux problématiques, si on les comprend *tels qu'ils ont été spécifiés*.

Or, c'est à nouveau cette étape de spécification que peut et doit refuser l'anti-représentationnaliste qui accepte les arguments de privauté et qui désire avancer un contre-exemple au représentationnalisme. En effet, comme nous l'avons dit, pour l'anti-représentationnaliste dualiste, une telle spécification, énoncée dans un langage dont les termes ont des critères publics de correction, *manquerait* nécessairement la nature réelle (privée et donc indescriptible) des propriétés qui déterminent le caractère phénoménal des expériences. Pour l'anti-représentationnaliste physicaliste, une telle spécification, fournie dans le langage « classique » dont disposent les philosophes et les introspecteurs en fauteuil, manquerait également son objet, puisque si cette spécification constituait une description douée de sens elle ne pourrait référer aux propriétés physiques intrinsèques qui déterminent réellement le caractère phénoménal des expériences en question.

Si l'on voulait réinterpréter l'un des contre-exemples d'une manière compatible avec ce style d'anti-représentationnalisme, on devrait ainsi s'arrêter avant l'étape de la spécification et présenter les choses, par exemple, de la manière suivante :

« L'orgasme (ou la douleur intense), sont des expériences qui sont vécues d'une manière qui ne peut être seulement déterminée par ce que ces expériences représentent du monde lorsque je les ai ; au contraire, leur phénoménologie excède ce qu'elles indiquent et disent du monde. Cette

intuition doit pouvoir être saisie telle quelle, sur la base de l'introspection, pour ceux qui y parviennent. Pour ceux qui n'y parviennent pas ou qui la refusent, pour ceux qui ne *voient pas ce dont je veux parler*, je ne peux pas spécifier les aspects de cette expérience qui ne sont pas déterminés par son contenu représentationnel, dans la mesure où je ne possède pas de termes qui permettent de désigner les aspects en question. Si je tentais de décrire ces aspects, cette description se ferait donc nécessairement en référence à des propriétés publiques d'objets publics classiquement accessibles (par exemple les objets et les propriétés généralement associés à ces expériences), ces objets et ces propriétés étant les seuls auxquels les termes que je suis capable d'employer dans un usage descriptif peuvent correctement référer. Cela pourrait inciter à croire que les aspects dont je veux parler sont simplement déterminés par la représentation de ces propriétés et objets publics classiquement accessibles; or, ce n'est précisément pas le cas »<sup>23</sup>.

Une telle présentation de contre-exemple prendrait sans aucun doute le risque d'être moins convaincante, et d'être moins à même de faire changer d'avis un philosophe convaincu, de manière préalable, par le représentationnalisme. Toutefois, il faut faire à ce propos trois remarques : d'une part, il n'est pas du tout sûr qu'un tel contre-exemple n'ait, pour autant, *aucun* pouvoir de conviction. Deuxièmement, ce type de contre-exemple empêche la formulation des réponses représentationnaliste classiques. Troisièmement, ce type de contre-exemple n'est pas le résultat d'une adaptation *ad hoc*, destinée justement à contrer les réponses représentationnalistes, mais elle correspond au contraire à la seule formulation cohérente de contre-exemples pour un anti-représentationnaliste qui accepterait les arguments de privauté.

## Conclusion

Le représentationnalisme à propos de l'expérience consciente, identifié à la thèse d'après laquelle le caractère phénoménal des expériences conscientes est entièrement déterminé par le contenu représentationnel de celles-ci, constitue une position assez largement acceptée. Si l'on ajoute à cela qu'elle constitue, sous cette formulation précise, une idée relativement récente, et qu'elle a été l'objet de débats vifs dans les trente dernières années – débats dont elle a souvent pu sembler sortir victorieuse – on peut comprendre pourquoi le représentationnalisme peut, à tous ces égards, apparaître comme un point relativement bien établi de la *doxa* contemporaine en philosophie de l'esprit.

Mon but, dans cet article, était de montrer pourquoi certains des arguments les plus convaincants et les plus influents en faveur du représentationnalisme pouvaient en réalité être contrés par l'anti-représentationnaliste, au prix d'une acceptation de certains arguments inspirés de l'argument wittgensteinien du langage privé. Je désirais également montrer pourquoi et comment l'anti-représentationnaliste pouvait refuser d'interpréter certains des succès dialectiques les plus éclatants du représentationnalisme comme des succès prédictifs, et donc des confirmations, de la conception représentationnaliste de l'expérience consciente.

Des arguments d'inspiration wittgensteinienne ont souvent pu être opposés à un grand nombre de théories de l'esprit, et notamment à l'ensemble de celles qui traitaient l'esprit comme un objet parmi d'autres, dont il est possible de produire une description scientifique objective. A ce

---

<sup>23</sup> Rappelons bien sûr que, pour un anti-représentationnaliste physicaliste qui accepterait les arguments de privauté, la spécification des propriétés non-représentationnelles qui déterminent le caractère phénoménal des expériences est *en dernière instance possible* ; mais cette spécification ne peut pas se faire dans le vocabulaire classique du philosophe qui introspecte dans son fauteuil. Une telle spécification se ferait plutôt dans les termes des neurosciences, et sur la base suivante : il s'agirait de repérer les propriétés cérébrales telles que (1) Elles sont corrélées avec certains types de caractère phénoménaux ; (2) Les sujets y sont introspectivement sensibles (ils peuvent les détecter introspectivement) ; (3) L'instanciation de ces propriétés ne porte en elle-même aucun contenu représentationnel, ne représente rien (ce qui peut vouloir dire différentes choses suivant la théorie du contenu représentationnel embrassée par le théoricien concerné). Ce seraient donc des propriétés physiques non-représentationnelles, qui bien entendu pourraient elles-mêmes être représentées, quoiqu'elles ne puissent pas être véritablement décrites dans un langage « classique ».

titre, les réflexions de Wittgenstein ont pu être considérées comme s'opposant autant à des formes classiques et modernes de représentationnalisme qu'aux théories anti-représentationnalistes, comme les théories des *qualia* non-physiques (Jackson, 1982) ou des *sense-data* (Russell, 1912), ou encore les théories des *qualia* intrinsèques mais de nature physique (Block, 2007). J'ai essayé de montrer ici que la prise en compte d'un aspect particulier des réflexions wittgensteiniennes (l'impossibilité d'un langage privé), sans acceptation de l'intégralité de la pensée de Wittgenstein, pouvait être utilisée, non pour argumenter directement contre le représentationnalisme en tant que tel, mais simplement pour montrer en quoi certaines des observations généralement convoquées pour appuyer celui-ci ne devraient pas en réalité être considérées comme des raisons de l'accepter.

Dans quelle position se trouve, à l'issue de ce cheminement, l'anti-représentationnaliste qui a accepté les arguments de privauté ? Il est, je crois, en position de répondre à certains arguments représentationnalistes ; il est aussi en mesure de formuler des contre-exemples au représentationnalisme sous une forme telle qu'il n'est pas aussi aisé que précédemment de leur apporter une réponse. Certes, une difficulté se présente à lui : il semble qu'il sera bien en peine de développer par lui-même une théorie positive des propriétés qui déterminent, peut-être en sus, ou peut-être en lieu et place des propriétés représentationnelles, le caractère phénoménal des expériences ; car il sera impossible, au moins dans un langage philosophique « classique », de fixer la référence des termes de cette théorie et de les utiliser de manière descriptive. Si cette impossibilité est sans aucun doute un désagrément, voire un inconvénient, elle ne permet en aucun cas de penser que la position qui l'implique est fautive.

## Bibliographie :

- Block, N. (1995). On a confusion about a function of consciousness. *Brain and Behavioral Sciences*, 18, 227–247.
- Block, N. (1996). Mental Paint and Mental Latex. *Philosophical Issues*, 7, 19–49.
- Block, N. (2003). Mental paint. In M. Hahn & B. Ramberg (Éd.), *Reflections and Replies: Essays on the Philosophy of Tyler Burge* (p. 165–200). MIT Press.
- Block, N. (2005). Two Neural Correlates of Consciousness. *TRENDS in Cognitive Sciences*, 9(2), 46–52.
- Block, N. (2007). Wittgenstein and Qualia. *Philosophical Perspectives*, 21(1), 73–115.
- Brentano, F. (1944). *Psychologie du point de vue empirique*. (M. de Gandillac, Trad.). Paris: Aubier.
- Byrne, A. (2001). Intentionalism defended. *Philosophical Review*, 110(2), 199–240.
- Candlish, S., & Wrisley, G. (2012). Private Language. In (E. Zalta, Éd.) *Stanford Encyclopedia of Philosophy (Summer 2012 Edition)*. Consulté à l'adresse <http://plato.stanford.edu/archives/sum2012/entries/private-language>
- Chalmers, D. (2004). The representational character of experience. In B. Leiter (Éd.), *The Future for Philosophy* (p. 153–181). Oxford University Press.
- Dretske, F. (1995). *Naturalizing the Mind*. MIT Press.
- Dretske, F. (2003). Experience as representation. *Philosophical Issues*, 13(1), 67–82.
- Fodor, J. (1975). *The Language of Thought*. Harvard University Press.
- Harman, G. (1990). The intrinsic quality of experience. In J. Tomberlin (Éd.), *Philosophical Perspectives, 4: Action Theory and Philosophy of Mind*. Atascadero, CA: Ridgeview.
- Horgan, T., & Tienson, J. (2002). The Intentionality of Phenomenology and the Phenomenology of Intentionality. In D. Chalmers (Éd.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford: Oxford University Press.
- Husserl, E. (1962). *Recherches Logiques, tome 2*. (H. Elie, L. Kelkel, & R. Schérer, Trad.) (Vol. 2). Paris: PUF.

- Husserl, E. (1969). *Méditations cartésiennes: introduction à laphénoménologie*. (G. Peiffer & E. Levinas, Trad.). Paris: J. Vrin.
- Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly*, 32(April), 127-136.
- Jackson, F. (1986). What Mary didn't know. *Journal of Philosophy*, 83(May), 291-5.
- Kriegel, U. (2013). *Phenomenal Intentionality*. Oxford: Oxford University Press.
- Lycan, W. (1996). *Consciousness and Experience*. MIT Press.
- Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, 83(October), 435-50.
- Peacocke, C. (1983). *Sense and Content: Experience, Thought, and Their Relations*. Oxford: Oxford University Press.
- Peacocke, C. (1992). *A Study of Concepts*. Cambridge (Mass.): MIT Press.
- Russell, B. (1912). *The Problems of Philosophy*. Oxford University Press.
- Seager, W., & Bourget, D. (2007). Representationalism about consciousness. In M. Velmans & S. Schneider (Éd.), *The Blackwell Companion to Consciousness*. Oxford: Blackwell Publishing.
- Searle, J. (1983). *Intentionality*. Cambridge: Cambridge University Press.
- Siewert, C. (1998). *The Significance of Consciousness*. Princeton: Princeton University Press.
- Speaks, J. (2010). Attention and Intentionalism. *Philosophical Quarterly*, 60(239), 325-342.
- Thalabard, E. (2013). Représentationnalisme et attention. *Dialogue*, 52(3), 551-586.
- Thau, M. (2002). *Consciousness and Cognition*. Oxford: Oxford University Press.
- Tye, M. (1995a). Blindsight, Orgasm and Representational Overlap. *Behavioral and Brain Sciences*, 18, 268-269.
- Tye, M. (1995b). *Ten problems of consciousness : a representational theory of the phenomenal mind*. Cambridge (Mass.): the MIT Press.
- Tye, M. (1996a). Orgasms Again. *Philosophical Issues*, 7, 51-54.
- Tye, M. (1996b). Perceptual Experience is a Many-Layered Thing. In E. Villanueva (Éd.), *Perception* (Vol. 1-117-1-126). Atascadero, CA: Ridgeview.
- Tye, M. (2002). Representationalism and the Transparency of Experience. *Noûs*, 36(1), 137-151.
- Wittgenstein, L. (2004). *Recherches Philosophiques*. (F. Dastur, M. Elie, J.-L. Gautero, D. Janicaud, & E. Rigal, Trad.). Paris: Gallimard.
- Wu, W. (2011). What is Conscious Attention? *Philosophy and Phenomenological Research*, 82(1), 93-120.