

---

# UN RETOUR SUR LES ORIGINES CLASSIQUES DU RÉALISME INTENTIONNEL

Cédric BRUN

Université Michel de Montaigne, Bordeaux, France

Lycée Charles de Gaulle, Ambassade de France,

70 Paris Caddesi, Kavaklidere, Ankara, Turquie

Mél : Cedbrun@pleksus.net.fr ou cedricbrun@yahoo.com

---

## Résumé

Les principaux acteurs de la révolution cognitive ont pris le parti de constituer les sciences cognitives en faisant table rase des théories classiques de la connaissance parce qu'ils jugeaient celles-ci trop simples et malencontreusement liées à des doctrines métaphysiques que la science ne peut prendre pour objet. Malgré ce discours de principe, de nombreux auteurs (comme J. Fodor par exemple) font appel à ces thèses au prétexte qu'elles permettent de saisir leurs propres théories — parce que les premières ne seraient que des versions simplifiées, des ébauches des dernières. Nous nous proposons ici de revenir aux représentationnalismes classiques sur l'exemple de la causalité mentale afin de mesurer ce que nous pouvons en apprendre. Nous tenterons de montrer que l'on peut trouver chez Locke une théorie de la causalité mentale qui ne dissocie pas type et contenu des attitudes mentales, mais fait de leur relation le moteur causal dont nous avons besoin pour construire une théorie de l'esprit qui rende compte des explications du réalisme intentionnel.

## Abstract

The cognitive revolution is an attempt to build a cognitive science by refusing all the old fashioned classical theories of knowledge because of their links with some metaphysical thesis contradictory with the main claims of the sciences. In spite of this principle statement, numerous authors (such as J. Fodor for instance) seem to need these old fashioned epistemological doctrines to give an account of their own — the old thesis should be some drafts of the new one. Our claim is that we should come back to the classical representational theories of mind in a more acute way in order to weight up what we can learn about mind from them to complete or correct our contemporary knowledges. We attempt here to point out that we can find in Locke's *Essay on Human Understanding* a theory of mental causality which don't radically separate the relation (hoping, believing, etc.) and the content of propositionnal attitudes. On the contrary, this theory make this connection between relation and content into a driving causal we are needing in order to build a theory of mind that naturalizes intentional explanations of human behaviour.

## 1. Introduction : intentionnalisme d'hier et d'aujourd'hui<sup>2</sup>

Une des voies empruntées pour construire la science cognitive repose sur un lieu commun : nous expliquons nos agissements et ceux de nos contemporains en faisant appel à des désirs, des croyances ou toutes autres attitudes propositionnelles. Certains auteurs se proposent de rendre compte de façon scientifique de cette intuition commune. Pour cela, l'une des pistes choisies est de tenter de montrer que l'efficacité causale des attitudes propositionnelles repose sur leurs propriétés intentionnelles (i.e. sur le fait qu'elles portent sur autre chose qu'elles-mêmes) et que ces propriétés intentionnelles sont explicables en termes non intentionnels ou naturels (i.e. dans les termes des sciences physiques).

L'un des promoteurs de cette entreprise, Jerry Fodor, insiste souvent sur l'inscription du projet de naturalisation de l'intentionnalisme dans une longue

tradition philosophique, il affirme par exemple :

*“ Je ne cesserai tout au long de mon propos, de souligner à quel point la science cognitive est une entreprise récidiviste. (...) Elle n'est à bien des égards rien d'autre que la redécouverte de doctrines familières — pour ne pas dire de lieux communs — de la tradition épistémologique classique, celle qui va de Descartes à Kant, en passant par Locke, Berkeley et Hume. ”* (Fodor, 1987a, pp. 301-302)

Nous voudrions montrer ici que certaines positions du réalisme intentionnel contemporain sont très proches des thèses classiques de la philosophie de l'action et de la connaissance. Plus encore, nous voudrions montrer l'utilité qu'il peut y avoir à renouer un dialogue précis et circonstancié avec les auteurs classiques autour de ces questions pour saisir les limites et les erreurs éventuelles des thèses contemporaines.

Nous présenterons d'abord quelques thèses caractéristiques du réalisme intentionnel et de la version représentationnaliste soutenue par Fodor. Puis, nous nous focaliserons sur la façon dont Locke parvient à rendre compte de la causalité mentale dans *l'Essai Philosophique Concernant l'Entendement Humain* (EEH). Enfin, nous tirerons les conséquences de ce travail pour montrer ce qu'une étude plus générale des intentionnalismes classiques peut apporter à la tentative

---

<sup>2</sup> Une version primitive de cet article a été présentée lors du “ 3<sup>ème</sup> Colloque jeunes chercheurs en sciences cognitives ” le 27 avril 1999 à Soulac en Gironde. Que tous ceux qui ont permis par leurs questions et leurs remarques l'évolution de mon propos soient remerciés et particulièrement P-Y Raccach, F. Lowenthal, J-M Roy, et mon directeur de thèse J. Terrel pour son dévouement et sa patience.

de naturalisation des explications intentionnelles de la psychologie du sens commun.

## 2. Intentionnalisme et psychologie populaire

### 2.1. Quelques enjeux philosophiques de l'étude de la psychologie populaire

Nous avons une faculté étonnante : nous pouvons prévoir et expliquer les comportements des hommes et des femmes qui nous entourent. Plus encore, nous sommes capables de déterminer avec une certaine réussite quels sont leurs états psychologiques (pourtant apparemment censés n'être accessibles que par introspection). Nous manipulons pour ce faire des notions et des termes décrivant leurs craintes, leurs désirs, leurs croyances, leurs perceptions et leurs douleurs. Nous inférons des prédictions de leurs comportements à partir de certaines observations de leurs comportements et de lois tacitement admises telles que :

(1) *Une personne qui éprouve de la douleur a tendance à vouloir la faire disparaître.*

(2) *Le visage d'une personne en colère rougit.*

La question de savoir en quoi consiste au fond cette psychologie populaire et quelle est sa valeur dans le projet de naturalisation de l'esprit est depuis longtemps au centre de nombreux débats ; il serait donc illusoire de vouloir, ici, régler ce problème<sup>3</sup>. Du moins devons-nous dire ce qu'il faut entendre par "psychologie du sens commun" ou "psychologie populaire" et éclairer les lignes de partage essentielles entre les diverses positions relatives à cette "psychologie".

#### 2.1.1. Comment définir la psychologie populaire ?

L'une des hypothèses les plus courantes consiste à dire que la psychologie populaire est une forme de compétence spontanée, voire innée, au même titre que notre *physique naïve* (l'ensemble des connaissances intuitives que nous avons à propos des objets de notre quotidien) mais appliquée aux comportements de nos contemporains (Dennett, 1987, chap.1). Une autre façon de caractériser la psychologie populaire consiste à la considérer comme un ensemble de capacités pratiques, de dispositions acquises, un art de la prévision des comportements des agents qui nous entourent. Nous nous en tiendrons ici à une caractérisation globalement acceptable par chacune des parties en présence. Nous considérerons que la psychologie populaire est l'ensemble des hypothèses et des notions présumées fondant notre utilisation dans la vie quotidienne d'un vocabulaire "psychologique", c'est-à-dire relatif à ce que nous nommons des "états mentaux" tels que la croyance, la crainte, etc., ce, quelle que soit la nature de ce réseau de notions et de présumés (qu'il s'agisse ou non d'une théorie, innée ou acquise, interne ou externe) et que ces états mentaux aient une existence physique (cérébrale) ou non.

#### 2.1.2. Quelle valeur scientifique accorder à la psychologie populaire ?

On peut établir une ligne de partage entre deux grands courants dans les sciences cognitives : d'un côté certains cherchent à conserver et à affermir la place de la psychologie populaire et d'un autre, certains cherchent à l'éliminer de toute construction d'une théorie naturalisatrice de l'esprit. La psychologie populaire a été étudiée tant par ceux qui se proposent de tenter de la réduire à des lois physiques que par ceux qui, au contraire, lui dénie toute valeur théorique. Pour ces derniers elle ne serait qu'une façon totalement impropre de décrire la "vie mentale" des hommes et femmes qui nous entourent de la même façon que les termes et les lois de notre physique naïve sont trompeurs et ont été contredits par les théories physiques académiques. Dès lors, l'objectif des sciences cognitives ne serait pas de traduire la psychologie populaire en une psychologie scientifique, mais il faudrait construire une science psychologique qui remplace le cadre inadéquat de notre psychologie naïve (Churchland, 1984, chap. 2, section 5).

Cette contestation de la valeur de la psychologie populaire s'appuie sur deux idées principales : d'une part sur la thèse issue du positivisme logique selon laquelle la signification de n'importe quelle phrase repose en dernière analyse sur les circonstances de l'observation qui permettraient de vérifier ou de confirmer cette phrase ; d'autre part sur l'hypothèse générale du courant analytique selon laquelle les problèmes philosophiques ne sont que des pseudo problèmes qui trouvent leur origine dans des confusions linguistiques ou conceptuelles et qui peuvent être résolus par une analyse du langage dans lequel ils ont été énoncés. D'autres raisons connexes sont avancées par les antiréalistes, mais elles ne sont que des compléments ou des formulations différentes de ces deux arguments principaux. Nous en exposerons brièvement certaines plus loin (voir 2.2.2.).

Le béhaviorisme philosophique — puisque tel est le nom donné à la tradition philosophique condamnant les ambitions scientifiques de la psychologie populaire — prétend que, lorsque nous parlons de nos croyances, de nos désirs ou de nos volontés, nous ne faisons qu'utiliser un raccourci linguistique (des "manières de parler") pour désigner des formes potentielles ou actuelles de comportement plutôt que des objets ayant une existence "mentale" effective (cela se double donc du point de vue ontologique d'un refus de toute forme de dualisme). Plus précisément, des auteurs comme Ryle et Wittgenstein ont soutenu l'idée que les explications de la psychologie ordinaire en termes d'attitudes propositionnelles n'ont que "l'air" d'être causales. De fait, pour un wittgensteinien, la connexion entre nos raisons et nos actions n'est pas une connexion causale mais une connexion "conceptuelle" ou "logique". L'idée centrale de cette position consiste à dire que nous ne connaissons pas nos motifs ou nos raisons d'action de la même manière que nous connaissons des causes. Les raisons ne peuvent pas être des causes dans la mesure où elles ne sont pas extérieurement observables (elles sont seulement accessibles à l'introspection). Ainsi nous ne pouvons pas les repérer comme "causes" étant donné qu'il faudrait, pour cela, pouvoir observer leur concaténation régulière d'un point de vue public

<sup>3</sup> On pourra se reporter à ce sujet à (Stich, 1978), (Churchland, 1984, chap. 3), (Horgan et Woodward, 1985) et (Fodor, 1986).

(l'introspection, nécessairement privée, est impuissante à cela). Cependant, elles ne peuvent pas plus être des causes "intérieures". Lorsque nous disons que la motivation correspond à une forme de causalité intérieure, nous transposons en fait à l'intérieur un mécanisme hypothétique que nous connaissons immédiatement alors que les causes ne *peuvent* pas être connues immédiatement. Comme Wittgenstein l'indique lui-même dans *The blue and the brown books* (Wittgenstein, 1958, p. 15), cette possibilité est logique et non empirique. C'est-à-dire que la connexion entre le constat de certaines actions et l'explication par certaines raisons fonctionne comme une quasi-définition de ces actions ou des concepts qui les désignent. C'est en ce sens que ces connexions sont "conceptuelles", "logiques" ou "grammaticales". D'après lui, le motif est énoncé comme une interprétation que nous donnons de l'acte (public). Cette interprétation est élaborée sur la base de certains critères c'est-à-dire sur la base de certaines manifestations comportementales — dans le cas de la douleur par exemple (Wittgenstein, 1958, p. 27) — conventionnellement ou publiquement reconnues comme critères.

En d'autres termes — et c'est ce que retiennent les héritiers actuels du béhaviorisme logique — il n'y a rien de "mental" derrière les termes de la psychologie ordinaire, il n'y a que des dispositions à agir physiquement. Il n'y a donc pas de différence qualitative entre la propriété dispositionnelle du sucre : "être soluble" et mon "état mental" exprimé ainsi : (1) "je veux un café". Dans les deux cas, il serait faux de dire que le sucre ou moi possédons un état interne de type mental ou autre qui explique que l'on dira (2) "le sucre est soluble" ou de moi que je veux un café. Les béhavioristes philosophiques nous dirons, "nous pouvons donner une définition opérationnelle de la proposition considérée : (2) est équivalent à (3) "si le sucre était mis dans un liquide non saturé, il se dissoudrait", de la même façon que (1) est équivalent à (4) "si on me demandait ce que je veux, je répondrais du café" et (5) "si on m'offrait un café je le prendrai et le boirais", etc."

Cette condamnation de la psychologie populaire ne nous semble pas pertinente pour deux raisons essentielles.

Premièrement elle nie la nature interne des états mentaux (Ryle, 1949, chap. 2, section 5). Pourtant, lorsque je me tape malencontreusement sur un doigt avec un marteau, ma douleur n'est pas qu'une simple disposition à crier, injurier mon marteau, porter mon doigt à ma bouche, le panser, etc. Il y a une réalité de la douleur comme phénomène interne révélé par l'introspection qu'une théorie de l'esprit ne peut ignorer de façon aussi radicale<sup>4</sup>

Deuxièmement (et surtout), si l'on veut vraiment préciser la disposition correspondant à la proposition simple (1) "je veux un café", on se rend compte que le nombre de propositions opérationnelles nécessaire est non

seulement très important mais nous engage sans doute dans une analyse indéfiniment longue. Les descriptions opérationnelles correspondant à la proposition (1) sont en nombre infini... De plus, on se rend compte que ces propositions opérationnelles sont insuffisantes pour caractériser la proposition. Il est indispensable de faire appel à des éléments de types mentaux dans la description opérationnelle pour en spécifier certaines conditions. Ainsi les deux propositions : (4) "si on me demandait ce que je veux, je répondrais du café" et (5) "si on m'offrait un café, je le prendrais et le boirais" ne sont pas des définitions opérationnelles suffisantes pour caractériser la proposition (1). En effet, la proposition (4) n'est vraie que si la proposition (6) "je cherche à cacher que je veux un grand verre de limonade" est fausse et la proposition (5) n'est vraie que si (7) "je crois que ce que l'on m'a offert est bien du café et non une boisson lyophilisée aromatisée au café" est vraie, etc. Les propositions (6) et (7) font appel à des termes (chercher à cacher, croire) qui appartiennent à la psychologie populaire et semblent désigner des états mentaux comme conditions psychologiques de la valeur de (4) et (5). Dans le cas où l'on ne fait pas intervenir de conditions psychologiques pour spécifier les propositions opérationnelles, on tombe dans des tautologies sans intérêt qui ne clarifient pas les expressions de la psychologie populaire<sup>5</sup>.

Malgré ce débat fondamental, nous utilisons toujours cette psychologie du sens commun aussi quotidiennement que nous nous brossons les dents. De la même façon, nous disons toujours que le soleil se couche et se lève sur l'horizon alors que nous savons pertinemment que ce n'est là qu'un mouvement apparent dû à la rotation de la terre autour du soleil. Sans doute l'efficacité de cette psychologie populaire dans notre activité journalière nous encourage-t-elle à ne pas la réformer. C'est là un des arguments essentiels de ceux qui veulent préserver la psychologie populaire tout en la naturalisant comme J. Fodor par exemple dans son article *Banish discontent* :

*"Nous ne pouvons pas nous passer de la psychologie des attitudes propositionnelles, d'une part parce qu'elle fonctionne et d'autre part parce que rien d'autre ne fonctionne."* (Fodor, 1986, p. 420 ; nous traduisons)

La psychologie du sens commun est apparemment le seul moyen dont nous disposons aujourd'hui pour interpréter, prédire, commenter, comprendre les comportements des êtres qui nous entourent. En cela, elle doit constituer une bonne base de départ (en tout cas la meilleure que nous ayons) pour la constitution d'une théorie de l'esprit. C'est là l'essence de la position réaliste à l'égard de la psychologie populaire défendue par Fodor (Fodor, 1985 et 1987). On peut donc considérer que la psychologie ordinaire n'est qu'une "proto-théorie" dont la psychologie cognitive conserverait des traits généraux, mais dont elle chercherait à déterminer les fondements dans la nature même de la vie mentale.

<sup>4</sup>Cet argument peut poser problème parce qu'il fait référence à la douleur alors que les béhavioristes philosophiques parlent généralement d'états mentaux de type propositionnel, mais nous le posons ici simplement pour mémoire et en réponse au passage du texte de Wittgenstein auquel nous renvoyons plus haut (Wittgenstein, 1958, p.27).

<sup>5</sup> Voir (Carnap, 1936-1937) pour une discussion plus globale de la question de la traduction d'un langage théorique en un langage opérationnel.

L'objectif de naturalisation n'est donc pas une simple formalisation de la psychologie populaire. Il s'agit de constituer une théorie de l'esprit qui assure un sauvetage théorique de la psychologie ordinaire et du réalisme intentionnel qu'elle semble admettre de façon spontanée. En d'autres termes encore, il s'agit de "sauver les phénomènes", dans la mesure où notre attitude première consiste bien à expliquer notre comportement par l'appel à des causes d'ordre psychologique.

## 2.2. Quelques enjeux philosophiques de l'étude de l'intentionnalité

Nous avons défini la psychologie du sens commun comme l'ensemble des hypothèses et des notions que nous présupposons qui fondent notre utilisation dans la vie quotidienne d'un vocabulaire "psychologique". Or, l'une des particularités de certains états mentaux — ou psychologiques — auxquels ce vocabulaire semble se référer est ce que Brentano (Brentano, 1874) appelait *l'intentionnalité*.

### 2.2.1. Qu'est-ce que l'intentionnalité ?

L'intentionnalité est la propriété de certains états mentaux (et de certaines choses comme les phrases, les mots ou les films) par laquelle ils sont "à propos de". Par exemple, ma pensée que "le café est noir" est "à propos" du café (et également à propos du fait d'être noir) et se distingue de ma pensée que "la neige est blanche" dans la mesure où cette pensée porte sur la neige et non sur le café ; sur le fait d'être blanc et non sur le fait d'être noir. Nos pensées manifestent donc de l'intentionnalité dans la mesure où elles sont "à propos de" quelque chose. Cela délimite certainement le concept d'intentionnalité mais ne le clarifie guère car l'expression "être à propos" reste vague.

De nombreux problèmes surgissent dès que l'on cherche à préciser la nature de l'intentionnalité : cela n'a fait qu'exciter la curiosité des philosophes et donne lieu à une abondante littérature. Nous esquisserons ici un bref descriptif des principaux éléments de l'étude contemporaine de l'intentionnalité dans le projet de naturalisation de la psychologie du sens commun afin de situer plus précisément la position fodorienne dans le débat actuel.

### 2.2.2. États intentionnels et états qualitatifs

Revenons brièvement<sup>6</sup> sur le fait relevé plus haut que seuls *certain*s états mentaux sont intentionnels. Nous devons en effet distinguer entre des états mentaux de types intentionnels (ils renvoient à des entités désignées par leur contenu sémantique) et des états mentaux de types qualitatifs non intentionnels. La différence est désormais classique mais il est bon de la préciser un peu. Les croyances, les désirs, les craintes et les intentions manifestent la particularité "d'être à propos" de ce qui constitue le contenu sémantique de la proposition qui les exprime. Par contre, si l'on considère des états comme l'expérience visuelle d'une violette, l'expérience olfactive d'un vin du Médoc, ou l'expérience gustative d'un morceau de réglisse ; on se rend compte

qu'ils sont caractérisés par des qualités subjectives intrinsèques (d'où leur nom de *qualia*) dépourvues de toute propriété sémantique et par conséquent incommunicables. Il est par exemple impossible de décrire une couleur à un aveugle, de la lui "donner à voir", par une description — aussi précise soit-elle — dans les termes de la physique. Ces états mentaux relèvent simplement de notre conscience phénoménale dans la mesure où ces sensations sont liées à "l'effet que cela fait de les avoir" (Dennett, 1987, p.107) ; tandis que nos états mentaux intentionnels supposent un accès à un contenu linguistiquement exprimable.

Les propriétés intentionnelles de ces derniers états mentaux relèvent, nous l'avons dit, du fait qu'ils "portent sur" autre chose qu'eux-mêmes, qu'ils entretiennent une certaine relation de type sémantique avec d'autres choses. Ces états mentaux sont appelés attitudes propositionnelles parce qu'ils peuvent être identifiés par leurs contenus propositionnels. Ainsi, ces états mentaux sont-ils transcritibles sous la forme d'une phrase faisant intervenir un sujet grammatical et un verbe désignant l'individu sujet de cet état mental et qualifiant le type de ce dernier (croyance, doute, crainte, espoir, désir, etc.) et un complément ou une proposition complémentaire déterminant le contenu de l'état mental :

- (3) *Je redoute d'avoir à manger des épinards.*
- (4) *Paul veut un café.*
- (5) *Alice croit que le café est noir.*

On remarquera plusieurs particularités que les attitudes propositionnelles sont réputées posséder.

Tout d'abord, on estime généralement qu'elles sont internes. C'est-à-dire qu'elles ne sont observables que par introspection. Vous pouvez certainement m'attribuer des attitudes propositionnelles par l'observation de ce que je dis, écris ou fais, mais vous ne pouvez avoir d'expérience directe de ce que sont mes croyances, mes désirs, mes craintes et mes espoirs.

Deuxièmement, elles peuvent porter indifféremment sur des objets qui existent ou sur des objets qui n'existent pas (voire qui sont impossibles). Par exemple de nombreux enfants croient au Père Noël bien que celui-ci n'ait aucune existence empirique ou historique. Mes croyances ou mes désirs peuvent ainsi être à propos d'états de choses impossibles ou fausses. Par exemple, il est possible de vouloir chevaucher un cheval ailé ou de croire que le plus grand des nombres entiers est un nombre premier. Dans ces cas là — des attitudes propositionnelles portant sur des choses impossibles ou fausses — il faut remarquer que ces états mentaux sont dépourvus de valeur de vérité.

Ces caractéristiques intentionnelles des attitudes propositionnelles ont des conséquences troublantes, sinon énigmatiques. Ces conséquences ne sont énigmatiques qu'à partir du moment où nous adoptons comme ici, et désormais de façon généralisée, une ontologie moniste physicaliste. Elles sont généralement mises en évidence par les partisans du béhaviorisme philosophique contre les réalistes intentionnels et tracent les contours d'un sujet de discorde au sein du projet de construction d'une théorie de l'esprit.

<sup>6</sup> Pour plus de détails, on se reportera à (Nagel, 1974), (Tye, 1986) et (Dennett, 1990)

### 2.2.3. Deux énigmes posées par l'intentionnalité

La première énigme est ontologique : " De quelle façon une chose matérielle, physique, peut-elle posséder des propriétés intentionnelles ou sémantiques en propre ? ". On voit en effet très bien qu'une idée porte sur autre chose qu'elle-même. C'est même une trivialité que de dire que nos croyances désignent des états de choses distincts d'elles-mêmes. La possibilité que nous avons de formuler linguistiquement nos prétendues attitudes propositionnelles montre bien qu'en plus des qualités directionnelles ou représentationnelles (elles sont mises pour, elles désignent...) qui sont les leurs, elles possèdent des qualités sémantiques. Mais justement, comment une chose physique peut-elle posséder toutes ces qualités ?

Selon Fodor, la question de savoir comment une chose matérielle peut posséder des qualités intentionnelles est loin d'être une énigme. Il se propose de montrer comment c'est possible par ce qu'il appelle la théorie computationnelle de l'esprit (TCRE).

Telle que Fodor la présente, elle repose sur le constat d'une parenté entre le fonctionnement à la fois causal et intentionnel des états mentaux et le fonctionnement causal et intentionnel des symboles. Les symboles linguistiques présentent en effet cette caractéristique d'être des objets physiques (sonores, graphiques, etc.) géométriquement ou physiquement caractérisables qui ont à ce titre des propriétés causales dispositionnelles (réfléchir la lumière, propriétés ondulatoires, etc.), mais qui ont aussi la particularité d'être intentionnellement reliés à certains objets.

Un énoncé tel que :

(6) *Voie sans issue.*

a plusieurs caractéristiques remarquables. Il possède des qualités physiquement caractérisables (masse de l'encre, indice de réfraction de la lumière...) qui ont des conséquences sur mon action (je n'emprunterai pas cette route si j'ai pu d'abord lire cet énoncé). Il porte sur quelque chose (le fait que cette route est un cul-de-sac), Et cet énoncé a aussi pour caractéristique de posséder une valeur de vérité. Au même titre que ma croyance que :

(7) *Cette voie est sans issue.*

l'énoncé ne sera vrai que si la route contextuellement pertinente est effectivement sans issue.

À partir du constat de cette ressemblance ou de cette parenté, Fodor envisage deux types de réductions théoriques :

*" Ce parallélisme entre les symboles et les états mentaux est tout simplement trop frappant pour qu'on puisse le considérer comme purement accidentel : il appelle de façon criante une réduction théorique. Deux stratégies sont à l'évidence possibles. Soit on montre que les propriétés intentionnelles des états mentaux dérivent des propriétés sémantiques des symboles ; soit on montre que les propriétés sémantiques des symboles dérivent des propriétés intentionnelles des états mentaux. "*  
(Fodor, 1987a, p. 307)

Fodor entrevoit donc deux thèses réductionnistes simples sur la base de la similitude ou de la parenté entre les propriétés intentionnelles des états mentaux et les propriétés intentionnelles des symboles linguistiques. L'une (thèse 1) consiste à réduire les propriétés intentionnelles des états mentaux aux propriétés intentionnelles des symboles linguistiques qui énoncent cette attitude propositionnelle (voir figure 1). L'autre (thèse 2) consiste, au contraire, à réduire les propriétés intentionnelles des symboles linguistiques aux propriétés intentionnelles plus originelles des états mentaux (suivant le principe selon lequel la parole exprime la pensée : voir figure 2).

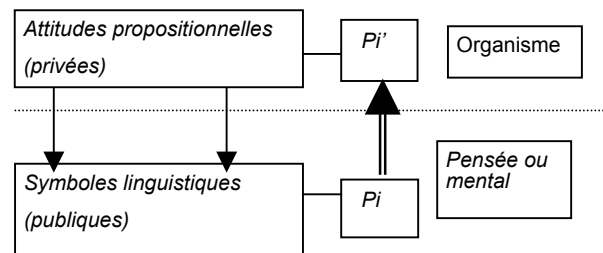


Figure 1 — Thèse 1 : (Béhaviorisme logique).

$P_i$  : Propriétés intentionnelles primitives  
 $P_i'$  : Propriétés intentionnelles dérivées  
 $\rightarrow$  : Réduction théorique  
 $\Rightarrow$  : Dérivation généalogique  
 ..... : Séparation entre le domaine de l'organisme support des états mentaux et son environnement.

Ainsi, nous rencontrons à travers la thèse 1 une idée familière au courant Wittgensteinien. Cela revient à dire, par exemple, que *la croyance* que : " César a franchi le Rubicon " est identique à une *disposition à dire* que César a franchi le Rubicon. Ce qui, en dernière analyse, implique *l'énonciation* de symboles linguistiques publics signifiant que César a franchi le Rubicon. Fodor résume cette position de la façon suivante :

*" La thèse est que c'est cette connexion "conceptuelle" entre les croyances et les énonciations qui confère un contenu intentionnel aux premières. "* (Fodor, 1987a, p. 308)

Fodor refuse la thèse wittgensteinienne pour deux raisons essentielles. Tout d'abord, admettre que la pensée dépend de la possession de compétences linguistiques exclut *a priori* la possibilité pour tout être vivant infra humain ou infra linguistique de posséder un esprit. Si nous ne pouvons pas expliquer le comportement animal en termes de désirs et de croyances, alors nous devons refuser aussi toute continuité entre les animaux et le genre humain, ce qui revient à invalider tout ou partie des thèses évolutionnistes qui semblent avoir fait leurs preuves dans les sciences de la nature. Cela pose donc un problème majeur de consistance épistémologique.

Ensuite, si l'on affirme que les propriétés intentionnelles des états mentaux sont dérivées de la capacité à l'énonciation linguistique, nous ne pouvons plus comprendre comment nous apprenons une langue. Un bébé est bien un organisme infra linguistique, mais cela ne l'empêche pas d'apprendre à parler, ce qui est la marque de la capacité à utiliser son esprit pour des

activités cognitives supérieures malgré l'absence de capacités linguistiques originelles. Ces deux arguments mériteraient certes d'être plus précisément et plus longuement décrits (dans leur double dimensions théorique et polémique), mais cela nous éloignerait du cœur de notre propos. Il suffisait ici de montrer que Fodor refuse la solution antiréaliste à la question de la nature des propriétés intentionnelles des états mentaux.

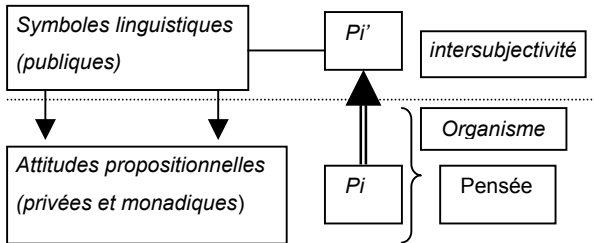


Figure 2 — Thèse 2 : Fonctionnalisme monadique.

- $P_i$  : Propriétés intentionnelles primitives  
 $P_i'$  : Propriétés intentionnelles dérivées  
 $\rightarrow$  : Réduction théorique  
 $\Rightarrow$  : Dérivation généalogique  
 ..... : Séparation entre le domaine de l'organisme support des états mentaux et son environnement.

Non seulement il refuse la thèse 1, mais Fodor estime aussi que la thèse 2 ne permet pas de saisir pleinement la nature du lien entre propriétés intentionnelles des états mentaux et propriétés intentionnelles des symboles linguistiques. Dériver les propriétés sémantiques du langage des propriétés intentionnelles de la pensée ne suffit pas. Certes, le langage a pour fonction d'exprimer la pensée. La phrase :

(7) *César a franchi le Rubicon.*

a la signification qu'elle a parce que nous l'utilisons pour exprimer une certaine croyance possédant un contenu intentionnel.

Cependant, si l'on s'en tient à dire que les propriétés sémantiques des signes linguistiques sont dérivées de propriétés intentionnelles plus fondamentales des états mentaux, alors nous sommes contraints à concevoir les attitudes propositionnelles comme des états fonctionnels monadiques, c'est-à-dire comme des états qui ne sont définis que par référence à leurs causes et à leurs effets typiques.

Autrement dit, chaque attitude propositionnelle serait parfaitement particulière dans son existence, causée par tel(s) état(s) de(s) chose(s) et causant elle-même tel(s) état(s) de(s) chose(s) ; c'est en ce sens que Fodor dit que ces états mentaux seraient *monadiques*. Nous n'aurions des attitudes propositionnelles que comme effets de certaines causes, leurs propriétés intentionnelles trouvant leur origine dans ce lien causal. Selon cette thèse, chaque type d'état mental consiste — étant donné certaines entrées (ou données) sensorielles et certains états mentaux — dans une disposition à avoir certains états mentaux et à agir d'une certaine façon.

Si l'on s'en tient à cette seule hypothèse nous ne pouvons pas penser deux propriétés importantes que sont réputées posséder les attitudes propositionnelles.

Le premier problème que ne peut résoudre de façon satisfaisante la simple dérivation des propriétés intentionnelles des symboles linguistiques à partir des propriétés intentionnelles des attitudes propositionnelles tient au fait que nous pouvons toujours distinguer au sein des attitudes propositionnelles des *types* (croyance, désir, crainte, etc.) et des *contenus intentionnels* (ce sur quoi porte l'attitude propositionnelle). Ce problème est celui des "contenus communs" (Fodor, 1987a, p. 311).

Si les états mentaux sont fonctionnels et monadiques et non relationnels, alors il est impossible de comprendre comment nous pouvons avoir des attitudes propositionnelles de types différents ayant pourtant des éléments communs. Ainsi, *le désir de faire du bateau* et *la peur de faire du bateau* sont deux états mentaux distincts qui ont pourtant des éléments communs (ils ont le même contenu). Si ces états mentaux sont fonctionnels et sont monadiques (c'est-à-dire s'ils fonctionnent comme des unités singulières déterminées par des causes et ayant des effets déterminés) alors nous devons considérer que l'élément commun doit être une propriété de leur rôle causal. Il nous faudrait découvrir un élément qui soit à la fois causé par le fait de désirer et de craindre faire du bateau et qui implique, du fait de cette relation causale elle-même, que ces deux états mentaux distincts portent sur l'action de faire du bateau. On peut difficilement résoudre ce problème si l'on s'en tient à la thèse 2 il semble donc, sans aller plus loin pour l'instant, que l'hypothèse d'états mentaux de type relationnel permette d'éviter ce premier écueil.

Le deuxième problème est celui que Fodor appelle "le problème de la constituance" (Fodor, 1987a, p. 311) ; c'est un aspect du problème plus général de la compositionnalité. Fodor observe que les croyances comme la plupart des attitudes propositionnelles semblent avoir une structure composée, l'exemple qu'il donne est le suivant :

*"La croyance qu'il pleut et que les rues sont mouillées semble être intimement liée à la croyance qu'il pleut et à la croyance que les rues sont mouillées. Mais le problème est le suivant : quelle est la nature de cette relation ?"* (Fodor, 1987a, p. 311).

Si les croyances et les autres états mentaux ont cette structure de constituance (s'ils sont composés de parties) la façon la plus simple de rendre compte de la nature de cette relation semble revenir à dire que la relation de constituance est une relation de partie à tout. Ainsi, dans l'exemple donné par Fodor, la croyance qu'il pleut est une partie de la croyance qu'il pleut et que les rues sont mouillées. De même, *mutatis mutandis*, pour la croyance que les rues sont mouillées.

Mais comme le remarque aussitôt Fodor, cette solution ne semble fonctionner que parce que l'exemple est mal choisi. L'exemple est une conjonction. Or toutes les conjonctions fonctionnent *a priori* sur un type de relation de partie à tout. Mais la propriété de constituance n'est pas la propriété des seules croyances conjonctives, on la trouve également dans les croyances disjonctives. Or, si ces croyances sont bien constituées par des éléments plus fondamentaux, elles ne peuvent avoir de relation du type "tout à partie" aux éléments qui les constituent. En effet, la croyance qu'il pleut ou que les rues sont mouillées, par exemple, n'est pas constituée par la

croissance qu'il pleut ou par la croyance que *les rues sont mouillées*. La croyance *il pleut ou les rues sont mouillées* est totalement différente (au sens ou elle n'est pas composée par les deux parties) de la croyance qu'il pleut et de la croyance que *les rues sont mouillées*, même si elle doit posséder une certaine parenté avec ces deux croyances. Fodor exprime cela assez clairement lorsqu'il écrit :

*“ Il semble que la croyance que P ou Q ait une certaine P-ité et une certaine Q-ité ; mais ce que peuvent être cette P-ité et cette Q-ité reste une énigme, puisque ce n'est ni la croyance que P ni la croyance que Q. ”* (Fodor, 1987a, p. 312)

La solution qu'entrevoit Fodor à ce problème est inspirée par l'observation du fonctionnement de la parole. Une assertion que P ou Q n'est pas l'assertion que P ou l'assertion que Q. On rend compte de ce phénomène par l'appel à la constituance symbolique : alors que l'on peut affirmer que P ou Q sans affirmer que P ou sans affirmer que Q, on ne peut pas produire le symbole “ P ou Q ” sans produire les symboles “ P ” et “ Q ”.

Ainsi, tant que nous pensons les états mentaux comme des états fonctionnels monadiques, nous ne pouvons comprendre le phénomène de constituance de façon simple. Par contre la solution qui consiste à saisir les états mentaux comme des relations à des symboles mentaux ou à des représentations mentales permet de rendre compte simplement de la propriété de constituance des attitudes propositionnelles. Si avoir la croyance qu'il pleut ou que les rues sont mouillées c'est être en relation avec une certaine représentation mentale, alors nous pouvons faire l'hypothèse qu'il s'agit d'une représentation complexe qui combine une représentation qui signifie que *les rues sont mouillées* et une représentation qui signifie qu'il pleut.

L'hypothèse selon laquelle les attitudes propositionnelles sont des relations avec des représentations mentales qui fonctionnent comme des symboles linguistiques permet donc d'éviter les deux écueils principaux (tenant à la compositionnalité des états mentaux) rencontrés par la thèse 2 : le problème de la nature de ce qui est commun aux états mentaux qui ont des contenus communs et le problème de la structure de constituants des états mentaux. Ainsi la thèse 2 qui consiste en une réduction simple des propriétés intentionnelles des symboles linguistiques communs aux propriétés intentionnelles plus fondamentales des états mentaux ne suffit-elle pas à rendre compte clairement de la nature des états mentaux. Fodor remarque que si cette thèse ne suffit pas, elle reste malgré tout nécessaire même s'il faut lui adjoindre une nouvelle réduction : la réduction des états mentaux à des représentations ou à des symboles mentaux. C'est là l'originalité de la thèse présentée par Fodor, et nous verrons que c'est en ce sens qu'elle constitue une théorie représentationnelle de l'esprit.

Nous reviendrons plus loin sur le détail de la stratégie choisie par Fodor et certaines de ses limites. Il nous suffisait ici de mettre en évidence que le débat sur la possibilité d'une naturalisation de la psychologie du sens commun exigeait une étude de l'intentionnalité sous son aspect ontologique.

Une autre facette du problème posé par la notion d'intentionnalité est non pas ontologique mais

épistémologique. Nous l'avons déjà rencontré à propos de la psychologie populaire et il faut dire que l'antiréalisme repose en fait bien souvent sur ce paradoxe épistémologique soulevé par les propriétés intentionnelles et nous ne pourrions que l'évoquer ici. Le paradoxe peut-être présenté de la façon suivante ; il résulte de la réunion de trois affirmations que nous avons présentées de façon éparse jusqu'à présent. Premièrement, on fait appel à des notions qui présupposent des états mentaux pourvus de propriétés intentionnelles véritables pour expliquer les comportements dits intelligents. Deuxièmement, ces qualités intentionnelles des états mentaux sont réputées ne pas être accessibles à l'observation empirique. Troisièmement on cherche à construire une science du comportement humain qui réponde aux réquisits essentiels de la scientificité (matérialisme ou physicalisme, méthodologie expérimentale, consistance théorique et cohérence empirique).

On voit assez vite que l'on doit soit abandonner l'idée de construire une science psychologique véritable, soit refuser toute valeur de vérité aux propositions qui font appel à des propriétés intentionnelles. Telle est, du moins, l'alternative présentée par certains antiréalistes comme les Churchlands. Leur position est clairement antiréaliste car, puisqu'il ne saurait être question de renoncer à construire une science véritable de la cognition, ils se sont prononcés en faveur d'une élimination pure et simple des lois et des propositions de la psychologie populaire et des notions intentionnelles de toute construction de la science cognitive.

Pour le dire autrement, la réduction de la psychologie populaire à une science psychologique matérialiste est aussi impossible à envisager que pourrait l'être la réduction de l'alchimie à la chimie. Dans les deux cas, on a affaire, selon la position éliminativiste des Churchland, à des concepts pseudo scientifiques dépourvus de toute référence qui ne peuvent être réduits à des concepts physiques empiriquement cohérents et théoriquement consistants. La discussion de cette théorie réclamerait une présentation plus précise et déborderait le cadre de cette étude. On retiendra cependant que l'unanimité est loin d'être acquise et que l'optimisme fodorien quant à la naturalisation de l'intentionnalité doit être tempéré par certaines critiques comme celle-ci.

### 2.3. Les fondements classiques du représentationnalisme

Avant d'aller plus loin dans la présentation de la position de Fodor et sa mise en perspective par un appel à la théorie lockienne de l'efficacité causale des contenus intentionnels des états mentaux, il nous faut justifier plus précisément le choix qui est le nôtre de rapprocher ainsi théories contemporaines du réalisme intentionnel et théories classiques de la connaissance et de l'action. Pour cela, nous ferons appel à deux idées : d'une part le fait que les tenants contemporains du représentationnalisme se réclament ouvertement d'une filiation classique ; d'autre part, en montrant qu'il y a effectivement une base représentationnaliste et intentionnaliste dans des philosophies classiques de la connaissance comme celle de Locke.

### 2.3.1. Un héritage revendiqué

Comme le rappelle Jerry Fodor, la science cognitive :

*“ n’est (...) qu’une redécouverte de doctrines familières (...) de la tradition épistémologique classique, celle qui va disons de Descartes à Kant en passant par Locke, Berkeley et Hume. ”* (Fodor, 1987a, p. 301-302.)

ou encore, elle :

*“ présente en effet cette curieuse particularité d’avoir plus en commun avec les théories psychologiques des XVIIème, XVIIIème et XIXème siècles qu’avec celles de la première partie du XXème siècle : de sorte que plusieurs de ses thèses caractéristiques auraient été plus recevables par John Stuart Mill ou William James que par Watson, Hull, Thorndike, ou, bien entendu, Skinner. ”* (Fodor, 1987a, p. 302.)

Cette affirmation est assez courante aujourd’hui et demande à être éclairée. Nous ne pouvons pas développer ce qui constitue un immense programme de recherche, mais nous tracerons quelques pistes et quelques frontières essentielles.

Malgré cette parenté reconnue et revendiquée par la communauté des chercheurs en sciences cognitives, on ne trouve pas de recherche proprement circonscrite sur la teneur et la consistance théorique et empirique des thèses développées dans les philosophies classiques de la connaissance. Si la plupart des philosophes de l’esprit contemporains revendiquent un lien de filiation remontant au moins jusqu’à Locke, il est frappant de voir à quel point les descriptions qu’ils donnent de sa position restent schématiques. On peut comprendre que ces références n’aient pas un but historique comme le rappelle Fodor :

*“ Je ne connais pas de meilleure façon de voir comment marchent les théories représentationnelles de l’esprit que d’examiner les versions relativement simples qu’en ont forgé ses premiers partisans. (...) Je me propose donc maintenant, très brièvement et sans aucun souci d’exactitude historique de jeter un coup d’œil à un des prototypes de la psychologie de la représentation mentale. Je l’appelle la “ théorie de Hume ” bien que ce soit peut-être diffamatoire pour Hume. ”* (Fodor, 1987a, p. 312-313. Nous soulignons.)

Cette utilisation purement instrumentale, à fin didactique, des thèses “ primitives ” de la théorie représentationnelle de l’esprit n’est certes pas condamnable si elle est clairement affirmée et assumée dans son caractère approximatif — comme c’est le cas ici. Cependant, elle ne doit pas non plus nous dédouaner d’une enquête plus rigoureuse sur la réelle consistance de ces théories représentationnelles primitives. Il nous semble, en particulier, qu’il faudrait revenir sur la prétendue “ simplicité ” des versions primitives des théories représentationnelles de l’esprit. C’est peut-être le point le plus troublant de l’affirmation de Fodor.

Si la position de Hume ou de Locke est si simple, on est en droit de se demander 1) si elle est vraiment pertinente pour saisir la complexité des théories représentationnelles contemporaines de l’esprit, 2) s’il ne

vaudrait pas mieux la décrire en détail afin de bien faire apparaître les difficultés qui lui sont inhérentes et de les éviter dans la théorie que l’on cherche à établir, 3) si cette théorie est si simple pourquoi est-il obligé de la schématiser afin de la rendre plus compréhensible. Toutes ces questions nous encouragent en fait à approfondir cette description et cette étude des représentationnalismes classiques. Mais il reste à justifier cette qualification de représentationnalisme pour les théories de la connaissance classiques telles que celle de Locke.

### 2.3.2. En quel sens parler de représentationnalismes classiques ?

Il faut bien reconnaître, avec Fodor une parenté sur au moins deux points avec certaines philosophies classiques de la connaissance. Premièrement, un bon nombre repose sur l’idée fondamentale que le comportement humain est dirigé par la volonté comprise comme faculté de l’esprit. De ce point de vue là elles sont toutes des formes de réalisme intentionnel. Deuxièmement, elles se sont toutes, de façon centrale ou superficielle, interrogées sur la nature de la relation de représentation afin de déterminer comment nous pouvons connaître les choses du monde qui nous entourent et/ou agir sur celui-ci. De ces deux points, nous ne développerons que le second.

Nous pouvons convenir que la plupart des philosophies classiques (toutes... ?) se sont interrogées sur la nature et les limites du pouvoir de connaître des hommes. Elles avaient sans doute à chaque fois des objectifs différents, mais ce qui les rapproche de Fodor, c’est que cette interrogation les a amenées à postuler une relation de type représentationnelle entre nos idées et les choses du monde extérieur qu’elles désignent. Autrement dit, si représentationnalisme il y a, il est d’abord dans l’ordre de la perception. Nos “ idées ”, nos “ perceptions ”, nos “ impressions ”, représentent dans notre esprit ce qui est nécessairement extérieur à lui (les choses). Dans la Lettre à Gibieuf du 19 janvier 1642, Descartes ne fait que mettre le doigt sur ce principe représentationnaliste :

*“ Je ne puis avoir aucune connaissance de ce qui est hors de moi que par l’entremise des idées que j’en ai eu en moi, je me garde bien de rapporter mes jugements immédiatement aux choses et de leur rien attribuer de positif que je ne l’aperçoive auparavant en leurs idées ; mais je crois aussi que tout ce qui se trouve en ces idées est nécessairement dans les choses. ”*<sup>7</sup>

Ici, le terme d’idée employé par Descartes est à saisir en un sens large d’objet de l’entendement. On retrouve chez Locke ou chez Hume des affirmations du même ordre. Locke, affirme par exemple :

*“ Il est évident que l’esprit ne connaît pas les choses immédiatement mais seulement par l’intervention des idées qu’il a de celles-ci. ”* (Locke, EEH, IV, IV, §3. Nous traduisons.)

Ou encore, selon Hume :

*“ La philosophie la plus vulgaire nous apprend qu’un objet extérieur ne peut pas se révéler à*

<sup>7</sup> Descartes, Lettre à Gibieuf du 19 janvier 1642, éd Adam et Tannery, tome III, p. 474, lignes 13 à 20.

*l'esprit immédiatement et sans l'intermédiaire d'une image ou d'une perception. Cette table qui, en ce moment même, m'apparaît, est seulement une perception, et toutes ses qualités sont les qualités d'une perception.* " (Hume, THN, p. 239. nous traduisons.)

À la différence terminologique près (Hume parle ici de perception et non d'idée) on retrouve la même affirmation de radicale incapacité de l'esprit humain à sortir de lui-même, et à atteindre les objets extérieurs sans médiateur représentationnel. L'accès immédiat à ces objets extérieurs nous est interdit, nous sommes condamnés à n'avoir affaire qu'à des représentations. Ce représentationnalisme de la perception n'a pas la même forme chez chacun de ces philosophes et a donné lieu à des interprétations différentes par les commentateurs, mais ce qui nous intéresse ici c'est que loin de s'en tenir à un représentationnalisme de la perception, Locke (comme Hume d'ailleurs) est allé jusqu'à proposer une théorie de la signification qui donne aux idées une place de médiateur entre nous et nos actions ou les phénomènes du monde et nous.

Le principe général reconnu par les trois philosophes que nous avons cité revient à dire que le monde ne peut être connu qu'en tant que nous pouvons nous le représenter, nos facultés cognitives (la mémoire, la perception, l'imagination, etc.) ne sont que des facultés différentes de représenter les choses. Il s'agit alors de savoir comment une chose peut en représenter une autre. Chez Locke comme chez Fodor les signes linguistiques vont servir de marchepied pour s'élever à une théorie de la connaissance réaliste.

Les mots fournissent en effet un bon moyen de nous représenter le monde ; mais dès que l'on s'interroge sur la nature de leur signification, on se rend compte qu'il est nécessaire de faire intervenir (du moins est-ce la thèse de Locke) quelque chose dont les mots soient les signes. Pour qu'un mot ait un sens il faut qu'il ait une certaine dénotation. Or les mots ne peuvent être les signes directs des choses, sans quoi nous ne pourrions construire de termes abstraits puisqu'un mot pour avoir un sens devrait signifier une chose et une seule (il faudrait postuler un langage nomenclatural ou adamique). Il faut donc supposer que les mots (au moins les termes généraux) sont signifiants dans la mesure où ils sont les signes d'idées (abstraites en l'occurrence). L'idée n'ayant aucune communauté de nature avec l'objet externe, elle prend donc le statut de signe, et même de signe arbitraire chez Locke. Les idées se comportent donc comme des d'éléments sémiotiques d'un langage privé en ce qu'elles possèdent un certain nombre de propriétés opératoires (une idée possède une compréhension (on peut lui associer d'autres idées qui en forment la définition et lui sont donc substituables) et une extension (la collection d'idées qui, toutes, possèdent cette idée comme attribut), etc.) et en ce qu'il existe différentes sortes d'idées (idées simples, idées complexes, etc.) que Locke définit au livre II. Cela revient à affirmer que les mots sont des signes de signes et donc que les idées sont elles-mêmes des signes.

*"Les signes dont nous nous servons communément sont de deux sortes ; savoir les idées et les mots."* (Locke, EEH, IV, v, §2. Italiques de Locke.)

Les idées sont donc des signes en ce qu'elles permettent de passer du particulier à une première sorte de généralité, propre à notre intériorité. Sans la généralité, aucune communication n'est concevable, et même si mes idées sont déjà générales, le "langage privé" qu'elles constituent pour mon esprit ne permet pas la communication avec autrui, puisqu'il faudrait qu'en plus d'avoir eu exactement les mêmes expériences sensibles et intérieures, nous ayons la possibilité de nous les communiquer par un "canal de communication" de type télépathique...

Le langage devient donc le moyen de traduire la pensée privée de son auteur de sorte qu'il puisse la faire partager à autrui. C'est là une idée développée par Locke :

*"Quand ils [les enfants] ont plié les organes de la parole à former des sons articulés, ils commencent à se servir des mots pour faire comprendre leurs idées aux autres."* (Locke, EEH, II, xi, §8.)

Ainsi, on trouve dans certaines philosophies classiques des thèses et des positions utilisant des entités mentales comme des éléments d'un mécanisme représentationnel du fonctionnement de l'esprit qui sont fortement semblables à ce que nous observons aujourd'hui dans certaines théories du réalisme intentionnel. Jusque là nous avons simplement remarqué que le représentationnalisme de Locke valait pour la perception et pour la théorie de la signification ; nous voudrions montrer dans ce qui suit que ce représentationnalisme vaut aussi pour les explications causales du comportement humain (voir 4). Mais avant cela il est nécessaire de préciser la nature de la théorie computationnelle représentationnelle de l'esprit (TCRE) fodorienne.

### 3. La causalité mentale dans la "théorie computationnelle représentationnelle de l'esprit" <sup>8</sup>

Pour base de notre description du problème de la causalité mentale, nous nous limiterons aux thèses principales développées par J. Fodor (Fodor, 1987a et 1987b).

#### 3.1. Le premier pas : l'intentionnalisme

Selon Fodor (Fodor, 1987b, chap.1), dans les limites d'une explication causale du comportement, les attitudes propositionnelles se conforment à des généralisations *ceteris paribus* de la psychologie ordinaire comme :

*"Toutes choses égales par ailleurs, si x craint que A alors x désire que non-A" (...) "Si x désire que A et si x croit que A n'advient pas à moins que B n'advienne et si x croit qu'il peut faire que B advienne, alors, toutes choses égales par ailleurs, x essaiera de faire en sorte que B advienne".*

Autrement dit, nous expliquons un comportement tel que le franchissement du Rubicon par César en postulant des entités intentionnelles qui donnent sens aux expressions de la psychologie ordinaire. Précisément, nous caractérisons pour cela la nature ou le **type** de ces

<sup>8</sup> Que nous abrègerons désormais par TCRE.

attitudes propositionnelles (croyance et/ou désir, etc.) et tenons compte de leur *contenu*. Ainsi César a franchi le Rubicon parce qu'il **voulait accroître son pouvoir** et parce qu'il **croyait que défier ainsi le sénat** assoirait son pouvoir de façon durable si celui-ci ne réagissait pas trop violemment. C'est non seulement le **type** mais surtout le *contenu* intentionnel de l'attitude propositionnelle qui est véritablement explicatif. C'est apparemment en fonction de ce contenu que l'explication est valable ou non. Si l'explication du comportement de César revenait à dire qu'il a franchi le Rubicon parce qu'il **croyait que la terre est ronde**, nous serions en droit de trouver cette explication absurde : le contenu intentionnel de cette attitude propositionnelle n'explique pas le comportement observé. Expliquer le comportement par l'évocation d'attitudes propositionnelles n'a de sens que si l'on admet que le lien causal entre le comportement et l'attitude propositionnelle dépend du contenu intentionnel de celle-ci. Nos raisons d'agir, présentées en termes de croyances, de désirs et d'espéros ont une valeur explicative seulement parce qu'elles sont des croyances, des désirs, des souhaits *que quelque chose* ou *de quelque chose*.

Je poste ma lettre une fois timbrée parce que je crois que sans timbre elle n'arrivera pas à destination, et non parce que je crois que Marcel Proust a écrit *À la Recherche du Temps Perdu...* Le rôle causal des attitudes propositionnelles semble donc bien être lié au contenu intentionnel de celles-ci. Par exemple, soit la proposition désignant l'explication de la psychologie du sens commun suivante :

- (8) *A fait x, parce que A désire y et croit que faire x lui procurera y.*

On peut schématiser la position intentionnaliste de la façon suivante :

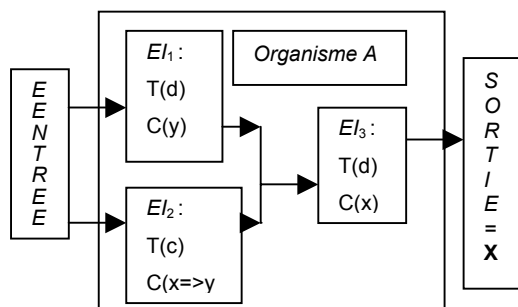


Figure 3 — Intentionnalisme.

$El_n$  : État intentionnel

$T(n)$  : Type d'état intentionnel (si  $n=d$ , alors  $T(d)$ =désir, si  $n=c$ , alors  $T(c)$ =croyance)

$C(n)$  : Contenu d'état intentionnel ( $C(x) \Leftrightarrow$  le contenu de l'état intentionnel est  $x$ )

$X$  : État de fait ou comportement  $x$ .

L'entrée correspond aux divers stimuli qui parviennent à l'organisme A, sur la base desquels les attitudes propositionnelles (ou états intentionnels) sont générées, la sortie correspond au comportement observable.

Le principe fondamental de l'intentionnalisme tel qu'il est présenté par Fodor (et tel que nous l'avons schématisé)

considère que le comportement d'un agent cognitif est causé par les propriétés intentionnelles de ses attitudes propositionnelles (ou états intentionnels notés  $E_n$ , voir figure 3). Cette première étape se conçoit comme une réduction de nos propositions descriptives de la psychologie ordinaire à des propriétés d'états mentaux intentionnels.

Dans l'exemple considéré, A va faire  $x$  parce que son désir de  $y$  et sa croyance dans le fait que  $x$  lui permettra d'obtenir  $y$  l'amènent à vouloir  $x$  et son comportement sera  $x$  dans la limite de ses capacités comportementales. Comme nous le remarquons plus haut, il y a une relation inférentielle entre les contenus intentionnels des attitudes propositionnelles qui constitue l'essentiel de la valeur explicative de (8). Notons au passage que ce type d'explication intentionnaliste permet de décrire non seulement la façon dont des états intentionnels causent des comportements mais aussi la capacité qu'ils sont présumés avoir de causer d'autres attitudes propositionnelles. Ainsi, les propriétés intentionnelles des attitudes propositionnelles du fait de leur composition en type (T sur le schéma) et contenu (C sur le schéma), permettraient de rendre compte en termes intentionnels de nos agissements.

Mais ce serait aller trop loin que de dire dès maintenant que les attitudes propositionnelles causent nos comportements. Bien sûr, c'est ce que cherche à montrer le réalisme intentionnel, mais, à ce moment de la position fodorienne, nous devons rester prudent car selon Fodor les contenus intentionnels des états mentaux se caractérisent par le rôle inférentiel qu'ils sont susceptibles de jouer (en tant qu'attitudes propositionnelles) et non par un rôle causal. Si l'on considère l'attitude propositionnelle :

- (9) *je crois qu'il pleut et que les rues sont mouillées.*

Son contenu intentionnel conjonctif peut être formulé dans la proposition :

- (10) *il pleut et les rues sont mouillées.*

Dans ce cas-là, nous pouvons affirmer que la nature du contenu intentionnel est déterminée — partiellement au moins — par le rôle inférentiel que possède ce contenu intentionnel. On peut inférer ce contenu intentionnel de la croyance qu'il pleut et de la croyance que les rues sont mouillées, et on peut en inférer à la fois qu'il pleut et que les rues sont mouillées. En bref, deux états mentaux qui sont censés avoir le même contenu intentionnel sont également censés avoir le même rôle inférentiel. Ainsi croire que "il pleut **ou** les rues sont mouillées" se distingue de croire que "il pleut **et** que les rues sont mouillées" dans la stricte mesure où ces deux croyances ont un rôle inférentiel distinct. On ne peut pas inférer les mêmes propositions de l'une et de l'autre. Le contenu intentionnel et le rôle inférentiel sont liés de telle sorte que l'on peut, au moins, affirmer la similarité de contenu lorsque l'on constate une similarité de rôle inférentiel.

À ce point, nous ne sommes donc pas encore parvenus à établir un lien entre causalité mentale et contenu intentionnel. Pour ce faire, il faudrait montrer que le rôle inférentiel peut être lié au rôle causal ou réciproquement. Or, comme le remarque Fodor :

“ La notion de rôle inférentiel est après tout une notion sémantique plutôt que causale : le fait que vous puissiez inférer (démonstrativement ou inductivement) une croyance à partir d’une autre, n’a en soi-même aucune implication causale du tout. ” (Fodor, 1987a, p. 318)

Aussi Fodor doit-il exécuter un pas supplémentaire dans le sens de la naturalisation de la psychologie populaire. Pour cela, il lui faut montrer qu’il y a des entités physiques supports (ou correspondants) des états intentionnels dont les propriétés physiques permettent de comprendre le fonctionnement causal du comportement cognitif.

### 3.2. Le Deuxième pas : le fonctionnalisme

De fait, Fodor a cherché à exprimer les propriétés intentionnelles des attitudes propositionnelles dans les termes de la science physique (dans des termes non intentionnels). Pour cela il rejoint un moment la thèse selon laquelle tout état intentionnel peut être réduit à son rôle fonctionnel (cf. 2.2.3., thèse 2), c’est-à-dire aux relations causales qu’il entretient avec d’autres attitudes propositionnelles, et les stimuli d’entrée et les comportements en sortie.

Ainsi, à chaque attitude propositionnelle correspond un état physique du système ( $EP_n$  (figure 4)) tel que les relations intentionnelles des attitudes propositionnelles se réduisent à des relations causales qu’un état physique de l’organisme A entretient avec d’autres états physiques de cet organisme et l’entrée et la sortie.

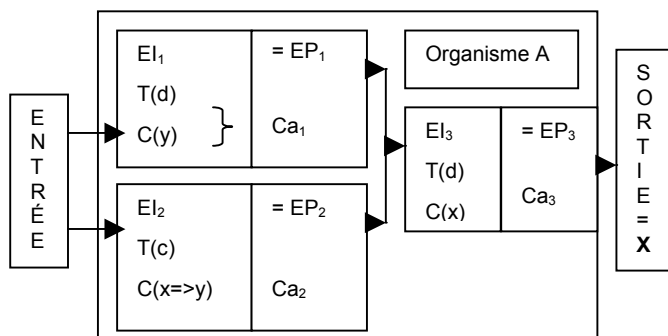


Figure 4 — le fonctionnalisme

- $El_n$  : État intentionnel.
- $EP_n$  : État physique de nature indéterminée.
- $T(n)$  : Type d'état intentionnel (si  $n=d$ , alors  $T(d)=$ désir, si  $n=c$ , alors  $T(c)=$ croyance)
- $C(n)$  : Contenu d'état intentionnel ( $C(x) \Leftrightarrow$  le contenu de l'état intentionnel est  $x$ )
- $Ca_n$  : Ensemble des propriétés causales de l'état physique  $EP_n$  avec les autres états physiques, l'entrée et la sortie.
- $x$  : État de fait ou comportement  $x$ .

À tout état intentionnel  $El_n$ , correspond nomiquement un état physique  $EP_n$  (quel que soit sa nature) dont les propriétés causales ( $Ca_n$ ) sont les corrélats des propriétés inférentielles des états intentionnels repérées à l'étape précédente ( $T(n)$  et  $C(n)$ ). Il faut insister sur le fait qu'ici, il n'est pas encore question de spécifier la nature exacte de ces états physiques ni la façon dont les propriétés physiques et causales de ces états permettent de rendre compte de l'efficacité causale de

nos attitudes propositionnelles sur notre comportement. On s'aperçoit dès lors que cette position demande à être complétée pour répondre aux critiques que nous avons déjà mentionnées (voir 2.2.3.) relatives au problème des “ contenus communs ” des attitudes propositionnelles et à ce que Fodor appelle “ le problème de la constituance ” (Fodor 1987a, p. 311) et pour envisager réellement une réduction de type physicaliste.

Pour mémoire, le problème des contenus communs naît de ce que plusieurs états mentaux peuvent porter sur le même objet de différentes manières (“ vouloir faire du bateau ” et “ redouter de faire du bateau ” par exemple) ce qui ne peut être expliqué si les états mentaux sont des états physiques fonctionnels monadiques (chaque état physique n'étant défini que par référence à ses causes et à ses effets typiques, il ne peut porter sur le même contenu de deux façons différentes : les mêmes causes provoquent les mêmes effets).

Le problème de la constituance, lui relève de la nature inférentielle et syntaxique des attitudes propositionnelles. Comme nous l'avons vu (voir 2.2.3.) la solution envisagée par Fodor consiste à refuser la structure fonctionnelle monadique des états mentaux et d'y substituer une nature relationnelle en termes représentationnels. C'est là la troisième étape de la position fodorienne qui en fait véritablement un représentationnalisme. Cela permet d'après Fodor de s'approcher encore un peu plus de l'objectif poursuivi de naturalisation des explications intentionnelles.

### 3.3. L'étape représentationnelle

La thèse fonctionnaliste est donc précisée par l'adjonction d'une composante représentationnaliste : les états physiques ne sont rien d'autre que des symboles mentaux dotés de propriétés causales ; c'est ce que l'on appelle l'hypothèse du langage de la pensée.

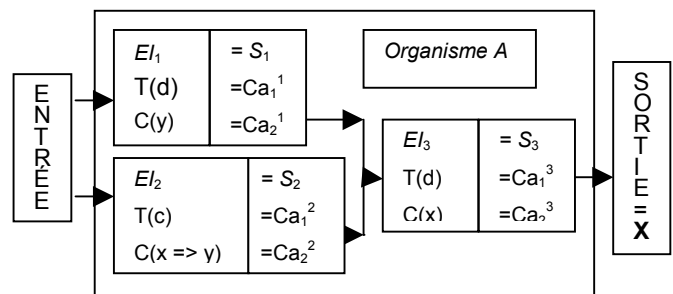


Figure 5 — le représentationnalisme

- $El_n$  : État intentionnel.
- $T(n)$  : Type d'état intentionnel (si  $n=d$ , alors  $T(d)=$ désir, si  $n=c$ , alors  $T(c)=$ croyance)
- $C(n)$  : Contenu d'état intentionnel ( $C(x) \Leftrightarrow$  le contenu de l'état intentionnel est  $x$ )
- $S_n$  : Symbole physique dont la nature est indéterminée et dont les propriétés substantielles sont indéterminées.
- $Ca_1^n$  : Ensemble des propriétés causales de  $S_n$  avec les autres symboles, l'entrée et la sortie.
- $Ca_2^n$  : Ensemble des propriétés causales (et dénotationnelles) de  $S_n$  avec l'objet correspondant au contenu intentionnel.
- $x$  : État de fait ou comportement  $x$ .

Cet amendement de la thèse fonctionnaliste est la pierre angulaire de l'objectif de naturalisation de l'intentionnalisme de Fodor. En effet, selon lui, on ne peut rendre compte de la productivité de l'esprit que si l'on fait l'hypothèse du fonctionnement symbolique de l'esprit. À chaque attitude propositionnelle doit correspondre un état physique de l'organisme qui doit être de nature symbolique. Seuls les symboles ont les propriétés à la fois sémantiques, intentionnelles et physiques nous permettant d'espérer parvenir à une naturalisation des processus mentaux auxquels la psychologie populaire fait référence.

On peut remarquer, au passage, que cela amène Fodor à revenir, *de facto*, sur sa condamnation de la thèse 1 puisqu'il réduit ici les propriétés intentionnelles des représentations mentales (censées être les plus fondamentales) à des propriétés sémantiques et intentionnelles de symboles *linguistiques* (les termes du langage de la pensée). Ce mouvement de réduction semblait pourtant être (dans une certaine mesure) le premier identifié et révoqué par Fodor. L'hypothèse du langage de la pensée à laquelle Fodor semble tenir pose donc un problème de cohérence interne aux thèses développées par celui-ci ; même s'il n'est pas de notre propos de faire ici une critique de la position de Fodor, il semble que ce point mériterait de faire l'objet d'une étude autonome.<sup>9</sup>

En fait la stratégie de réduction des propriétés intentionnelles observables (celles des expressions de la psychologie populaire) à des propriétés intentionnelles non observables (celles des attitudes propositionnelles) se double maintenant d'une réduction de ces phénomènes intentionnels à des propriétés elles-mêmes non-intentionnelles (naturelles) : les propriétés causales des représentations mentales. On voit (cf. figure 5) qu'à chaque propriété des attitudes propositionnelles (contenu et type) correspond une propriété causale ( $Ca_1$  ou  $Ca_2$ ) du symbole mental ( $S_n$ ). Au type de l'attitude propositionnelle correspond la totalité des propriétés causales du symbole  $S_n$  avec les autres symboles et les stimuli d'entrée ou les comportements de sortie ( $Ca_1^n$ ), tandis qu'au contenu de l'attitude propositionnelle correspond l'ensemble des propriétés causales (et/ou dénotationnelles) du symbole  $S_n$  avec l'état du monde correspondant au contenu ( $Ca_2^n$ ). Tout cela revient à affirmer que des symboles physiques peuvent être au fondement de la causalité mentale, mais cela ne nous dit pas comment.

Ce principe représentationnel ne spécifie pas encore la nature du fonctionnement causal (au niveau symbolique) permettant de naturaliser les expressions de la psychologie du sens commun. D'ailleurs, Fodor en est conscient et le relève avant d'entreprendre le dernier pas vers ce qu'il appelle une théorie computationnelle de l'esprit :

*“ Ce que nous voulons (...) c'est une théorie des processus mentaux qui explique comment les relations causales entre les états mentaux peuvent être liées d'une manière ou d'une autre*

*avec les contenus intentionnels de ces derniers. ”*  
(Fodor, 1987a, p. 317)

Il est donc nécessaire de mettre en évidence un mécanisme permettant de lier le rôle causal des états mentaux au rôle inférentiel des contenus intentionnels de ces mêmes états mentaux. Fodor trouve ce mécanisme dans les propriétés syntaxiques des symboles complexes. Ceci lui permet donc de soutenir à nouveau sa thèse de la nature symbolique des états mentaux et vient renforcer sa thèse du “ langage de la pensée ” déjà évoquée plus haut.

### 3.4. La thèse computationnelle

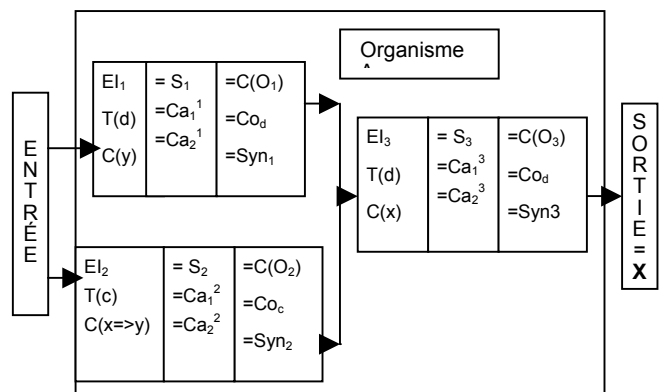


Figure 6 — le Computationalisme  $EI_n$  : État intentionnel.

$T(n)$  : Type d'état intentionnel (si  $n=d$ , alors  $T(d)=\text{désir}$ , si  $n=c$ , alors  $T(c)=\text{croyance}$ )

$C(n)$  : Contenu d'état intentionnel ( $C(x) \Leftrightarrow$  le contenu de l'état intentionnel est  $x$ )

$S_n$  : Symbole physique dont la nature est indéterminée et dont les propriétés substantielles sont indéterminées.

$Ca_1^n$  : Ensemble des propriétés causales de  $S_n$  avec les autres symboles, l'entrée, et la sortie.

$Ca_2^n$  : Ensemble des propriétés causales (et dénotationnelles) de  $S_n$  avec l'objet correspondant au contenu intentionnel.

$Co_d, Co_c$  : Propriétés computationnelles du symbole mental  $S_1$ . (La lettre en indice désigne les relations déterminées par ces propriétés computationnelles : “ d ” pour désir, “ c ” pour croyance, etc.)

$Syn_n$  : Propriétés syntaxiques du symbole mental  $S_n$   
 $C(O_n)$  : Relations causales de  $S_n$  avec un objet  $O_n$ .

$x$  : État de fait ou comportement  $x$ .

L'argumentation de Fodor pour l'introduction de la thèse computationnelle peut être résumée de la façon suivante. On peut constater — dans le cas des langages symboliques possédant une structure syntaxique interne — que la syntaxe est liée à certaines propriétés physiques des signes composant le symbole complexe qui elles-mêmes déterminent son rôle à la fois inférentiel et causal.

C'est-à-dire qu'un symbole complexe constitué suivant des règles de syntaxe est constitué de différents symboles dont les caractéristiques physiques (telles que la configuration géométrique qui intervient

<sup>9</sup> Notons que ce point a fait l'objet d'une intervention de J.-M. Roy lors du Twentieth World Congress of Philosophy, Boston, Août 1998 : *Cognitive turn and linguistic turn*.

essentiellement dans la composition syntaxique du symbole complexe) déterminent le rôle causal et inférentiel de chaque élément.

Les processus mentaux doivent être de nature computationnelle ou calculatoire pour que nous comprenions comment le contenu intentionnel des états mentaux est lié aux relations causales que l'on suppose communément entre les états mentaux. Dire que ces processus sont computationnels revient à dire qu'ils correspondent à des séquences de manipulations de symboles qui sont gouvernées par des règles portant sur la structure interne des symboles manipulés. Ainsi, c'est seulement dans la mesure où des symboles sont syntaxiquement organisés que leurs rôles inférentiels peuvent être liés à leurs rôles causaux. Nous pouvons donc établir un lien entre le rôle inférentiel du contenu intentionnel des états mentaux et leur rôle causal si ces états mentaux sont des symboles ou des représentations mentales d'un "langage de la pensée". Ainsi, de même qu'un ordinateur (du type machine de Turing) exécute des opérations physiques en fonction des propriétés computationnelles et syntaxiques des symboles qui composent à la fois sa ligne de commande et sa ligne d'entrées ; on doit pouvoir réduire les explications causales du comportement d'un organisme intelligent à des opérations de calcul sur des symboles mentaux.

On voit ici (cf. figure 6) que les états intentionnels ( $EI_n$ ) dont on a fait l'hypothèse dans l'étape intentionnaliste sont réduits à des symboles physiques ( $S_n$ ). Chaque symboles est causalement relié à l'objet du contenu de l'état intentionnel correspondant ( $C(O_1)$  par exemple est lié à  $y$ ). Pour le dire plus clairement, si mon état intentionnel est le désir d'un café, il faut que le symbole physique correspondant à ce désir ait un lien causal avec un certain café (passé, potentiel, imaginé, etc.). Cela vient illustrer au niveau physique ce que l'on trouve au niveau représentationnel à savoir  $Ca_2^n$ , c'est-à-dire des propriétés causales et dénotationnelles entre la représentation mentale et l'objet de celle-ci. De plus, chaque représentation mentale doit posséder des propriétés syntaxiques ( $Syn_n$ ) qui lui confèrent les propriétés computationnelles ( $Co_c$  et  $Co_d$ ) nécessaires pour expliquer l'efficacité causale de l'état intentionnel sur le comportement et d'autres états intentionnels.

Fodor parvient ainsi à rendre compte de la causalité mentale c'est-à-dire de l'efficacité causale des attitudes propositionnelles sur notre comportement en réduisant les propriétés intentionnelles des attitudes propositionnelles à des propriétés syntaxiques et computationnelles. Nous avons donc d'un côté des lois psychologiques qui sont intentionnelles et de l'autre des processus mentaux qui sont des processus computationnels.

Mais ces processus computationnels ne nous permettent pas de penser que la nature du contenu intentionnel des attitudes propositionnelles influe sur leur type. Fodor disjoint ce qui semble lié dans la psychologie ordinaire : le contenu et le type de l'attitude propositionnelle. Au contraire, on est frappé par le parallélisme qui semble interdire à tout niveau de l'explication proposée par Fodor un rapport causal entre type et contenu de l'état intentionnel. Pourtant c'est bien seulement parce qu'il s'agit d'épinards et que je n'aime pas cela que je ne désire pas en manger. Ce lien entre

contenu et type de l'état intentionnel est mis en évidence dans l'explication de la psychologie ordinaire mais Fodor l'élimine de sa TCRE.

Nous voudrions montrer que Locke, quelques siècles auparavant a fait face à cette question, et que sa réponse, loin d'être aussi simpliste que de nombreux commentateurs ont voulu le croire, peut être l'occasion d'une réflexion renouvelée sur la causalité mentale.

#### **4. Représentationnalisme et causalité mentale chez Locke — Essai concernant l'entendement humain**

Il faut noter que l'objet de l'EEH<sup>10</sup> n'est pas le même que celui des recherches de Fodor. En effet, Locke tente d'établir que l'homme est capable de connaître de façon certaine la loi naturelle, c'est-à-dire l'ensemble des lois et règles morales qui s'imposent à l'homme par nature. Cet objectif critique conduit Locke à s'interroger sur les facultés cognitives de l'être humain en des termes nouveaux. Son empirisme le contraint en effet à montrer que toute connaissance provient en dernière analyse de la connaissance sensible. Or, il affirme simultanément que toute connaissance est générale et que le réel n'est constitué que de particulier. Ainsi est-il astreint à chercher comment nous produisons la généralité de la connaissance. Sa réponse est simple : la généralité est dans les signes. Et les signes les plus utilisés sont les mots. Mais pour qu'un mot ait un sens, il faut qu'il puisse être attaché régulièrement au même référent. Chaque référent étant particulier, il est nécessaire d'avoir un intermédiaire entre la sensation toujours particulière et le mot toujours général. Cet intermédiaire sera l'idée ou la représentation mentale. Dès lors, la question de Locke sera de comprendre comment des idées sont capables de causer notre action par rapport à la connaissance de la loi naturelle. Et nous retrouvons ici le point de vue intentionnaliste qui nous intéresse. Locke ne cherche donc pas à naturaliser l'intentionnalité mais examine comment notre esprit parvient à déterminer son action, dans l'ordre de la pensée ou dans l'ordre du mouvement corporel. Aussi la question de l'efficacité causale de nos états intentionnels est-elle examinée par Locke à l'occasion de son enquête sur la liberté humaine.

##### **4.1. La distinction entre contenu et type de l'état intentionnel dans l'EEH**

Locke n'identifie jamais explicitement ni la notion d'état intentionnel ni celle de représentation mentale, mais utilise indifféremment pour les deux le terme d'idée. Ainsi, le terme d'idée désigne-t-il aussi bien l'attitude propositionnelle : "je veux une montagne d'or" que la représentation mentale <montagne d'or>. On estime que ceci est le signe qu'il n'a pas clairement distingué le contenu et le type de l'attitude propositionnelle. Pourtant, il semble bien réserver des traitements différents dans des passages distincts de l'EEH à ce que nous dénommons contenu et type de l'état intentionnel.

En effet, le chapitre VIII du livre II et le chapitre III du livre III, présentent la nature et la genèse des contenus des attitudes propositionnelles. Locke montre en effet

---

<sup>10</sup> Nous abrègerons ainsi désormais *L'Essai philosophique concernant l'Entendement Humain* de Locke.

quelle mécanique est à l'origine de la naissance de nos idées (au sens de représentation mentale) en faisant appel à ce que nous pourrions appeler aujourd'hui une théorie de la covariation (nous laissons à des recherches ultérieures cet intéressant aspect de la thèse représentationnelle lockéenne). Ses développements sur le plaisir et l'inquiétude — ou malaise (*uneasiness*) — dans le livre II chapitres VII, XX et XXI et la probabilité (dans le livre IV, chapitre XV) s'inscrivent dans une description du processus causal de détermination du type de l'attitude propositionnelle à partir des qualités du contenu de l'attitude propositionnelle. Cette étude dissociée du contenu et du type des attitudes propositionnelles montre bien que Locke les distingue — au moins de façon implicite. Mais plus encore il montre que Locke a tenté de les lier causalement de façon formelle dans les processus de choix d'actions.

#### 4.2. Locke distingue deux genres de types d'attitudes propositionnelles

Selon Locke, on peut distinguer deux grands types d'attitudes propositionnelles en vertu des "sentiments" qui accompagnent régulièrement les contenus intentionnels de ces attitudes propositionnelles. Cette distinction permet à Locke de ne pas réduire les processus mentaux à l'arbitraire du principe d'association des idées, mais d'élaborer une "mécanique des idées" qui explique comment nos actions sont déterminées par les contenus de nos attitudes propositionnelles. D'après Locke, ce sont les sentiments de plaisir, d'*uneasiness*, et de probabilité qui jouent le rôle de "moteurs causaux" pour les types d'attitudes propositionnelles. Le type de l'attitude propositionnelle serait causé par des qualités du contenu. Ainsi, la distinction type/contenu que nous avons identifiée dans la TCRE de Fodor n'est pas *stricto sensu* identique à celle de Locke. En d'autres termes, s'il y a bien chez Locke comme chez Fodor une distinction entre type et contenu des états mentaux, cette distinction ne prend pas la forme, dans l'EEH, d'une séparation de deux ordres causaux parallèles. C'est ce point précis que nous souhaitons mettre en lumière.

##### 4.2.1. La première espèce de type d'attitudes propositionnelles : les volitions

La question de la causalité mentale est d'abord abordée par Locke dans le livre II, chapitre XXI (de la puissance) où il essaye de montrer la liberté de l'homme. Pour lui, être libre c'est être :

*"capable d'agir ou de ne pas agir en conséquence de notre choix ou volition."* (Locke, EEH, II, XXI, § 27)

Mais pour que cette définition de la liberté ait un sens, il faut préciser quel est le fonctionnement des volitions. Selon Locke, le plaisir et la peine ou le contentement et l'*uneasiness* accompagnent quasiment toutes nos idées (qu'elles proviennent des sens ou de la réflexion) : de ce fait, c'est conformément à l'un de ces sentiments joint à une idée que nous *désirons* ou *redoutons* ce qu'elle représente ou que nous *espérons* ou la *regrettons*.

Locke explique en effet que :

*"Le plaisir ou le malaise<sup>11</sup> (uneasiness) sont, l'un ou l'autre, joint à presque toutes nos idées, qu'elles proviennent de la sensation ou de la réflexion ; et il n'y a quasiment aucune affection des sens par l'extérieur ni aucune pensée dans notre esprit qui ne soit pas susceptible de produire en nous du plaisir ou de la peine. (...) L'auteur de notre être dans son infinie sagesse nous a donné la puissance sur plusieurs parties de notre corps ; de les mouvoir ou de les tenir au repos comme nous le voulons, et par leurs mouvements de nous mouvoir ainsi que les corps qui nous sont contigus, en quoi consistent toutes les actions de notre corps. Il a aussi accordé à notre esprit le pouvoir de choisir, en différentes occasions, parmi ses idées celle dont il veut faire l'objet de ses pensées (...). Afin de nous porter à ces actions de mouvement ou de pensée, il a eu la bonté d'attacher une perception de plaisir à plusieurs de nos pensées et de nos sensations. Si cette perception était pleinement séparée de toutes nos sensations externes et de nos pensées internes, nous n'aurions aucune raison de préférer une pensée ou une action à une autre ; de préférer la négligence à l'attention ou le mouvement au repos. Ainsi nous ne mettrions pas notre corps en mouvement, pas plus que nous n'emploierions notre esprit ; mais nous laisserions (si je puis employer cette expression) dériver nos pensées sans aucune direction ou dessein, et nous subirions les idées de notre esprit comme de vaines ombres y apparaissant sans que nous y fassions attention."* (Locke, EEH, II, VII, §2-3. nous traduisons.)

Ainsi, le plaisir et la peine sont-ils les moteurs causaux de notre activité. Nous redoutons quelque chose dont l'idée provoque en nous du malaise tandis que nous désirons quelque chose qui provoque en nous du plaisir. Locke ajoute :

*"Qu'est ce qui détermine la volonté ? Le sens de cette question se réduit à ceci. Qu'est ce qui pousse l'esprit, dans chaque occasion particulière, à déterminer sa puissance de direction à ce mouvement ou à ce repos particulier plutôt qu'à un autre ? À quoi je réponds que le motif qui nous pousse à persister dans le même état ou dans la même action est seulement la satisfaction présente en l'un ou en l'autre. Le motif qui pousse au changement est toujours quelque "malaise" [uneasiness] ; rien ne nous portant au changement d'état ou à une nouvelle action sinon quelque "malaise"."* (Locke, EEH, II, XXI, §29. nous traduisons.)

Ainsi, c'est parce que l'on admet qu'une propriété du contenu d'une attitude propositionnelle est cause du type de cette attitude propositionnelle que l'on peut comprendre comment le contenu détermine le type de l'attitude propositionnelle. Cette idée, d'ailleurs, semble assez convaincante dans la mesure où il s'agit, pour

<sup>11</sup> L'anglais donne *uneasiness* que Coste traduit par "inquiétude" resté classique. Mais ce terme ne rend pas l'idée d'une position intenable. C'est pourquoi nous lui préférons le mot à mot "malaise" même si le sens français en est différent.

tous ces types déterminés par le plaisir ou le déplaisir accompagnant le contenu intentionnel, de types qui appartiennent à la classe des volitions (désirer, redouter, souhaiter, etc.), et dans la mesure où vouloir quelque chose suppose bien une certaine propriété dans la chose qui fait qu'elle me semble désirable.

*“Ainsi, en réfléchissant sur la pensée que nous avons du plaisir qu'une chose présente ou absente peut produire en nous, nous avons l'idée de ce que nous appelons aimer. Car lorsqu'un homme déclare en automne, lorsqu'il mange du raisin ou au printemps, lorsqu'il n'y en a pas, qu'il aime cela ; il ne dit rien de plus que le goût du raisin lui procure du plaisir. Mais si une altération de sa santé ou de sa constitution anéantit le plaisir qu'il prenait à leur goût, on ne pourra plus dire de lui qu'il aime le raisin.”* (Locke, EEH, II, XX, §4. nous traduisons.)

Si une attitude propositionnelle est caractérisée par le type de l'amour, c'est dû au plaisir qui accompagne actuellement l'idée constituant le contenu de cette attitude propositionnelle. Ce plaisir naît de la relation particulière d'une qualité physique de l'objet dont nous avons la sensation ou l'idée avec nos organes sensoriels. Locke élabore donc une position intentionnaliste complexe reposant sur une liaison intime entre des états intentionnels (vouloir, désirer, rechercher, redouter, etc.) et des états qualitatifs (plaisir, douleur, *uneasiness*, contentement, etc.) eux-mêmes reliés aux contenus intentionnels de ces états intentionnels.

Cette originalité de la position de Locke doit d'abord attirer notre attention sur la difficulté qu'a toujours constituée la notion d'état qualitatif pour les théories représentationnelles de l'esprit (au premier rang desquelles figure celle de Fodor). Pierre Jacob note par exemple :

*“La thèse représentationnaliste semble s'appliquer aux attitudes propositionnelles beaucoup mieux qu'aux sensations. (...) L'un des défis auxquels doit répondre la thèse représentationnaliste est donc de montrer qu'une théorie des propriétés sémantiques des attitudes propositionnelles peut contribuer à la compréhension du problème soulevé par l'existence des sensations conscientes.”* (Jacob, 1997, chap. 1)

Le problème ici soulevé par Locke est le suivant. Il se trouve que nous avons des attitudes propositionnelles dont les types sont caractérisés par certains états qualitatifs accompagnant régulièrement le contenu de cet état intentionnel. Ainsi Nino Ferrer aurait-il peut-être raison de relever dans sa chanson *Madame Robert*<sup>12</sup> :

*“Mon frère n'aim' pas les épinards.  
Et c'est heureux pour mon frère car,  
S'il les aimait, il en mangerait,  
Or il ne peut pas les supporter...”*

Il y a manifestement un lien entre des propriétés qualitatives de ma relation à certains objets et le type

d'attitude propositionnelle que je peux avoir à propos de cet objet. Mais si Locke affirme apparemment une sorte de mécanique rigide de la détermination du type de l'état propositionnel par des propriétés qualitatives de son contenu (ou de notre rapport à l'objet distal correspondant à ce contenu), il ne peut plus penser la variation de type d'attitude propositionnelle pour un même contenu.

Dans l'orthodoxie fodorienne cet écueil est évité par le fait que le contenu n'a pas d'influence directe (ou mécanique) sur le type de l'état intentionnel. Le type de mon attitude propositionnelle au sujet de ce contenu est le résultat de l'intervention d'autres attitudes propositionnelles. Cela permet d'expliquer pourquoi certains d'entre nous désirent regarder les rencontres de football tandis que d'autres non, ou pourquoi à un moment donné je désire regarder une rencontre alors qu'à un autre je le regrette... C'est le système d'attitudes propositionnelles sous-jacent du sujet qui explique ces variations de types pour les mêmes objets intentionnels et donc la relativité ou la versatilité de nos désirs ou de nos volontés.

On pourrait ici reprocher à Locke d'ignorer ce problème et de s'interdire de penser les revirements de type d'attitude propositionnel quant à un même objet. Les deux derniers passages cités nous semblent cependant offrir un aperçu d'une alternative sinon valable du moins intéressante à la position fodorienne.

En effet, cette position nous semble poser un problème fondamental : si nous avons un certain type d'attitude mentale face à un objet seulement en fonction d'autres croyances, nous ne pouvons jamais poser de terme fixe à l'origine de nos croyances ou de nos désirs. Une sorte de régression à l'infini s'impose d'elle-même et nos explications du sens commun deviennent proprement inopérantes car, autant que les descriptions opérationnelles des behavioristes logiques, nous ne pouvons jamais terminer notre explication. Si l'on peut montrer que le type de certaines de nos attitudes propositionnelles est déterminé par des états qualitatifs produits par les objets eux-mêmes, alors nous pouvons espérer mettre un terme à nos explications causales du comportement et ainsi “sauver les phénomènes”.

S'il est vrai, comme le montre l'expérience, que je peux à un moment aimer le raisin (qu'il y en ait ou pas à ma disposition), c'est-à-dire que le raisin (comme collection de sensations ou comme contenu d'un état mental) provoque en moi du plaisir et me porte à en désirer et, à un autre moment ne pas aimer le raisin et donc ne pas le désirer — voire regretter son occurrence. C'est, prétend Locke, seulement dans la mesure où ma conformation physiologique entre le premier moment et le deuxième a changé, ma relation gustative, par exemple, peut être altérée (par une maladie, l'absorption d'un produit au goût très prononcé, etc.) et m'amener à craindre ce que je désirais auparavant.

Cette introduction d'une possibilité de modification des qualités subjectives causées par un même objet à différents sujets ou à un même sujet à différents moments permet à Locke d'envisager une réponse originale au problème de la variation des types des attitudes propositionnelles. Si elle ne permet pas, sans doute, la plasticité explicative de la position fodorienne, elle a l'avantage d'expliquer certains cas laissés pour

<sup>12</sup> Nino Ferrer / G. Verlor, *Madame Robert*, (p) 1966 Barclay, éd. Beuscher.

compte par celle-ci. En effet, si je n'aime pas les épinards ce n'est pas un corps d'attitudes propositionnelles qui explique le fait que je redoute d'avoir à en manger lorsque je suis convié à un dîner. C'est bien le déplaisir de la sensation gustative (et éventuellement le déplaisir de l'idée d'avoir à éprouver cette sensation) suscitée en moi qui en est la cause.

L'efficacité de la stratégie de Locke semble à première vue limitée aux relations entre états qualitatifs et types des attitudes propositionnelles ayant pour contenu les objets causant ces sensations. Cependant, un des aspects du propos de Locke consiste à dire que cela vaut aussi pour les états qualitatifs suscités non par des qualités de l'objet lui-même mais par des qualités de sa représentation mentale ou de son idée elle-même. Il s'agit d'éclairer ce point car si la position de Locke à ce sujet était viable, elle offrirait manifestement une alternative élégante à la position de Fodor.

Pour cela il est nécessaire de se rappeler que pour Locke toute idée s'accompagne nécessairement de conscience et plus précisément encore de la conscience de l'effet qualitatif qu'elle provoque en mon esprit. Selon lui, avoir une idée peut provoquer plaisir, douleur, ou *uneasiness* rappelons le passage déjà évoqué plus haut :

*“ Le plaisir ou l'inquiétude (uneasiness) sont, l'un ou l'autre, joint à presque toutes nos idées, qu'elles proviennent de la sensation ou de la réflexion ; et il n'y a quasiment aucune affection des sens par l'extérieur ni aucune pensée dans notre esprit qui ne soit pas susceptible de produire en nous du plaisir ou de la peine. Par plaisir et par peine j'entends signifier tout ce qui nous plaît ou nous incommode, que cela procède des pensées en notre esprit ou d'une chose qui agit sur notre corps. ”* (Locke, EEH, II, VII, §2.; nous traduisons et soulignons)

Ainsi, parce que nos représentations mentales s'accompagnent de la conscience du plaisir ou du déplaisir qu'elles provoquent en nous, nous pouvons comprendre l'affirmation lockéenne d'une influence du contenu sur le type de l'attitude intentionnel. Cette explication, aussi éloignée qu'elle soit de la position de Fodor montre cependant qu'on peut donner une réponse au problème de la variation du type des attitudes propositionnel qui permette de prendre en compte tous les cas de figure (lien entre qualités de l'objet et type ou entre qualités de la représentation mentale et type).

On voit que ce qui distingue ici la position de Locke de celle de Fodor tient aux différences qu'entretiennent leurs projets épistémologiques respectifs. Locke est amené à décrire et à expliquer les processus mentaux de détermination causale du type de l'attitude propositionnelle parce qu'il cherche à déterminer les limites de la connaissance humaine, et donc les conditions épistémologiques de la connaissance vraie. Cela doit lui permettre à terme de décrire les règles normatives de la pensée juste (comme le montrent bien certains passages du dernier livre de l'EEH : (Locke, EEH, IV, VIII : “ Des propositions frivoles ” [*trifling* en anglais, c'est-à-dire “ *of slight importance or little value : insignificant* ”]). Pour décrire ces règles permettant de distinguer les idées fausses des vraies, les idées bien formées des idées absurdes (et donc les actions justes

des injustes, les actions droites des confuses) il lui faut assurer une certaine objectivité et une certaine régularité des choix et des comportements des hommes. Au contraire, Fodor cherche à rendre compte de ce qui est et non à établir des règles de la pensée en rapport avec la nature et les limites des capacités de connaissance des hommes. Il cherche à élaborer une théorie intentionnaliste “ naturalisée ” du comportement cognitif, non à déterminer les “ bonnes ” façons de penser étant donnée la nature de l'esprit humain.

#### 4.2.2. La deuxième espèce de type d'attitudes propositionnelles : les assertions

Le problème reste que certains états intentionnels ne peuvent être déterminés par des perceptions de plaisir ou de peine. Les types d'attitudes propositionnelles tels que : “ croire que ”, “ douter que ”, “ estimer que ”, semblent se distinguer des volitions.

Aussi, Locke élabore-t-il, à l'occasion de sa réflexion concernant la probabilité, un ensemble de thèses qui décrit le processus de détermination du type de l'attitude propositionnelle par la propriété de probabilité de la vérité du contenu intentionnel. D'après Locke, si l'objet d'une pensée est peu probable, nous allons en douter, tandis que s'il est très probable nous allons lui accorder notre assentiment, croire en le contenu de cette attitude propositionnelle. Le chapitre XVI du livre IV développe la question des degrés d'assentiment que nous devons accorder à une idée en fonction de son degré de probabilité. Nous avons donc là un indice de plus tendant à montrer que c'est une théorie des types d'attitudes propositionnelles présentant des aspects normatifs importants que fournit ici Locke :

*“ Les fondements de la probabilité que nous avons proposés dans le chapitre précédent étant la base sur laquelle est construit notre assentiment, ils sont aussi la mesure par laquelle les différents degrés de l'assentiment sont ou doivent être réglés. ”* (Locke, EEH, IV, XVI, §1.; nous traduisons, souligné par Locke.)

Locke insiste en fait sur un bon usage de l'assentiment. Le but poursuivi est double : d'une part limiter les possibilités de fausses connaissances (connaissances admises pour vraies alors qu'elles sont improbables), et d'autre part réserver un type d'assentiment différent pour les connaissances révélées de la religion (dans lesquelles le degré de probabilité n'intervient pas) et pour la connaissance des “ *matters of facts* ” (Locke, EEH, IV, XVI, §12).

De ce fait, les types d'attitudes propositionnelles qui sont déterminés par la probabilité sont seulement la croyance, le doute, l'assurance, la conjecture, l'assentiment, l'estimation et tout autre type d'attitude propositionnelle relatif à la valeur de vérité que nous accordons au contenu de cette attitude propositionnelle. Ici, comme dans le cas des types d'attitudes propositionnelles déterminés par le plaisir ou l'*uneasiness*, la croyance est déterminée par une propriété du contenu intentionnel de l'attitude propositionnelle.

Ici une objection formelle assez forte semble pouvoir être formulée à l'encontre de la position de Locke : il ferait appel à la probabilité comme à une qualité objective de l'objet de ma croyance ou de mon doute

alors qu'elle n'est en fait qu'une croyance que le sujet formule à propos de l'objet considéré. Il y aurait ici une ambiguïté entre probabilité subjective et probabilité objective. La critique consiste à faire remarquer que la probabilité dont parle Locke est la probabilité subjective que j'accorde à l'objet de ma croyance et que cette probabilité n'est rien d'autre qu'une autre de mes croyances sur ce contenu.

De ce point de vue là, Locke serait plus fodorien que dans son explication du lien entre contenus et types volitifs des attitudes propositionnelles. Il rejoindrait Fodor sur l'idée que ce qui fait le lien du contenu au type de l'état mental n'est rien d'autre que les croyances que je possède sur ce contenu.

Examinons de plus près ce que dit Locke à propos de la probabilité et à propos de la façon dont elle arrache notre assentiment pour déterminer si, oui ou non, la détermination de notre croyance ou de notre doute repose sur des propriétés du contenu et non sur d'autres croyances. Il faut faire une remarque préliminaire qui jette d'ores et déjà une certaine lumière sur ce que nous avons dit. Cette remarque est en fait une citation d'un passage de la lettre sur la tolérance de Locke :

*“ Croire que ceci ou cela est vrai n'est pas dans les limites du pouvoir de la volonté. ”* (Locke, 1968, p. 121. nous traduisons.)

L'idée de Locke est donc que nous ne pouvons pas choisir en quoi nous croyons. Nous croyons seulement lorsque nous accordons une valeur vraie à une proposition eu égard à certains motifs (*enducements*). Ces motifs sont de deux ordres : la conformité avec nos propres connaissances (tirées des vérités révélées ou des démonstrations), observations et expériences et les témoignages des autres appuyés sur ce qu'ils connaissent ou ce qu'ils ont expérimenté.

*“ Une proposition est en elle-même plus ou moins probable selon que la conformité de notre connaissance, que la certitude de nos observations que la fréquence et la constance de nos expériences et que le nombre et la crédibilité des témoignages s'accordent plus ou moins avec elle. ”* (Locke, EEH, IV, XV, §1)

Ici, la proposition est réputée par Locke être en elle-même plus ou moins probable en fonction des différents indices ou motifs que nous avons de lui reconnaître cette probabilité.

L'objection pourrait être poursuivie en faisant remarquer que ces motifs invoqués par Locke ne sont rien d'autre que des objets auxquels je dois accorder mon crédit. En d'autres termes, cela revient à dire que la crédibilité d'une chose n'est pas une qualité de la chose mais, en fait, un corpus de croyances à propos de l'objet intentionnel lui-même. Dès lors, nous tomberions dans une régression à l'infini du même type que celle repérée plus haut mais appliquée aux croyances et non aux volitions. À la question “ pourquoi crois-tu que la terre est ronde ? ” Je ne peux pas donner de réponse objective, car je dois toujours remonter de croyance en croyance sans pouvoir sortir de moi-même. Cette régression est un risque majeur pour la valeur des explications causales du comportement de la psychologie populaire. Si nous sommes condamnés à faire reposer une croyance sur d'autres croyances et

ainsi de suite, nous donnons raison aux antiréalistes car sans ancrage dans des qualités des choses, nous ne pouvons pas espérer naturaliser ces explications dans toutes leurs dimensions.

Si cette thèse de Locke n'est pas une solution viable en elle-même (l'indistinction entre probabilité objective et probabilité subjective est patente et problématique), elle a au moins le mérite de pointer un problème fondamental pour toute tentative de naturalisation de la psychologie.

## 5. Conclusion ou pourquoi revenir aux doctrines classiques

Ainsi, Locke distingue-t-il deux grandes familles de types d'attitudes propositionnelles : les unes portent sur l'occurrence effective de l'objet d'une sensation ou d'une pensée qui nous procure du plaisir ou de la peine (toutes ces attitudes propositionnelles sont d'après lui des espèces de volitions : aimer, vouloir, désirer, souhaiter, craindre, redouter...), les autres concernent la valeur de vérité à accorder aux contenus des attitudes propositionnelles (ces attitudes propositionnelles sont donc dépendantes du degré de probabilité de la vérité du contenu de l'attitude propositionnelle : croire, estimer, douter, admettre, etc.). Dans le premier cas, c'est le plaisir ou la peine accompagnant la pensée ou la sensation qui détermine le type de l'attitude propositionnelle, et dans le second, c'est la probabilité qui le détermine. Dans l'un la solution qu'il envisage nous semble viable, dans l'autre, non.

## 6. Conclusion

Plus largement, on se rend compte qu'en liant comme il le fait la nature du type de l'attitude propositionnelle à des propriétés du contenu intentionnel de celle-ci, Locke permet de penser de façon réaliste la causalité mentale. Cependant, cette voie peut nous apparaître comme condamnée d'avance dans la mesure où elle suppose que tout état intentionnel devrait être conscient. La thèse de Locke n'est donc pas exempte de tout problème et elle ne saurait être absolument valable. Mais au bout du compte, il nous révèle à la fois un risque majeur d'échec de l'entreprise de naturalisation de la psychologie populaire et des pistes de recherche permettant d'éliminer ce risque.

Reste à suivre ces pistes en procédant à une analyse plus précise des problèmes révélés et en approfondissant les stratégies évoquées à la lumière des progrès des sciences cognitives.

## Références bibliographiques

[Brentano, 1874] Brentano F. (1974). *Psychologie vom Empirischen Standpunkt*. 3 vol. Leipzig. Felix Meiner Verlag. Trad. franç. De Gandillac M. (1944). *Psychologie du point de vue empirique*. Aubier-Montaigne. Paris.

[Carnap, 1936-1937] Carnap R. (1936-937). *Testability and Meaning*. In *Philosophy of science* n°3/1936 419-471 et n°4/1937 1-40.

[Churchland, 1984] Churchland P. M. (1984). *Matter and Consciousness* MIT Press / Bradford Books. Cambridge, Mass.

**[Dennett, 1987]** Dennett D. (1987). The intentional stance. MIT Press / Bradford Books, Cambridge Mass. 6<sup>th</sup> ed.

**[Dennett, 1990]** Dennett D. (1990). Quining Qualia. In *Mind and Cognition*. W. Lycan, ed., Blackwell. Oxford. 519-548.

**[Fodor, 1986]** Fodor J. (1986). Banish Discontent.. In Butterfield J. éd. *Language, Mind and Logic*. Cambridge University Press. Repris In Lycan W. éd. *Mind and Cognition*. Blackwell, Oxford. 420-438.

**[Fodor, 1987a]** Fodor J. (1987a). Mental representation : an introduction. In N. Resher (Ed.) *Scientific inquiry in philosophical perspective*. University Press of America. Trad. Franç. Roy J-M (1992). Introduction au problème de la représentation mentale, In *Les études Philosophiques*, n°3, 301-322.

**[Fodor, 1987b]** Fodor J. (1987). Psychosemantics. MIT Press / Bradford Books. Cambridge Mass.

**[Hume, THN]** Hume D. (1739). A Treatise of Human Nature. L-A Selby-Bigge and P-H Nidditch (Eds.). Oxford University Press. Oxford. 30<sup>th</sup> ed. (1994).

**[Horgan et Woodward, 1985]** Horgan T., et Woodward J. Folk (1985). Psychology is here to stay. In *The Philosophical Review*. XCIV, n°2. Repris In Lycan W. (1990). Éd. *Mind and Cognition*. Oxford. Blackwell. 399-420.

**[Jacob, 1997]** Jacob P. (1997). Pourquoi les choses ont-elles un sens ? Odile Jacob, Paris.

**[Locke, 1968]** Locke J. (1968). Letter on Toleration. Klibansky R. et Gough J. (Ed.). Oxford University Press. Oxford.

**[Locke, EEH]** Locke, (1689 - 1710). An Essay concerning Human Understanding. Edited with an Introduction and index by P. Nidditch. (1975) Clarendon Press. Oxford. Trad. Franç. Coste (1700) Essai philosophique concernant l'Entendement Humain. Réimpr. Vrin. Paris.(1972).

**[Nagel, 1974]** Nagel T. (1974). What is it like to be a Bat? *Philosophical Review*, 83, 435-456.

**[Ryle, 1949]** Ryle G. (1949). The concept of Mind. Hutchinson & Co, Londres.

**[Stich, 1978]** Stich S. (1978). Autonomous Psychology and the Belief-Desire Thesis. In *Monist*. n°61. 573-591. Repris In Lycan W. éd. (1990). *Mind and Cognition*. Blackwell. Oxford. 345-361.

**[Tye, 1986]** Tye M. (1986). The Subjective Qualities of Experience. *Mind* n°95/1986. 1-17.

**[Wittgenstein, 1958]** Wittgenstein L. (1958) The blue and the brown books. Blackwell. Oxford. Trad. franç. (1962) : Le cahier bleu et le cahier brun. Gallimard, Paris.