

Les expériences de Burge et les contenus de pensée

Michel SEYMOUR*

Résumé

Les expériences de pensée de Tyler Burge révèlent une tension entre deux sortes d'arguments visant à prouver la détermination de l'environnement sur les contenus d'états intentionnels. On se propose dans un premier temps de montrer que l'expérience de 1979 est plus fondamentale et moins compromettante du point de vue métaphysique, et qu'il faut en ce sens insister davantage sur les déterminations de l'environnement social plutôt que de l'environnement physique. On fait valoir ensuite avec Loar que l'expérience de 1979, comme telle, n'a pas d'incidence sur les contenus de pensée tels qu'individués dans la psychologie cognitive et n'affecte que les attributions d'attitudes dans la psychologie populaire. Il est démontré enfin que l'admission d'une théorie citationnelle des contenus de pensée renforce l'argument de Burge, permet de contourner l'objection de Loar, et résout plusieurs des difficultés auxquelles était confrontée la version initiale de l'expérience.

Les expériences de pensée («thought experiments») récemment proposées par Tyler Burge dans plusieurs articles ont un objectif commun. Il s'agit de montrer que les états mentaux ne sont pas individués de façon individualiste et qu'ils sont, bien au contraire, individués en relation avec l'environnement social ou physique. Ces expériences de pensée peuvent être rapprochées de l'expérience des terres jumelles de Putnam au sens où ce dernier cherche à montrer que les significations ne sont pas individuées de façon individualiste¹. Mais une des prémisses dans l'argument de Putnam implique que, d'une terre à l'autre, les individus et leurs répliques ont les mêmes états psychologiques parce qu'ils sont dans les mêmes états fonctionnels². Une fois que ces états fonctionnels ont été mis en corrélation avec certains états neuro-physiologiques, les états mentaux prennent alors la forme de propriétés internes de l'individu et sont donc individués de façon individualiste. Vu sous cet angle, les expériences de Burge vont plus loin que celles de Putnam puisque ce sont dé-

* Université de Montréal, Canada.

¹ Putnam, Hilary, «The Meaning of Meaning», dans *Philosophical Papers*, vol. 2, *Mind, Language and Reality*, Cambridge, Cambridge University Press, 1975, 215-271.

² Putnam, Hilary, «The Meaning of Meaning», ibidem, 224.

sormais les états mentaux eux-mêmes et non seulement les significations qui ont un caractère public.

Je voudrais dans un premier temps me prononcer sur l'importance relative à accorder à ces diverses expériences de pensée. Après avoir imposé une certaine perspective d'ensemble, je ferai la suggestion qu'elles doivent être comprises comme portant au sujet des *attribution*s d'attitudes psychologiques («mental ascriptions») et non au sujet des attitudes elles-mêmes. J'essaierai ensuite de répondre à l'objection de Brian Loar qui est, à mon avis, l'objection la plus importante qui ait été adressée à Burge³.

Loar profite justement du fait que ces expériences n'ont qu'une incidence sur les *attribution*s d'attitudes et n'affectent que la psychologie populaire. La nature intrinsèque des attitudes étant sous la responsabilité de la psychologie cognitive et non de la psychologie populaire, il serait toujours possible, selon Loar, d'accorder un droit de cité à l'individualisme en psychologie cognitive tout en reconnaissant le bien-fondé des expériences de Burge. Je chercherai à répondre à cet argument en prétendant que les contenus des états mentaux sont de nature linguistique. Je terminerai la discussion en montrant les multiples avantages qui découlent de l'adoption d'une théorie linguistique des contenus mentaux pour la théorie de Burge.

I

Rappelons brièvement l'expérience de pensée la plus connue. Elle porte sur des croyances au sujet de l'arthrite. Le choix du terme «arthrite» est relativement arbitraire et l'expérience pourrait donner des résultats analogues pour un très grand nombre d'expressions. Burge nous propose de considérer un agent qui entretient un certain nombre de croyances en rapport avec l'arthrite. Il croit qu'il est atteint par cette maladie qu'est l'arthrite. Il croit que son père a lui aussi des problèmes d'arthrite. Il croit qu'il vaut mieux souffrir de l'arthrite que du cancer. Mais il croit aussi par erreur avoir de l'arthrite dans la cuisse. Cette croyance est fausse puisque l'arthrite est une maladie qui n'affecte que les articulations. Burge nous invite à admettre qu'un agent puisse entretenir une telle croyance au sujet de l'arthrite même s'il est sémantiquement compétent. Cette dernière suggestion peut sembler prêter à la controverse, mais nous reviendrons sur ce point un peu plus loin. On stipule ensuite une situation contrefactuelle dans laquelle l'agent en question reste physiquement le même. Il a une même histoire biologique, une même constitution phy-

³ Loar, Brian, «Social Content and Psychological Content», dans R.H. Grimm, D.D. Merrill, (eds), *Contents of Thought*, Arizona Colloquium in Cognition, Tucson, The University of Arizona Press, 1988, 99-110.

siologique, est habité par les mêmes événements neuro-physiologiques et est responsable des mêmes mouvements corporels. La seule chose différente est que la communauté a adopté une autre convention linguistique concernant le terme «arthrite». Le terme n'est pas employé seulement pour signifier un problème au niveau articulaire, mais plutôt pour couvrir un ensemble de douleurs rhumatismales pouvant être ressenties à différents endroits sur le corps et en particulier au niveau des cuisses. La communauté en question n'utilise pas d'expression désignant spécifiquement une maladie qui n'affecte que les articulations. On peut supposer que notre notion d'arthrite, celle que nous utilisons dans le monde actuel, n'est réservée dans ce monde qu'à une élite spécialisée dans les douleurs rhumatismales.

La thèse de Burge est que, dans une telle situation contrefactuelle, l'agent en question ne pourrait entretenir les mêmes croyances. Mais puisque seules les conventions linguistiques ont été modifiées, il s'ensuit qu'elles ont une incidence réelle sur la structuration des contenus de pensée. D'un monde possible à l'autre, sa croyance qu'il a de l'arthrite dans la cuisse est devenue une croyance vraie. Par hypothèse, le monde est resté le même et le seul changement est survenu au niveau des conventions linguistiques. Puisque le terme affecté par le changement est utilisé dans les attributions de croyance, notre intuition nous incite à conclure que les croyances exprimées ne sont pas les mêmes. Force est de constater que, dans la situation contrefactuelle, l'agent n'a pas les mêmes croyances même si, par hypothèse, il est resté physiquement le même. La conclusion inévitable est alors que les états mentaux ne sont pas individués de façon individualiste. Corrélativement, on peut prétendre que les états mentaux de l'agent ne sont pas dans une relation de dépendance psycho-physique («supervenience») à ses états neuro-physiologiques.

L'expérience que nous venons de considérer est celle développée dans «Individualism and the Mental»⁴. La conclusion de l'argument m'apparaît justifiée même si, telle que je viens de la formuler, elle donne lieu à des objections qui pourraient rendre l'argument non valide. Cependant, je laisse ces objections de côté parce qu'aucune ne m'apparaît décisive.

Les autres expériences font valoir un même type d'argumentation mais à des niveaux différents. Dans «Intellectual Norms and the Foundations of Mind», Burge prétend que non seulement les règles linguistiques, mais aussi les normes régissant l'emploi des termes empiriques et exprimant les théories standards au sujet de leurs dénotations, ont un rôle structurant à jouer sur les

⁴ Burge, Tyler, «Individualism and the Mental», dans P.A. French et al (eds), *Midwest Studies Vol. 4, Studies in Metaphysics*, Minneapolis, University of Minnesota Press, 1979, 73-121.

contenus d'états mentaux⁵. Dans «Cartesian Error and the Objectivity of Perception», il prétend cette fois-ci que les objets qui s'offrent à une expérience perceptuelle déterminent eux aussi les contenus de pensée en vertu de leur propre constitution interne⁶. Enfin, dans «Other Bodies», Burge va jusqu'à prétendre que la nature intrinsèque des objets physiques, perçus ou non, peut jouer un rôle structurant analogue⁷.

II

Lorsqu'on les considère comme formant un tout, les expériences de Burge révèlent une tension entre deux points de vue fort différents. Cette tension apparaît déjà dans la thèse essentielle que les contenus d'états mentaux sont en partie déterminés par les environnements physique et social. Le fait d'insister sur ces deux types de déterminations est susceptible de nous entraîner sur des terrains philosophiques fort différents. Burge [1979] et Burge [1986] développent le thème des facteurs socialement déterminants alors que Burge [1982] et Burge [1988] analysent les facteurs physiquement déterminants de l'environnement. Dans tous les cas, l'expérience est sensiblement la même. On stipule un monde possible dans lequel tout ce qui a trait à la réalité proprement physique de l'individu est fixé. Simultanément, on imagine que des variations ont lieu dans les conventions linguistiques, les normes gouvernant l'emploi des termes empiriques, la nature des objets de perception ou même la constitution interne des objets physiques. A chaque fois, une seule conclusion s'impose. Les états psychologiques attribués à l'agent seront autres même si les formes verbales utilisées pour spécifier les contenus de croyance sont les mêmes d'un monde à l'autre.

Derrière cette apparente unité se cache pourtant une tension entre deux points de vue qui cohabitent difficilement. Lorsque les expériences sont appréciées sous l'angle du problème de l'individuation, l'expérience initiale concernant les variations dans les règles sémantiques pour des termes comme «arthrite» apparaît comme cruciale et comme l'énoncé minimal de la thèse anti-individualiste, alors que celle concernant les objets physiques tels que l'aluminium prête davantage à la controverse. L'argument de «Other Bodies», s'il est retenu, nous place en effet dans une position favorable pour prétendre que les objets physiques eux-mêmes, qui sont susceptibles de transcender les

⁵ Burge, Tyler, «Intellectual Norms and the Foundations of Mind», *Journal of Philosophy*, vol. LXXXIII, no 12, 1986, 697-720.

⁶ Burge, Tyler, «Cartesian Error and the Objectivity of Perception», in *Contents of Thought*, ibidem, 62-76.

⁷ Burge, Tyler, «Other Bodies», dans A. Woodfield, (ed), *Thought and Object, Essays on Intentionality*, Oxford, Clarendon Press, 1982, 97-120.

pouvoirs cognitifs des agents, jouent un rôle décisif dans la détermination des contenus de pensée.

Avec cette dernière expérience, on est en mesure de tirer des conclusions philosophiques beaucoup plus ambitieuses: en l'occurrence, mettre en échec le scepticisme et prouver le réalisme. Mais ces conclusions ne découlent pas de la première expérience. Il est moins problématique de prétendre, comme c'est le cas dans la première expérience, que les règles linguistiques de la communauté ont une incidence sur la structuration des états psychologiques, car elles ne transcendent pas en principe les pouvoirs cognitifs des agents sémantiquement compétents.

Mais la perspective s'inverse complètement lorsque l'on présuppose une ontologie réaliste des états mentaux ou une thèse interactionniste. Les états psychologiques apparaissent alors inscrits dans un vaste réseau de relations causales et il semble qu'il faille reconnaître au moins qu'un état psychologique différent détermine un pouvoir causal différent. Il faut en somme reconnaître une relation de dépendance à ce pouvoir causal. En outre, une modification dans la composition d'un objet physique va entraîner une modification dans le réseau complexe de relations causales dans lequel l'état mental se trouve inscrit et il devient naturel de supposer que cela entraîne un changement dans la nature même de l'état mental. Mais il est plus difficile de prétendre qu'une simple variation dans les règles sémantiques soit susceptible d'en faire autant. Il est en effet plus controversé de prétendre qu'un changement survenu seulement au niveau des stipulations sémantiques de la communauté modifie d'une quelque façon le réseau de relations causales et donc la nature des états mentaux qui s'y trouvent inscrits⁸. Dans une telle perspective, l'énoncé minimal de la thèse anti-individualiste semble être celui de «Other Bodies», et l'expérience de «Individualism and the Mental» apparaît comme une variante problématique.

Même si les deux points de vue sont possibles, il ne fait pas de doute que la thèse sur le caractère socialement déterminé des états mentaux occupe la première place et représente la version minimale de la thèse anti-individualiste. Plusieurs raisons peuvent être invoquées à l'appui de cette interprétation.

L'argument de Burge tire une bonne partie de sa force du fait de pouvoir être formulé sans présupposer une position ontologique particulière. Dans la version de «Individualism and the Mental», l'argument est, par exemple, compatible avec le rejet ou l'adoption de la dépendance psycho-physique en général. Il est même compatible avec un rejet de l'interactionnisme. Lorsqu'il

⁸ Fodor, Jerry, *Psycho-Semantics. The Problem of Meaning in the Philosophy of Mind*, Cambridge, MIT Press, A Bradford Book, 1987; voir ch. 2, 33-34.

est formulé comme impliquant la détermination de l'environnement physique sur les états mentaux, il ne manifeste pas la même indépendance. Prétendre en effet que l'environnement physique joue un rôle prépondérant dans l'individuation des états psychologiques s'accorde mal avec un rejet de l'interactionnisme. Les expériences montrant le caractère déterminant de l'environnement physique sur les états mentaux tirent toute leur force intuitive du fait de présupposer l'interactionnisme car, dans une telle perspective, il est naturel de supposer qu'un état mental se trouvant dans un rapport causal à un objet différent soit un état mental différent.

L'admission du caractère déterminant de l'environnement physique sur les états mentaux semble aussi présupposer l'existence de ceux-ci. Cela découle directement du fait de souscrire à l'interactionnisme. Burge fait intervenir explicitement l'interactionnisme comme prémissse dans son argument visant à prouver le rôle déterminant des objets de perception sur les contenus de pensée⁹. Mais qu'en est-il des objets physiques avec lesquels l'agent n'entretient aucun rapport direct? Il semble bien que Burge doive supposer que l'agent est en rapport causal avec d'autres agents dont le contenu de pensée est déterminé par l'objet physique en question¹⁰. Sinon, il serait problématique d'affirmer qu'un objet physique puisse quand même déterminer ses contenus de pensée. Mais comment peut-on admettre de telles déterminations causales sur les états mentaux sans être engagé du même coup à l'existence de ceux-ci? Le réalisme des états mentaux n'engage sans doute pas à l'interactionnisme, mais il est difficile de voir comment il serait possible d'être interactionniste sans endosser simultanément ce réalisme.

On notera aussi que l'expérience concernant le caractère déterminant de l'environnement socio-linguistique est suffisante pour montrer la nature publique des contenus d'états mentaux et pour réfuter la dépendance psycho-physique individualiste. L'argument pourrait alors être interprété comme une première étape dans la réfutation de la dépendance psycho-physique en général, et ce, sans qu'il soit nécessaire de souscrire au dualisme. Une réfutation de la dépendance psycho-physique en général est certes à la disposition de celui qui présuppose le réalisme des états mentaux, mais cela ne peut se faire qu'en souscrivant au dualisme.

Ensuite, le propos essentiel de Burge est de procéder à une critique de l'individualisme. Or, les expériences qui traitent des déterminations de l'environnement physique sur les états psychologiques affectent en plus le solipsisme méthodologique de Putnam selon lequel les contenus des états mentaux ne

⁹ Burge, Tyler, «Cartesian Error and the Objectivity of Perception», *ibidem*, 70-71.

¹⁰ Burge, Tyler, «Other Bodies», *ibidem*, 109.

présupposent l'existence de rien d'autre que celle du sujet¹¹. L'expérience de «Individualism and the Mental» reste neutre sur ce point.

Enfin, l'argument anti-individualiste de «Other Bodies» accomplit aussi beaucoup plus en ce qui a trait aux débats métaphysiques traditionnels dans la mesure où, ainsi qu'en l'a fait remarquer plus haut, il permet de mettre en échec le scepticisme et de prouver le réalisme¹².

Pour toutes ces raisons, il convient de lire l'argument de Burge dans la perspective que j'ai suggérée. Il m'est avis que telle était la perspective de Burge lui-même à l'époque de «Individualism and the Mental». Les autres expériences de pensée s'y trouvent déjà contenues en germe, mais c'est l'argument concernant le caractère déterminant des conventions linguistiques qui occupe le devant de la scène. Il ne s'agit pas de prétendre que l'expérience de «Other Bodies» est incohérente par rapport à celle proposée dans «Individualism and the Mental». Je veux seulement attirer l'attention sur le fait qu'elle met davantage de thèses en jeu et qu'elle constitue une extension plus discutable de l'expérience initiale.

De la discussion qui précède, il ressort clairement que les expériences sur les déterminations environnementales physiques font intervenir un certain nombre de présupposés philosophiques qui compliquent singulièrement l'enjeu théorique. Le réalisme des états mentaux et l'interactionnisme sont invoqués comme prémisses et le résultat des expériences de pensée a des répercussions immédiates non seulement sur la question de l'individualisme mais aussi sur des questions épineuses comme le solipsisme méthodologique de Putnam, le scepticisme et la question du réalisme. La question se pose de savoir si les expériences sur les déterminations sociales ne se suffisent pas à elles-mêmes pour réfuter l'individualisme. Il se pourrait même qu'elles permettent à elles seules une généralisation concernant l'ensemble des états mentaux. L'expérience de «Individualism and the Mental» constituerait l'énoncé minimal de la théorie anti-individualiste et ne nécessiterait pas le support d'une autre expérience.

III

Si la perspective adéquate des thèses de Burge est bel et bien celle que je propose, il est possible d'aller plus loin et de prétendre que l'argument doit être lu au départ comme une thèse portant sur le langage et non comme une thèse portant sur des états mentaux réels. Dans la mesure où Burge aspire à

¹¹ Burge, Tyler, «Other Bodies», *ibidem*, 115-116.

¹² Burge, Tyler, «Other Bodies», *ibidem*, 116-117.

une certaine neutralité à l'égard des diverses positions ontologiques, les expériences ne devraient pas présupposer la fausseté des options éliminationniste ou instrumentaliste en postulant des propriétés intentionnelles extra-linguistiques. Le problème de l'individuation devrait alors être formulé comme celui de la relation entre les concepts psychologiques et leurs exemplifications ou, encore mieux, entre les prédictats psychologiques et les entités qui les satisfont, plutôt qu'entre les propriétés intentionnelles et chacune de leurs instances.

Pour s'accorder avec l'élimination, l'instrumentalisme ou même l'anti-réalisme, l'expérience de Burge doit pouvoir être lue comme une expérience portant au sujet de nos attributions d'attitudes psychologiques¹³. Puisque l'argument tire une bonne part de sa force du fait de ne pas présupposer de doctrine ontologique particulière, il importe d'assurer cette version linguistique. Il ne faut d'ailleurs pas s'étonner que certains auteurs, comme Loar par exemple, aient vu dans l'expérience de Burge un argument concernant le langage ordinaire.

Il y a d'autres raisons pour interpréter les expériences linguistiquement. Burge prétend être relativement neutre quant à la nature des contenus d'attitudes propositionnelles. Ces derniers peuvent être assimilés à des propositions ou à des énoncés-types¹⁴. Toutefois, une étape essentielle dans l'argument repose sur le fait que, d'un monde possible à l'autre, le contenu de l'attitude est spécifié par les mêmes formes verbales («word forms»)¹⁵. Les formes verbales <arthrite>, <sofa> ou <aluminium> serviront à désigner respectivement la warthrite, les safos et du twaluminium dans les situations contrefactuelles stipulées. La notion de forme verbale permet d'illustrer ce qui reste inchangé d'un monde possible à l'autre dans la spécification du contenu de l'attitude. Les contenus sont spécifiés par les mêmes formes verbales, mais celles-ci ont des valeurs sémantiques distinctes et déterminent par conséquent des contenus mentaux distincts. Il est assez difficile d'écartier l'idée que les formes verbales interviennent elles-mêmes dans les contenus mentaux. La seule possibilité est d'interpréter l'argument comme portant sur des attributions d'attitudes et non sur les attitudes elles-mêmes. Les formes verbales seraient alors ce que nous utilisons pour spécifier les contenus d'attitudes et non ce que sont ces contenus en eux-mêmes.

¹³ Burge distingue «mental attributions» et «mental ascriptions». C'est cette dernière expression que je traduis par «attribution d'attitudes psychologiques». Seule la première implique l'existence de propriétés mentales. Voir Burge, Tyler, «Individualism and the Mental», *ibidem*, 75.

¹⁴ Burge, Tyler, «Individualism and the Mental», *ibidem*, 74.

¹⁵ Burge, Tyler, «Other Bodies», *ibidem*, 101; «Intellectual Norms and the Foundations of Mind», *ibidem*, 708.

On aurait tort de prétendre, comme le fait Lynn Rudder Baker, que la prémissé requise est seulement que nous pensions parfois à l'aide de mots¹⁶. Avec une telle prémissé, la conclusion de l'expérience de Burge serait seulement que, parfois, les contenus d'états mentaux ne sont pas individués de façon individualiste. Burge prétend être en mesure de tirer des conclusions générales concernant les croyances qui font intervenir obliquement des notions comme l'aluminium ou l'arthrite. Ses remarques ne se limitent pas au cas où l'agent accède à de telles croyances en acquiesçant à une phrase particulière. Certes, il n'est pas encore entièrement parvenu à prouver sa thèse anti-individualiste pour toutes les sortes de croyance. Il lui resterait, par exemple, à montrer que les croyances portant sur des vérités mathématiques, des vérités logiques ou des sensations internes telles que la douleur ont elles aussi un caractère public. Mais les autres sortes de croyance tombent toutes sous la portée de l'argument et elles ont par conséquent un caractère essentiellement public.

J'en conclus que pour préserver, ainsi que Burge le souhaite, la neutralité quant à la nature des contenus d'états mentaux, l'expérience doit être interprétée au départ comme portant sur des attributions d'attitudes et non sur les attitudes elles-mêmes. Je m'empresse d'ajouter qu'il s'agit ici d'un point d'exégèse. La neutralité concernant la nature des contenus de pensée n'est pas recommandable, bien au contraire. Comme on le verra plus loin, l'adoption d'une théorie linguistique des contenus de pensée s'avère hautement recommandée. Mon point est seulement que Burge est lui-même forcé de formuler sa théorie en des termes qui ont une pertinence pour les attributions d'états psychologiques, étant donné sa prétention à la neutralité au sujet des contenus de pensée.

IV

Jusqu'ici, j'ai prétendu peu de choses. J'ai seulement voulu montrer que, dans une perspective d'économie ontologique et dans l'hypothèse où la doctrine visée est l'individualisme, la première expérience de pensée est centrale. Vient ensuite celle de 1986. Les autres font intervenir des présupposés additionnels qui remettent en question d'autres doctrines que l'individualisme. J'ai montré ensuite qu'il fallait donner une formulation linguistique à l'expérience. La thèse de Burge doit pouvoir être formulée comme une thèse portant sur les attributions d'attitudes pour maintenir les exigences d'économie ontologique. Mais, ce faisant, l'argument prête flanc à l'objection de Loar et

¹⁶ Rudder Baker, Lynne, *Saving Belief. A Critique of Physicalism*, Princeton, Princeton University Press, 1987, 29.

on ne voit plus clairement comment la théorie peut avoir une quelconque incidence sur la nature des attitudes propositionnelles elles-mêmes.

C'est ici qu'une théorie citationnelle des contenus de pensée permet de venir à la rescousse de Burge. Il s'est déclaré neutre quant à la nature des contenus de pensée, mais l'expérience de «Individualism and the Mental» appelle très naturellement une théorie citationnelle. La théorie suppose que les contenus de pensée sont au moins en partie constitués d'items appartenant à la langue publique. Elle peut être mise à contribution pour répondre à certaines objections. Un premier avantage est de permettre le transfert des traits caractéristiques gouvernant les attributions psychologiques aux états psychologiques eux-mêmes. En particulier, les expériences qui concernent la détermination du contenu des attributions d'états mentaux par l'environnement social auront désormais une répercussion immédiate sur le contenu des attitudes elles-mêmes. Lorsqu'une théorie citationnelle vient s'ajouter à l'expérience initiale, on est en mesure de montrer que, si les états mentaux intentionnels existent, ils ne sont pas individués de façon individualiste.

Il est vrai que le saut du langage intentionnel aux états mentaux intentionnels ne peut se faire automatiquement. Il faut tout d'abord invoquer une théorie linguistique des contenus mentaux, ainsi que je viens de le suggérer. Mais il faut aussi tenir compte du fait que les contenus attribués sont susceptibles de transcender les capacités cognitives des agents. Cela se produit, par exemple, lorsque le prédicat de croyance attribué à un agent contient, au niveau de la subordonnée, des déictiques qui font intervenir la perspective de celui qui fait l'attribution. Cela se produit aussi lorsqu'on applique des principes de substitution ou d'équivalence sémantique que l'agent est susceptible d'ignorer, ou encore lorsqu'on fait valoir certaines conséquences logiques obtenues à partir de ses autres croyances.

Plusieurs de ces cas peuvent être considérés comme des cas d'attribution «pseudo *de re*», pour employer une expression de Kaplan¹⁷, mais d'autres pourraient tout aussi bien être considérés comme des cas d'attribution «pseudo *de dicto*». C'est notamment ce qui se produit lorsqu'on attribue à un agent certaines des conséquences logiques de ces croyances. Qu'il apparaisse sur le mode *de re* ou *de dicto*, cet usage correspond à ce que j'ai nommé ailleurs l'attribution matérielle d'attitudes psychologiques, par opposition à l'attribution intentionnelle qui, elle, reproduit fidèlement le contenu de l'attitude. Mais je ne suis pas disposé à adopter la terminologie péjorative de Kaplan parce que j'estime que l'usage matériel est primitif et que ce que nous appelons l'usage intentionnel peut être considéré comme une espèce particulière d'usage matériel.

¹⁷ Kaplan, David, *Demonstratives*, UCLA, 1977, Unpublished Manuscript.

Plusieurs autres facteurs de moindre importance peuvent entrer en ligne de compte pour empêcher que les propriétés des attributions intentionnelles puissent être transférées aux états mentaux intentionnels eux-mêmes. Il y a, par exemple, les cas où l'énoncé d'attitude est utilisé pour performer autre chose qu'une assertion. Mais une fois que ces obstacles ont été écartés, il n'y a pas de raison d'interdire une application des remarques de Burge aux états mentaux eux-mêmes. L'intérêt de procéder de la sorte, en commençant par se prononcer sur les attributions d'états mentaux pour ensuite tirer des conclusions au sujet des états mentaux eux-mêmes via une théorie citationnelle, est que l'on peut de cette manière défendre un point de vue anti-individualiste tout en restant ontologiquement neutre. On s'engage certes à un point de vue particulier concernant la nature des contenus d'états mentaux, mais on reste neutre sur le problème crucial de l'identité des états mentaux eux-mêmes.

Supposons qu'un agent attribue à Alfred la croyance qu'il a de l'arthrite dans la cuisse. Pour que l'expérience de Burge puisse être appliquée à l'état mental d'Alfred lui-même, il faut que l'énoncé d'attitude soit utilisé de façon assertorique. Il faut en plus supposer un usage intentionnel du verbe d'attitude propositionnelle et non un usage matériel. Il faut aussi que les déictiques apparaissant dans la subordonnée aient été contextuellement éliminés. Supposons enfin avec Burge que les expressions ont une occurrence oblique et servent à déterminer une attribution *de dicto*. Cette contrainte importe même une fois que les usages «pseudo *de re*» ont été écartés et qu'on s'en tient à des usages intentionnels. Sinon, la conclusion de l'expérience pourrait être rendue triviale.

Une attribution de croyance intentionnelle *de re* au sujet de l'aluminium deviendrait en effet trivialement autre s'il était stipulé que la constitution interne du métal change, et ce, pour la raison suivante. Dans une attribution de croyance, les expressions ne servent jamais à spécifier un mode de présentation privé et, conséquemment, ne fonctionnent pas comme des expressions d'un langage privé. Cela est vrai aussi pour les rapports de croyance intentionnelle *de re*. Même si un rapport intentionnel *de re* implique l'existence d'une expérience perceptuelle au sujet de l'aluminium, la forme verbale <aluminium> ne sert pas à référer à cette expérience particulière et sert plutôt, là comme ailleurs, à désigner l'aluminium. Certes, puisque l'attribution est intentionnelle, la forme verbale <aluminium> peut intervenir en partie elle-même comme constituant du mode de présentation que se fait l'agent, et c'est la raison pour laquelle la substitution d'une expression co-désignative est interdite, mais le référent du terme est toujours la substance comme telle. Or la particularité d'un rapport de croyance *de re* est que la substance intervient elle-même dans le rapport. Elle n'intervient pas à titre de constituant proposi-

tionnel, mais son existence est présupposée ainsi qu'une spécification de sa nature. Il ne fait pas sens d'attribuer une croyance *de re* au sujet de l'aluminium et de nier simultanément que l'aluminium existe. On échoue notre attribution de croyance *de re* si la substance n'existe pas. D'un point de vue logico-sémantique, il faudrait dire que les rapports intentionnels *de re* sont référentiellement opaques (échec du principe de substitution) et quantificationnellement transparents (application de la généralisation existentielle). Et puisque l'existence du métal est présupposée, on présuppose également une spécification de sa nature. On réalise donc une attribution différente de croyance intentionnelle *de re* si la nature du métal est transformée, que cela soit en vertu d'une stipulation ou d'un progrès dans nos connaissances. Voilà pourquoi un changement dans la constitution interne du métal entraîne trivialement une attribution intentionnelle *de re* différente. Pour que nous soyions en mesure de transférer les conclusions de l'expérience au niveau des contenus d'états mentaux de l'agent, il faut donc que l'attribution intentionnelle soit *de dicto* et non *de re*.

Supposons que toutes ces conditions soient satisfaites. En ce cas, il ne nous reste plus qu'à adopter l'hypothèse linguistique concernant les contenus d'attitudes pour justifier la conclusion que les états mentaux eux-mêmes sont individués en relation avec l'environnement social.

L'expérience nous a déjà convaincu que des changements apparaissant au niveau des conventions linguistiques affectent de façon significative les attributions de croyance à un individu y compris lorsque les descriptions physiques de cet individu sont restées les mêmes. En particulier, cela se produit lorsqu'une nouvelle convention linguistique survient et affecte l'emploi de la forme verbale <arthrite>. Puisque, par hypothèse, Alfred est lui-même mis en rapport avec la forme verbale <arthrite> et que cette dernière constitue une partie de la substance même de son contenu de pensée, il faut conclure qu'Alfred n'a pas les mêmes contenus de pensée. Et puisque nos descriptions de ses états physiques sont restées les mêmes, on doit conclure aussi que les états mentaux particuliers d'Alfred ne sont pas individués de façon individuliste.

V

Je voudrais montrer maintenant que l'adoption d'une théorie citationnelle des contenus de pensée entraîne plusieurs changements et améliorations à la théorie de Burge. Le premier est que l'expérience de «Individualism and the Mental» se suffit désormais à elle-même pour montrer que l'anti-individuisme est vrai en général. Il n'est plus nécessaire de faire intervenir des expé-

riences de pensée plus compromettantes du point de vue ontologique comme celle de «Other Bodies». Il n'est plus nécessaire de se demander si des variations dans l'environnement physique ont un impact sur la structuration des contenus de pensée. La question peut désormais rester ouverte.

Il suffit de reconnaître que des expressions telles que «aluminium», «sofa», «eau», etc., ont une signification linguistique ou un caractère au sens de Kaplan. Dans la plupart des cas, le caractère fait intervenir une spécification de la catégorie d'objets désignés par l'expression et du type d'entités auquel elle s'applique. En outre, la compréhension du caractère doit, pour se qualifier comme compétence sémantique, être accompagnée d'une capacité à identifier des instances du concept exprimé par le terme. Il peut être difficile d'établir, pour chaque expression, ce qui en constitue le caractère, mais ce n'est pas une raison pour douter de son existence. La difficulté s'explique par le fait que les significations n'ont pas de réalité objective en dehors des stipulations de la communauté. D'un groupe à l'autre, la signification d'une expression peut changer et c'est pourquoi il n'est pas facile de se prononcer sur la signification de tel ou tel mot particulier.

Une fois que les significations linguistiques ont été admises, il est facile de reproduire l'expérience de «Individualism and the Mental» à toutes les expressions du langage, y compris celles qui déterminent des espèces naturelles¹⁸. On peut imaginer un agent qui entretient un ensemble de croyances au sujet de l'aluminium. Il croit que le mât du bateau de son oncle est en aluminium, que l'on doit souvent recouvrir les aliments de papier d'aluminium, etc. Toutes ces croyances sont vraies. En outre, il fait reposer la tâche de spécifier le concept d'aluminium sur la communauté scientifique à laquelle il appartient et se rapporte, le cas échéant, à un dictionnaire, pourvu que ce dernier enregistre justement le concept en vigueur dans la communauté scientifique. Enfin, il est aussi en mesure d'identifier correctement des échantillons d'aluminium. Bref, il se comporte comme un locuteur normal du français. Mais il croit aussi par erreur que l'aluminium n'est pas un métal. Supposons que cela aille à l'encontre de la pratique des locuteurs et que ceux-ci stipulent plutôt que la notion de métal intervient dans la signification linguistique du terme. En somme, bien qu'il soit un locuteur français compétent et que sa croyance soit intentionnelle, l'individu en question manifeste une maîtrise imparfaite du mot «aluminium» et ne saisit pas correctement sa signification linguistique. On n'a ensuite qu'à imaginer une situation contrefactuelle dans laquelle la même forme verbale, bien qu'extensionnellement équivalente, ne servirait pas à désigner un métal. Cela ne pourrait se produire que dans une

¹⁸ Burge, Tyler, «Individualism and the Mental», *ibidem*, 79.

communauté qui est dans un état d'ignorance relative quant à la nature des métaux, mais nous n'avons pas à nous prononcer sur les raisons qui expliquent pourquoi un concept est en vigueur à un moment donné dans une communauté. La conclusion est que, dans une telle situation contrefactuelle, l'agent a des croyances intentionnelles très différentes.

Mais qu'en est-il de l'expérience de Burge elle-même telle qu'elle est formulée dans «Other Bodies»? Pour répondre à cette question, il faut stipuler un monde possible dans lequel on ne trouve pas d'aluminium comme tel, mais seulement un matériau qui a toutes les apparences de ce métal. Dans ce monde, la forme verbale <aluminium> se voit toujours associée la même signification linguistique. Il s'agit, pour les membres de la communauté vivant dans ce monde, d'un métal léger ayant les propriétés phénoménales habituelles, et qu'on parvient à identifier de la même manière qu'on parvient à identifier l'aluminium dans notre monde. Dans un tel monde, un agent entretiendrait-il des croyances différentes quant à l'aluminium? Il me semble que non.

Il ne faut pas oublier que pour reproduire intégralement l'expérience de pensée, l'agent doit être resté physiquement le même. En particulier, il doit avoir les mêmes dispositions à acquiescer ou s'opposer aux énoncés contenant la forme verbale <aluminium>, y compris celles qui servent à spécifier sa signification linguistique. Puisque les expressions qui servent à spécifier sa signification ne sont pas affectées par un changement de signification, la signification de la forme verbale <aluminium> reste la même. Burge n'a donc pas raison, dans ce cas particulier, de prétendre que les croyances de l'agent seraient différentes. C'est du moins la conclusion qu'on doit tirer si l'on s'en tient à une lecture intentionnelle des rapports de croyance. Cela ne veut pas dire que l'expérience de «Other Bodies» soit dépourvue de tout caractère intuitif, mais la force de l'argument s'explique peut-être par le fait que l'on tend à faire une lecture matérielle des rapports de croyance. Dans la situation contrefactuelle où un métal semblable à l'aluminium (le twaluminium) a remplacé celui que nous désignons par la même forme verbale, Alfred croit en fait que le twaluminium n'est pas un métal. Si sa croyance *de dicto* est rapportée matériellement, elle ne sera pas la même d'un monde possible à l'autre. Dans un rapport de croyance matérielle, l'agent est mis en relation avec les conditions de vérité de l'énoncé. Or la forme verbale <L'aluminium est un métal> aura des conditions de vérité différentes si le terme <aluminium> ne sert plus à désigner l'aluminium, mais plutôt le twaluminium.

Il ne faut pas se méprendre sur la conclusion à tirer. En distinguant le caractère et le contenu, je ne cherche pas à séparer radicalement les significations linguistiques des théories scientifiques, mais plutôt d'une soi-disant réa-

lité transcendance. Plus exactement, j'estime qu'il n'est pas nécessaire d'invoquer la réalité transcendance pour expliquer le caractère public des significations linguistiques et, par voie de conséquence, des états mentaux. Si un changement dans l'environnement physique peut se traduire par un changement dans le contenu des états mentaux, ce n'est pas à cause de la détermination d'une réalité transcendance sur les contenus de pensée, mais plutôt parce qu'une communauté choisit d'ajuster son vocabulaire sur celui des spécialistes et que ceux-ci ont pris connaissance de ce changement.

VI

Dans cette dernière section, je voudrais considérer d'autres avantages, pour la théorie de Burge, de l'adoption d'une caractérisation linguistique des contenus de pensée. On pourrait vouloir s'en prendre à la prémissse de l'argument selon laquelle l'agent croit qu'il a de l'arthrite dans la cuisse. Burge doit supposer, dans de tels cas, que l'agent a une maîtrise imparfaite des concepts qu'il utilise, et cela peut paraître problématique¹⁹. Le problème ne se pose plus si l'on précise que l'agent accède à ces «concepts» via des expressions linguistiques. La maîtrise imparfaite des concepts équivaut à une maîtrise imparfaite des expressions. C'est là un fait banal et courant, y compris pour des locuteurs sémantiquement compétents. On a donc tort de considérer la première prémissse de l'argument comme problématique. La prémissse problématique serait de stipuler qu'un agent sémantiquement compétent puisse, de façon systématique, utiliser incorrectement les expressions de son propre langage.

On peut aussi faire valoir que l'anti-individualisme de Burge s'accorde mal avec le principe cartésien selon lequel un agent sait toujours ce à quoi il est en train de penser. Ce principe serait bien entendu faux si le verbe «penser» était utilisé au sens matériel. Mais il semble bien qu'il y ait intuitivement une bonne part de vérité dans le principe, et que celui-ci soit vrai lorsque l'on fait une lecture intentionnelle du verbe. Le principe pose une difficulté réelle à Burge surtout à cause des expériences qui visent à montrer la détermination par l'environnement physique des contenus d'états mentaux. Si une différence dans la nature d'un objet transcendant détermine une différence dans le contenu de l'état mental qui lui correspond, comment faire pour éviter de postuler des contenus de pensée qui transcendent les capacités cognitives des agents?

¹⁹ Bach, Kent, *Thought and Reference*, Oxford, Clarendon press, 1988, ch. 13; «Burge's New Thought Experiment: Back to the Drawing Room», *Journal of Philosophy*, vol. LXXXV, no 2, 1988, 88-97.

Burge n'est pas engagé à dire que les objets physiques eux-mêmes font partie intégrante des contenus d'états mentaux, mais il ne peut pas exploiter aussi facilement l'option qui paraît s'offrir à lui et qui consiste à postuler des concepts frégéens pour rendre compte du principe de Descartes. Les concepts frégéens sont les valeurs cognitives associées aux expressions. Ils peuvent varier d'un locuteur à l'autre et il peut sembler au départ que cela permette de régler la question. Un agent peut en effet difficilement ignorer quels concepts il associe à telle ou telle expression et sait donc ce à quoi il est en train de penser. Mais, dans les expériences, on stipule au départ que l'agent reste physiquement identique d'un monde possible à l'autre et cela implique qu'il est responsable des mêmes comportements. Il doit donc avoir les mêmes dispositions à acquiescer aux mêmes formes verbales. En particulier, il acquiescera ou s'opposera de la même manière aux formes verbales contenant la forme <aluminium>. En outre, puisqu'il reste physiquement le même, il doit être atteint par les mêmes stimulations sensorielles. Il peut donc difficilement se forger un concept différent. Il a les mêmes croyances générales et il associe aussi une même signification linguistique à la forme <aluminium>. On voit mal alors comment une différence conceptuelle pourrait surgir en de telles circonstances.

Dans un monde possible où les règles sémantiques sont modifiées, les agents sont amenés à affirmer des énoncés différents et les modifications sémantiques vont alors de pair avec certaines modifications au niveau de la réalité physique. Mais l'idée ici est que tout cela pourrait se produire sans que ne survienne une modification dans la constitution physique de l'individu Alfred. Cela est rendu possible par le fait qu'Alfred peut avoir une maîtrise imparfaite de certaines expressions du langage tout en étant sémantiquement compétent. La définition idiosyncratique et erronée qu'il a du terme <aluminium> devra alors être rendue dans un vocabulaire pleinement maîtrisé. Alfred reste donc physiquement le même, mais ses croyances sont désormais très différentes. Il en est ainsi à tout le moins pour ce qui est de celles que l'on spécifie en utilisant le terme <aluminium>, puisqu'il s'agit d'une expression ayant une tout autre signification. J'accepte le résultat de cette expérience, mais je mets en doute le résultat que Burge croit être en mesure d'obtenir lorsqu'il ne fait varier que l'environnement physique et laisse intactes les significations linguistiques.

Burge peut répondre que, dans la situation contrefactuelle, l'agent entre en contact avec le nouveau métal par l'intermédiaire des contenus de perception. Ces derniers sont des représentations qui spécifient le twaluminium comme tel et non des contenus représentationnels comme des sense data, par exemple. L'idée est que l'on ne peut pas avoir une expérience visuelle de l'alu-

minium dans un environnement où l'aluminium n'existe pas. Burge a défendu cette thèse dans un article récent²⁰. Bien qu'il s'agisse d'une question controversée, je suis prêt à accepter la thèse pour les fins de la discussion. Le problème est que, dans ce cas, il est difficile d'admettre les prémisses de départ de l'expérience tout en préservant le principe cartésien. On ne voit pas comment Burge pourrait prétendre que, dans l'expérience, Alfred sait qu'il est en train de penser à du twaluminium. De deux choses l'une: ou bien une différence dans la réalité transcendante entraîne une différence dans les contenus de pensée, et alors Alfred ne sait pas ce à quoi il est en train de penser, ou bien Alfred le sait, mais alors on ne voit plus comment préserver la prémissse qu'Alfred a les mêmes dispositions à acquiescer ou s'opposer aux énoncés contenant la forme verbale <aluminium>.

La solution consiste à reconnaître qu'Alfred a bel et bien les mêmes contenus de pensée d'un monde possible à l'autre. Il en est ainsi parce qu'Alfred associe une même signification aux expressions linguistiques et que celles-ci forment la substance de ses contenus de pensée. Le principe cartésien est satisfait parce qu'Alfred est un locuteur sémantiquement compétent. On se rappelle le principe néo-cartésien de Kaplan selon lequel si je pense la théorie des démonstratifs, alors je suis²¹. Un principe analogue peut être invoqué ici. Celui-ci énonce que si je suis un locuteur sémantiquement compétent, alors je sais ce à quoi je suis en train de penser.

Il est vrai, de cette manière, l'expérience de «Other Bodies» doit être rejetée. Plus exactement, répétons-le, l'expérience doit être rejetée lorsque les verbes épistémiques sont utilisés intentionnellement. La conclusion de l'expérience est acceptable lorsque les verbes épistémiques sont utilisés matériellement. L'expérience de «Individualism and the Metal» reste vraie puisqu'elle repose sur des variations concernant les conventions linguistiques. Une théorie citationnelle des contenus de pensée non seulement s'accorde avec le résultat de cette expérience, mais en fournit en plus l'explication.

Une autre conséquence est que l'introduction de concepts frégéens n'est plus nécessaire. En essayant de montrer que les objets physiques eux-mêmes déterminent les contenus de pensée, Burge ne veut pas faire la preuve de l'existence des propositions singulières et suggérer qu'une proposition concernant le Mont Blanc a, comme l'un de ses constituants, la montagne elle-même, intégralement, avec tous ses cristaux de neige. Les concepts frégéens sont les véritables constituants propositionnels et sont censés jouer le rôle d'intermédiaire entre l'agent et les choses elles-mêmes. Mais l'introduc-

²⁰ Burge, Tyler, «Cartesian Error and the Objectivity of Perception», *ibidem*, 69.

²¹ Kaplan, David, *op. cit.*

tion de ces entités fait problème. La difficulté n'est pas seulement que les concepts frégéens entraînent une surcharge ontologique, ont une identité obscure et ne nous sont pas très utiles, comme on l'a vu, pour rendre les expériences compatibles avec le principe cartésien. Le problème le plus grave, à mon avis, est qu'ils finissent par rendre imperceptible sinon inexistante la frontière entre le contenu pragmatique et le contenu se mantique des expressions. En admettant, en lieu et place, des contenus linguistiques, on s'épargne de telles difficultés et on se donne des entités plus faciles à apprivoiser.

Il peut sembler à première vue qu'une tension persiste entre l'expérience de Burge ainsi reconstruite et le principe cartésien, et ce, même une fois que les contenus de pensée ont été déclarés linguistiques. En effet, l'expérience de «Individualism and the Mental» ne requiert-elle pas que l'agent viole le principe? Dans la prémissse initiale, Alfred se trompe sur la signification du mot arthrite (ou aluminium). Si les contenus sont linguistiques et qu'il ignore la signification du mot dans sa communauté, comment peut-il savoir ce à quoi il est en train de penser? La réponse est qu'effectivement le principe est alors violé. Le principe cartésien est présupposé aussi longtemps que la compétence sémantique est elle-même présupposée. Lorsque la pensée d'un agent est articulée à l'aide d'un vocabulaire qui n'est pas totalement maîtrisé, il est normal de conclure qu'il ne sait pas exactement ce à quoi il est en train de penser. Même si sa croyance est intentionnelle et qu'il est un locuteur compétent du français, il peut utiliser incorrectement certaines des expressions de son propre langage.

L'erreur linguistique ne constitue pas un problème pour une théorie linguistique des contenus de pensée, bien au contraire. Notre point de vue général est que les règles linguistiques en vigueur déterminent les contenus de pensée et non l'inverse. Si les contenus de pensée étaient antérieurs, il n'y aurait pas vraiment place pour l'erreur linguistique. Tout écart par rapport aux usages en vigueur pourrait équivaloir à la création d'un idiolecte. A vrai dire, il pourrait ne pas vraiment y avoir de langage, ni de règles linguistiques, parce qu'il n'y aurait pas de distinction entre «suivre une règle» et «croire que l'on suit une règle». Pour celui qui pense à l'inverse que les contenus de pensée sont déterminés par le langage, les règles sémantiques jouent un rôle déterminant et c'est seulement à cette condition qu'il peut y avoir des erreurs linguistiques.

L'admission d'erreurs linguistiques ne constitue pas un problème pour la théorie que je propose, et cela revient à autoriser des définitions idiosyncratiques de la part des agents. Mais de telles définitions ne sont possibles que si les expressions utilisées dans le definiens sont utilisées de façon compétente par lui. Il peut y avoir un écart entre la pensée de l'agent et les règles linguistiques

en vigueur comme cela arrive lorsqu'on commet une erreur, mais cela n'est possible que sur l'arrière-plan d'un accord systématique avec ces règles. Et c'est parce qu'il y a une telle conformité avec les règles qu'on est en mesure de souscrire au principe cartésien. Même lorsque l'agent a un contenu de pensée idiosyncratique qui implique un écart entre la signification qu'il associe à un mot et sa signification admise par la communauté, son contenu de pensée n'aura de sens que s'il peut être exprimé à l'aide du vocabulaire admis par la communauté. Voilà pourquoi la violation du principe cartésien n'était vraie que localement. Lorsque la compétence sémantique est systématiquement présupposée chez un agent, ce dernier sait toujours ce à quoi il est en train de penser.